Biosynthetic regulation of the major opiates in *Papaver somniferum*

Fergus Meade

Doctor of Philosophy

University of York

Biology

September 2015

Abstract

Opium poppy, *Papaver somniferum*, is the sole source of the analgesic alkaloids morphine and codeine as well as thebaine, a precursor for semi-synthetic opiates. *T6ODM* (thebaine 6-O-demethylase) and *CODM* (codeine O-demethylase) are dioxygenases involved in morphine biosynthesis and represent promising targets for metabolic engineering of the morphinan alkaloid pathway through reverse genetic screening. An EMS (ethyl methanesulfonate)-mutagenised population of a morphine accumulating cultivar (>4000 plants) was screened for mutations in *CODM* and *T6ODM*. Although nonsense mutations were found in both, complete metabolic blocks and codeine and thebaine were not observed owing to the presence of multiple copies of these genes in the genome. Crosses and further mutagenesis were attempted to produce new cultivars of opium poppy with increased yields of codeine and thebaine.

Table of Contents

List of Figures

List of Tables

Acknowledgements

First and foremost, I would like to thank my supervisors Ian Graham and Carol Walker. Ian has always been a fountain of knowledge and a source of encouragement while always helping to focus on the bigger picture. Due to the time difference, Carol has endured many late night telecons over the years, but her commitment and passion for the project was contagious. Thank you for being such a great host during my time in Tasmania and for teaching me so much about the industry and the island in general. I would also like to reserve a special thanks to Thilo Winzer for always being so approachable and keen to discuss the finer details of the project. I am really grateful for the all the constructive discussions we have had, particularly in these past few months during the writing process.

I would to thank both the BBSRC and GSK Australia for the funding to carry out this research. I would to thank all the staff I met at GSK in Latrobe, Tasmania for their help with the field trials and also the staff at Port Fairy, Victoria for carrying out phenotypic analysis. A special mention should also go to Tim Bowser. Thank you for allowing me the opportunity to pursue this research and I really hope the company can use some of the outcomes in future.

I also would like to thank my TAP members- Louise Jones, Rob Edwards and Seth Davis- who contributed to great discussions at meetings and brought new perspectives.

Thanks to the rest of the poppy group, past and present, including Roxana Teodor, Sue Heywood, Samantha Donninger, Marta Mendes, Marc Cabry, Valeria Gazda, Dave Harvey, Tony Larson, Yi Li, Marcelo Kern and Linda Sainty. It has been a pleasure to work with you all and I have enjoyed discussing the finer details of our efforts to understand and manipulate this intriguing plant. Thanks also to Judith Mitchell for making sure everything ran smoothly.

Finally, thanks to my wife Chitra, for your support, patience and understanding, and my boys, Ciarán and Inesh, who both came into this world during the lifetime of this project and kept things interesting!

Author's declaration

I declare that the work submitted in this thesis is my own unaided work, unless otherwise acknowledged in the text. No part of this thesis has been previously submitted for any other degree, at this or any other university.

Chapter 1- Introduction

1.1 The opium poppy, Papaver somniferum L.

Papaver somniferum is an annual herb of the *Papaveraceae* family belonging to the basal eudicots (Linné and Salvius, 1753). Mature plants can reach heights of up to 1.5 m. Lengths of the semi-erect leaves shorten towards the top of the stem which is solid and may have hairs. The four petals of the flower vary in colour from red to white and in Tasmanian varieties flowers are a pale pink with a darker purple dot at their base. A multi-rayed stigma occupies the centre of the flower and multiple stamens surround it. The fruit is a porous capsule. After fertilisation, the petals drop off and the seed capsule swells and takes on a distinctive shape i.e. spherical with a star-shaped flattened top. The top lifts up as the capsule dries out to aid in dispersal of seed. The entire growth cycle takes approximately 120 days.

Figure 1. Opium poppies growing in the field in Tasmania. Plants are at different stages of development. These include the characteristic hook stage prior to flowering. The pale pink petals of flowering plants are also seen with a darker purple dots towards the base. When these petals fall the seed capsule swells and the top flattens. Photo credit: Carol Walker

Papaver somniferum has a chromosome number of eleven (Kaul et al., 1979; Lavania and Srivastava, 1999; Wakhlu and Bajwa, 1987). Its genome has an estimated size of 3724 Mb (Bennett and Smith, 1976) but it has yet to be sequenced. Plants are mainly self-pollinating but are able to outcross to generate variation (Bhandari, 1990; Kumar and Patra, 2010; Patra et al., 1992).

Opium poppy produces latex in highly specialised cells called laticifers. The latex contains a range of benzylisoquinolines alkaloids (BIAs) many of which have analgesic and sleep-inducing properties. Morphine is perhaps the most well known example of a BIA from opium poppy. These characteristics led Carl Linnaeus to name the plant *somniferum*, which means sleep-inducing in Latin. The fact that there are no real wild opium poppy varieties points to the close interaction between the plant and humans over the last few millennia. It is still one of the most important medicinal plants in use today.

1.2 Historical importance of opium poppy

Opium poppy has been used for the medicinal properties conferred by some of its alkaloids for many thousands of years. Sumerian artefacts from 4000BC bear images of opium poppies. It was known as *hul gil*, the plant of joy (Brownstein, 1993). Opium was widely used by the Minoans for cult rituals or healing purposes since at least the 5th century BC (Askitopoulou et al., 2002). Images of poppies are depicted in Egyptian pictography and Roman sculptures. Egyptian priests promoted the use of opium preparations called *thebacium*, named after the city of Thebes where the poppies were cultivated (Tibi, 2003). Seeds were widely used as food before the Greeks recognised the sleep-inducing effects of the capsules and latex. Opium comes from the Greek word *opus* meaning latex. By the 8th century AD, opium use had spread to Arabia, India and China and the Arabs had organised its trade. Ignition and smoking of opium to obtain a more powerful medicinal effect began in Islamic culture. The Chinese also smoked opium, often in combination with tobacco. Indians ordinarily ingested opium. The pain relieving properties of opium in surgery were recognised by the Greeks and Romans, while Arab culture developed the preparations still further, frequently using opium for the treatment of diarrhoea (Aragon-Poce et al., 2002), coughs and colds and ocular remedies (Tibi, 2003)

Medicinal use of opium became more widespread in the 16th century after the Swiss-German alchemist Paracelsus discovered that the alkaloids in opium are more soluble in alcohol than in water. A tincture of opium or laudanum had use in pain relief (van Ree et al.,

1999). Various preparations of laudanum were later used in combination with everything from saffron and sugar to mercury and whiskey for the treatment of practically every ailment known to man. By the 18th century the British East India Company controlled prime poppy growing areas of India and dominated the trade controlling supply and setting prices. Importation and recreational use of opium in China was banned by the Imperial Chinese Court, but efforts to enforce this were stepped up in 1839 by the Emperor Tao Kwang. This led to the First Opium War, when the British Empire responded with force. Defeated China was forced to sign the Treaty of Nanjing in 1842 allowing the opium trade to expand with the opening of five new ports. Hong Kong was also ceded to the British. In 1856 the Second Opium War broke out over demands that the trade be expanded even further. Again, China was defeated and forced to sign the Treaty of Tientsin in 1858, formally legalising opium imports to China. By the end of the century, it is estimated that over a quarter of all Chinese males were addicted to opium. These events in history point to the potential abuse of the plant. Nevertheless, use of the plant for its medicinal properties also flourished at this time and efforts were being made to identify the active ingredients.

Although laudanum use continued to flourish, there were often problems with purity and its precise content in active ingredients could not be measured. A milestone in pharmacy was reached when morphine was first isolated by Friedrich Sertürner (Sertürner, 1806). He named it after Morpheus, the Greek god of sleep. Morphine was the first ever medicinal alkaloid isolated in pure form from a plant (Luch, 2009). This is testimony to the medicinal importance of opium poppy in the history of mankind. The discovery ushered in modern pharmacology, because for the first time it was possible to precisely measure the dose-effect relationship of a drug. Furthermore, the pharmaceutical industry developed as some pharmacists realised specialising in the extraction and supply of specific alkaloids to other pharmacists could be a profitable enterprise. Merck started marketing morphine commercially in 1827. In the wake of the isolation of morphine, additional alkaloids with medicinal applications were isolated from the opium poppy plant, including the analgesic and antitussive codeine (Robiquet, 1832), the non-opiate alkaloid noscapine (Robiquet, 1817)

and the smooth muscle relaxant papaverine (Merck, 1848). By the 1850s pure alkaloids rather than crude opium preparations were being used for the relief of pain, cough and diarrhoea. Use expanded after the invention of the hypodermic syringe in 1857.

Besides the use of its alkaloids for pharmaceutical applications, its seeds are a source for poppy seed oil, which has established uses as cooking oil, lamp oil and varnish and is also used to make paints and soaps. More recently, pharmaceutical and medicinal diagnostic applications as an adjuvant add to its importance as a high quality edible oil (Krist et al., 2005). Poppy seeds are also used by the bakery and confectionery industries.

However, the primary reason opium poppy is cultivated today is for its alkaloids. It is the sole commercial source of the analgesics morphine and codeine. It is also the only source for thebaine and oripavine, which are used as precursors for a range of semisynthetic opiates (Berenyi et al., 2009).

Morphine is easily converted to O,O-diacetylmorphine, otherwise known as heroin, which is more potent and is used as a recreational drug for the intense euphoria it induces. Tolerance soon develops and users require larger doses to achieve the same effect. Heroin was first made from morphine in England in 1874 by CR Wright but it was some time later in 1898 when Bayer marketed the substance as a highly effective drug for the treatment of coughs, asthma, bronchitis and the discomfort associated with tuberculosis (Munsey, 2005). Employees of Bayer reported feeling 'heroic' after consuming the drug which led to its brand name 'Heroin'. By 1913 Bayer had stopped selling it as it became clear that, contrary to initial claims, it was addictive and a culture of recreational use had developed. Today, illicit cultivation of opium poppy remains a major problem. Much of the heroin consumed in the world originates in politically unstable Afghanistan. Illicit cultivation has also been documented in countries such as Iran, Pakistan, Thailand, Burma and Mexico. Continued use of heroin leads to dependence and associated social problems and this highlights the potential abuse of medicinal products from *Papaver somniferum*. The International Council for Security and Development has proposed that Afghanistan be afforded a preferred supplier

status in a bid to break its reliance on the opium trade and boost the supply of opiate based medicines. This proposal would imply that global production of morphine and thebaine is inadequate to satisfy demand. However, the International Narcotics Control Board (INCB), which is charged with monitoring the implementation of United Nations drug control conventions, has stated that production and stocks of morphine and thebaine are in excess of demand and notes that producing countries actually have plans to increase poppy production (INCB, 2010).

1.3 Benzylisoquinoline alkaloids

Benzylisoquinoline alkaloids (BIAs) are a large and structurally diverse group of alkaloids with approximately 2500 members (Facchini et al., 2007; Hagel and Facchini, 2013). They are secondary metabolites generally thought to play roles in plant defence against herbivores and pathogens. Much of the research on BIAs has focussed on their pharmacology in humans rather than their physiological roles in the plant. However, there are some reports of BIA roles as antifeedants to insects (Krug and Proksch, 1993; Park et al., 2000; Sellier et al., 2011; Shields et al., 2008) and inhibitors to the propagation of bacteria and viruses (Schmeller et al., 1997). The occurrence of BIAs is thought to have a monophyletic origin (Liscombe et al., 2005). These alkaloids are most common among the order Ranunculales, specifically the Papaveraceae, Ranunculaceae, Berberidaceae and Menispermaceae families (Ziegler and Facchini, 2008). Protoberberine (e.g. (S)-Scoulerine), benzophenanthridine (e.g. sanguinarine) and 1-benzylisoquinoline alkaloids have a broad distribution and have even been described in a gymnosperm, *Gnetum parvifolium* (Xu and Lin, 1999). In opium poppy (*Papaver somniferum; Papaveraceae*) sieve elements and specialised laticifers of the phloem are the sites of synthesis and storage of BIAs (Bird et al., 2003; Samanani et al., 2006) In contrast phloem tissues are not involved in BIA synthesis in meadow rue (*Thalictrum flavum; Ranunculaceae*; Samanani et al., 2005).

1.3.1 Alkaloids

The term alkaloid was introduced by Carl Friedrich Wilhelm Meißner in 1819 and it is derived from the Latin word *alkali* which in turn came from Arabic *al qali* or "ashes of plants" (Meißner, 1819). Alkaloids are a diverse group of nitrogen-containing plant secondary metabolites present in approximately 20% of plants (Facchini et al., 2007). Many of the approximately 12,000 known plant alkaloids have been exploited since the start of civilisation for their potent biological activities as potions, medicines, and poisons.

Pharmacological properties of alkaloids useful to mankind include antimalarial (quinine), antiasthma (ephedrine), anticancer (vincristine, vinblastine), vasodilatory (papaverine), muscle relaxant ((+)-tubocurarine), analgesic (morphine), antihyperglycemic (piperine) and antibacterial (sanguinarine). Many are used in traditional or modern medicine, or as starting points for drug discovery. Alkaloids are also used as stimulants (caffeine, nicotine, cocaine), narcotics and poisons (aconitine, tubocurarine).

Specialised plant secondary metabolites such as alkaloids are a reflection of adaptation of plants to their environment as many are thought to provide protection against various biotic and abiotic stresses. Alkaloids are thought to have important ecochemical functions in the defence of the plant against pathogenic organisms and herbivores or, as in the case of pyrrolizidine alkaloids, as pro-toxins for insects, which further modify the alkaloids and then incorporate them into their own defence secretions (Hartmann, 1999; Hartmann and Ober, 2000; Ober and Hartmann, 2000).

Chemical synthesis of many alkaloids is difficult due to their complex structures which often contain multiple asymmetric centres. For example, the structure of morphine, which has five such asymmetric centres, was not solved (Gates and Tschudi, 1952) until 146 years after it was first isolated. Since the 1950s, advances in technology have enabled researchers to study alkaloid biosynthesis in greater detail, and enzymes involved in the biosynthesis of indole, isoquinoline, tropane, pyrrolizidine, and purine classes of alkaloids have been identified and characterised (Kutchan, 1995). Tools such as feeding of plants with

radiolabelled precursors, analysis with nuclear magnetic resonance spectroscopy and plant cell culture techniques have aided our understanding of alkaloid biosynthesis. Plants are still largely superior chemists and we rely on alkaloid extraction from plants rather than difficult and uneconomical chemical synthesis of complex structures such as the morphinans (morphine, codeine, thebaine) from *Papaver somniferum* and anticancer agents (vinblastine and vincristine) from *Catharanthus roseus*.

1.3.2 Society's use of BIAs

Prominent members of the BIA family include emetine (an antiprotozal from the ipecac root, *Carapichea ipecacuanha*), colchicine (a microtubule disrupter used to treat gout from the autumn crocus, *Colchicum autumnale*) and berberine (used as an antimicrobial in eye drops and also in the treatment of intestinal disorders, cancer, HIV and diabetes).

Benzophenanthridine alkaloids, a subclass of the BIAs, are produced by a number of species within the *Papaveraceae.* These include bloodroot (*Sanguinaria canadensis*), opium poppy (*Papaver somniferum*), and California poppy (*Eschscholtzia californica*). Bloodroot extracts rich in the intensely red benzophenanthridine alkaloid sanguinarine are used in oral hygiene products because of its antiplaque properties. Berberine and sanguinarine have been considered as anticancer drugs (Bessi et al., 2012). Benzophenanthridine alkaloids have been shown to accumulate in cell cultures of *E. californica* in response to various fungal elicitors (Schumacher et al., 1987).

The 1-benzylisoquinoline papaverine has use as for treatments of cerebral vasospasm (Kassell et al., 1992) and erectile dysfunction (Dinsmore, 2005). Dimers of 1-benzylisoquinolines, known as bisbenzylisoquinolines often have use as pharmaceuticals e.g. dauricine found in Chinese herbal preparations of Moonseed (*Menispermum dauricum*) has use as an antiarrhythmic agent (Qian, 2002) and (+)-tubocurarine is used as a neuromuscular blocking agent in anesthesia (Wenningmann and Dilger, 2001) .

Morphinan and promorphinan alkaloids are produced by certain genera of the Papaveraceae, Menispermaceae and Berberidaceae (Liscombe et al., 2005) and include

valuable pharmaceuticals such as the narcotic analgesics codeine and morphine. Thebaine is not itself used as an analgesic, but it is converted into a wide range of semi-synthetic opiates such as oxycodone, oxymorphone, etorphine and burprenorphine (Berenyi et al., 2009). Thebaine is also used to manufacture a range of opiate antagonists including naltrexone, methylnaltrexone, nalmefene as well as naloxone.

Phthalideisoquinoline alkaloids are found in *Eschscholzia, Papaver, Dicentra* and *Glaucium* within the Papaveraceae, and *Cocculus* within the Menispermaceae (Liscombe et al., 2005). Noscapine is present in the latex of certain opium poppy chemotypes and other *Papaver* species (Sariyar, 2002) and has antitussive and anticancer properties (Barken et al., 2008; Mahmoudian and Rahimi-Moghaddam, 2009). The ability of opium poppy chemotypes to synthesise noscapine is conferred by the presence of a complex gene cluster encoding many of the enzymes required for its synthesis (Winzer et al., 2012).

Pavine, isopavine, rhoeadine and cularine alkaloids are smaller BIA subgroups found in most Papaveraceae genera and certain members of the Menispermaceae and Ranunculaceae, such as meadow rue (*Thalictrum flavum*) (Gözler et al., 1983; Liscombe et al., 2005; Montgomery et al., 1983).

1.3.3 Roles of BIAs in the plant

Various studies have offered potential roles for BIAs in plant defence. Berberine, for example, has antiherbivore properties towards a variety of insects (Park et al., 2000; Sellier et al., 2011; Shields et al., 2008). Sanguinarine and chelerythrine (a benzophenanthridine present in the greater celandine, *Chelidonium majus* and the plume poppy, *Macleaya cordata*) demonstrated activities against economically important pathogentic fungi (Liu et al., 2009). Similarly, antifungal and antimicrobial activities were attributed to aporphine (Zhang et al., 2012) and protopine alkaloids (Singh et al., 2009). Studies have also suggested BIA roles in the deterrence of larger pests such as the South American primate *Callithrix flaviceps* (Simas et al., 2001) and pocket gophers (Watts et al., 2011). A defensive ecophysiological role for morphine has been suggested since it was shown that, when opium poppy plants are

wounded, morphine is rapidly metabolised to two forms of bismorphine, which are incorporated into the cell wall to increase resistance to pectinase hydrolysis (Morimoto et al., 2003, 2001). Production of bismorphine is catalyzed by a bismorphine-forming peroxidase (BFP) that pre-exists in the capsules and leaves, allowing this defence system to be rapidly induced in response to mechanical damage. BFP does not accept codeine and thebaine as substrates and it was also demonstrated to be unable to catalyse the oxidation of various phenolics. Therefore, it appears to be a unique peroxidase with restricted substrate specificities as opposed to other plant peroxidases which catalyse a range of substrates.

In response to a fungal elicitor, sanguinarine production in opium poppy cell cultures was shown by microarray and RNA blot experiments to involve not only an increase in all known sanguinarine biosynthetic gene transcripts, but also an increases in the abundance of metabolic enzymes in primary metabolism such as sugar catabolism, the shikimate pathway, and aromatic amino acid metabolism (Zulak et al., 2006). The most abundant transcripts in response to the elicitor were for defence related proteins and enzymes involved in S-adenosylmethionine (SAM), alkaloid, and phenylpropanoid biosynthesis. A later study, which involved pyrosequencing elicitor treated opium poppy cell cultures, counted transcripts of proteins involved in metabolism, defence, signalling, transport and cellular structure as most abundant (Desgagné-Penix et al., 2010). The defence related transcripts encoded chitinase, β-lactamase, polyphenol oxidase, zyloglucanase inhibitor, peroxidase, and pathogenesis-related (PR) proteins. These were similar to results generated with an opium poppy seedling cDNA library in terms of ESTs associated with plant stress/defence and metabolism (Ziegler et al., 2005). While evidence for specific roles of BIAs is scarce, the available results indicate roles in plant defence against herbivores, insects and pathogens.

1.4 Biosynthesis of BIAs in opium poppy

Synthesis of specific alkaloids is under developmental control. Sanguinarine and dihydrosanguinarine are reported to be the most abundant root alkaloids (Frick et al., 2005) whereas noscapine and papaverine accumulate only in aerial organs (Facchini and De Luca,

1995). Although morphine and codeine are predominant in the aerial organs, they have also been found in roots.

1.4.1 Structural diversity from a common origin

All BIAs share a common precursor in (S)-norcoclaurine, which is formed by condensation of two derivatives of tyrosine, namely, dopamine and 4-hydroxyphenylacetaldehyde (4-HPAA) (Stadler et al., 1989, 1987)(Figure 2).

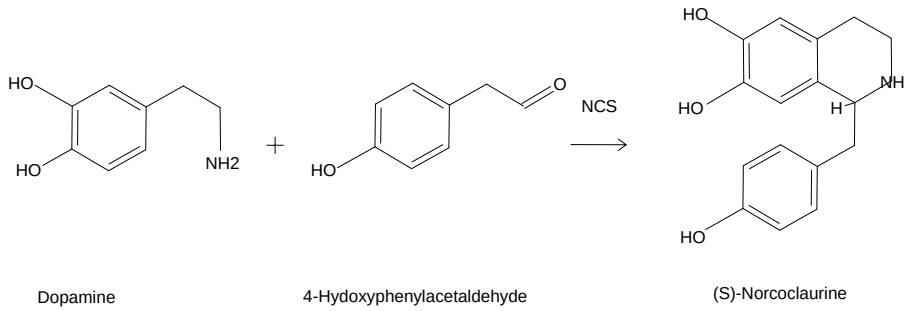


Figure 2. Condensation reaction leading to the common precursor of all plant benzylisoquinolines, (S)-norcoclaurine. The reaction is catalysed by norcoclaurine synthase (NCS).

The reaction is catalysed by norcoclaurine synthase (NCS), a member of the pathogenesis-related (PR10) and Bet v 1 protein families (DeLoache et al., 2015; Lee and Facchini, 2010; Liscombe et al., 2005; Samanani et al., 2004). NCS is a promiscuous enzyme that can also cyclise dopamine with a wide variety of acetaldehydes yielding various substituted tetrahydroisoquinoline alkaloids (Ruff et al., 2012). Subsequent 6-O-methylation of (S)-norcoclaurine followed by N-methylation, cytochrome P450 dependent 3'-hydroxylation, and 4'-O-methylation results in the formation of the central branch point intermediate, (S)-reticuline (Frenzel and Zenk, 1990; Pauli and Kutchan, 1998; Sato et al., 1993).

Various C-C or C-O coupling reactions involving one or two 1-benzylisoquinoline units yield the core backbone structures of other BIA subgroups (Figure 3). These backbones can then undergo modifications such as aromatic ring hydroxylation, or N-methylation generating a tertiary or quaternary amine, to yield the diversity of compounds. Hydroxyl groups enable downstream O-methylation, O-acetylation and methylenedioxy bridge formation.

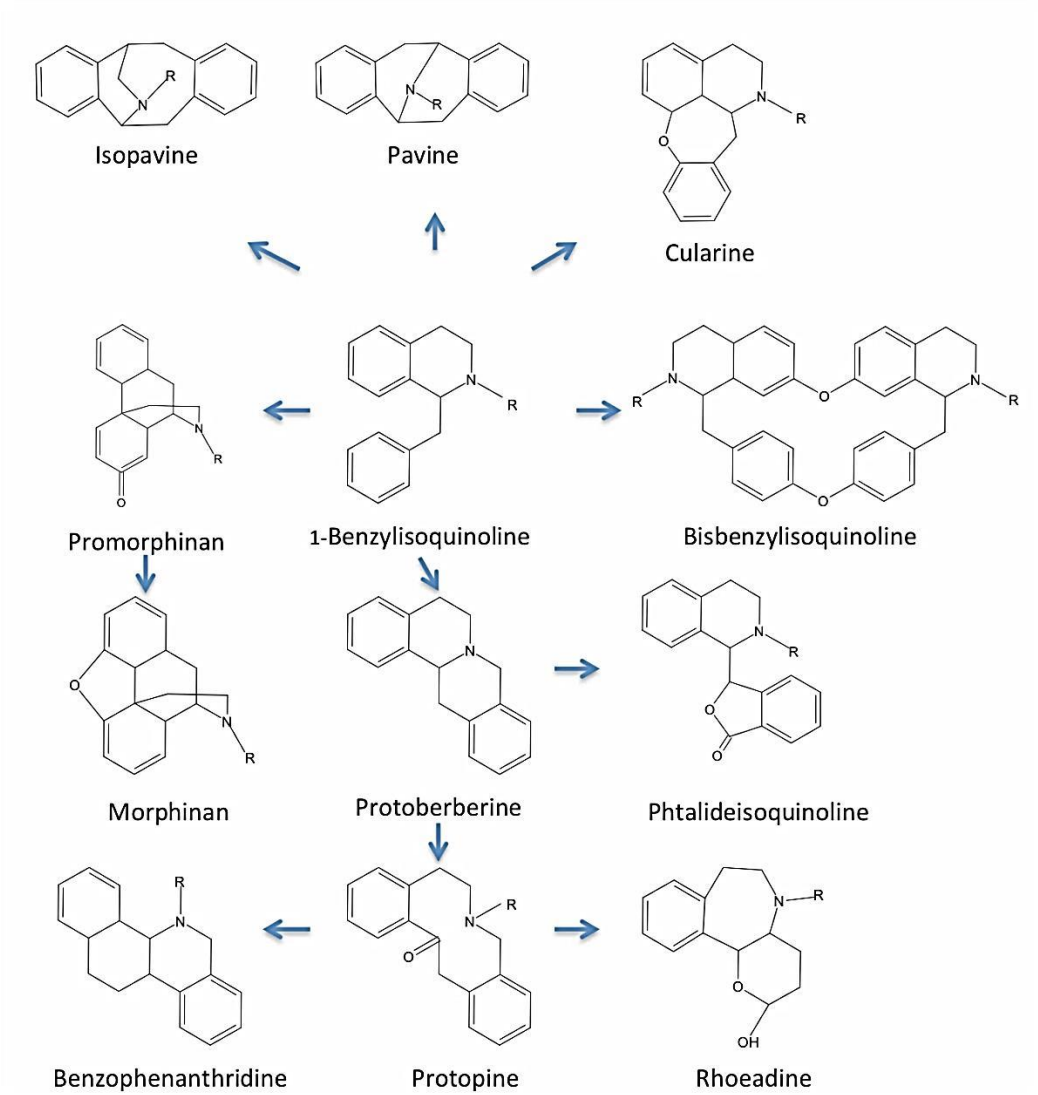


Figure 3. Representative benzylisoquinolines alkaloid subclasses with diverse backbone structures produced from the 1-benzylisoquinoline intermediate.

1.4.2 Benzophenanthridine and protoberberine biosynthesis

The first step in benzophenanthridine and protoberberine alkaloid biosynthesis is the conversion of (S)-reticuline to (S)-scoulerine catalysed by the berberine bridge enzyme (BBE) (Facchini et al., 1996b). Progress from (S)-scoulerine to sanguinarine is achieved by six enzymes including three CYPs (CheSyn, StySyn, P6H; Díaz Chávez et al., 2011; Ikezawa et al., 2007; Takemura et al., 2013), an N-methyltransferase (TMNT, Liscombe and Facchini, 2007) and a flavoprotein oxidase (MSH, Beaudoin and Facchini, 2013). The final oxidation of dihydrosanguinarine to sanguinarine is catalysed by dihydrobenzophenanthridine oxidase (DBOX, Hagel et al., 2012). The reverse reaction is catalysed by sanguinarine reductase (SanR), which may reduce the cytotoxic effects of sanguinarine within the cytoplasm (Weiss et al., 2006).

1.4.3 Phthalideisoquinoline biosynthesis

Alternatively, in poppy chemotypes that accumulate the phtalideisoquinoline noscapine, (S)-scoulerine undergoes 9-O-methylation by (S)-scoulerine O-methyltransferase (SOMT) to yield (S)-tetrahydocolumbamine (Takeshita et al., 1995). The introduction of a methylenedioxy bridge by canadine synthase (CanSyn; Ikezawa et al., 2003) and addition of a methyl group by tetrahydroprotoberberine-*cis*-*N*-methyltransferase (TNMT; Liscombe and Facchini, 2007) produce N-methylcanadine, a precursor of noscapine. Nine of the ten genes resident on the noscapine gene cluster participate in the synthesis of noscapine from N-methylcanadine (Winzer et al., 2012).

1.4.4 Promorphinan and morphinan biosynthesis

Morphine biosynthesis begins with the epimerisation of (S)-reticuline to (R)-reticuline (Farrow et al., 2015; Winzer et al., 2015). The reaction is catalysed by a P450-oxidoreductase fusion protein called STORR ((S)-to-(R)-reticuline). The P450 module catalyses the conversion of (S)-reticuline to a 1,2-dehydroreticuline intermediate which is then converted to (R)-reticuline by the oxidoreductase module.

Salutaridine synthase (SalSyn; Gesell et al., 2009) catalyses the conversion of (R)-reticuline to salutaridine, which is subsequently reduced to salutaridinol by salutaridine reductase (SalR; Ziegler et al., 2006). O-acetylation by salutaridine-7-O-acetyltransferase (SalAT; Grothe et al., 2001) is followed by with pH-dependent spontaneous rearrangement to thebaine. An enzyme facilitating this rearrangement has also been proposed (Fisinger et al., 2007).

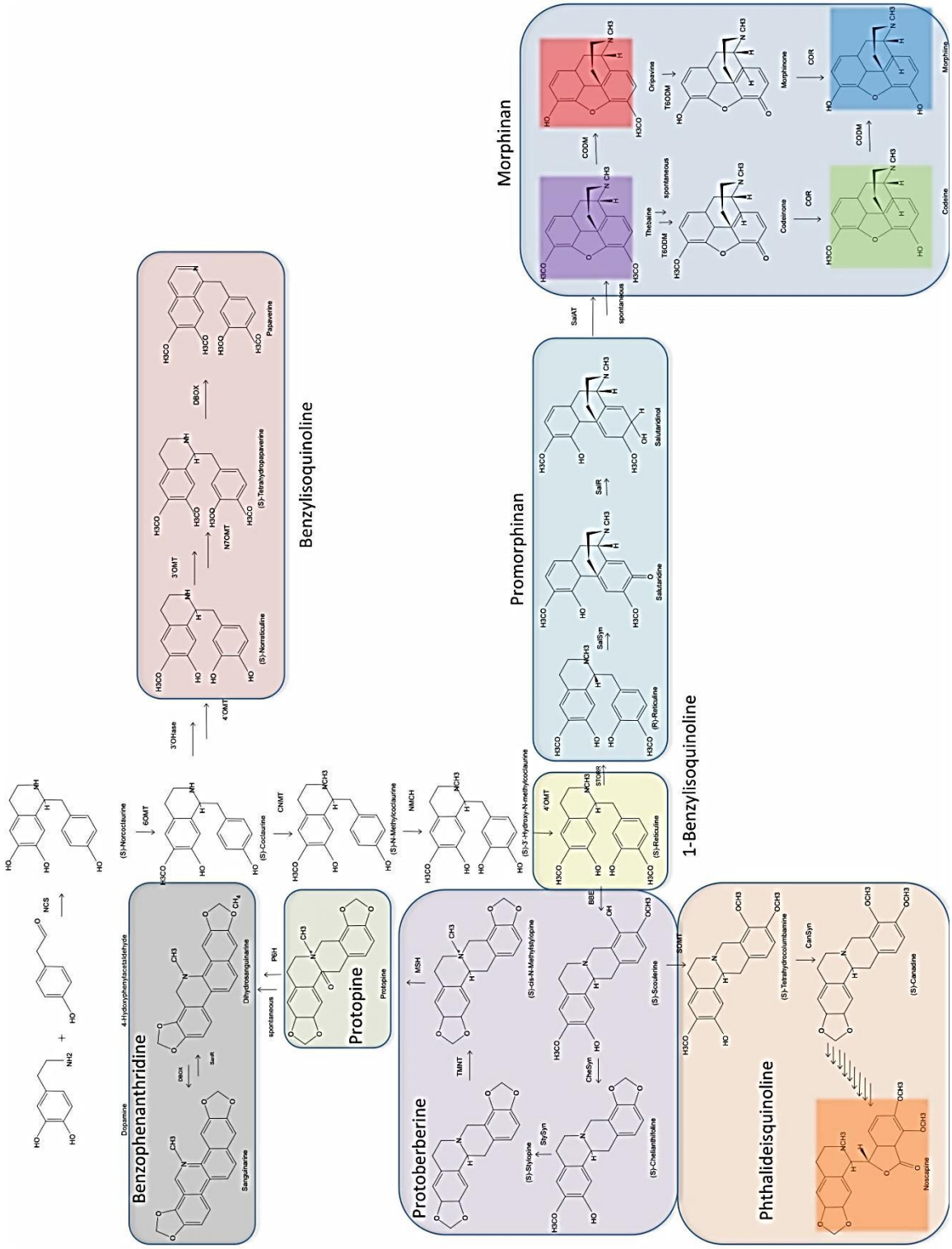


Figure 4. Biosynthesis of the major benzylisoquinoline alkaloids from derivatives of tyrosine dopamine and 4-Hydroxyphenylacetaldehyde. The major subclasses formed from the central branchpoint intermediate (*S*)-Reticuline are highlighted. Colours for the commercially valuable alkaloids noscapine, thebaine, oripavine, codeine and morphine will be maintained throughout the thesis. Acronyms for the genes involved are explained in the text.

Although the pathway is bifurcated at this point, the major route to morphine from thebaine (Figure 4) is thought to occur via codeinone (Nielson et al., 1983). The minor route involves the intermediates oripavine and morphinone (Brochmann-Hanssen, 1984). Codeinone reductase (COR) reduces codeinone to codeine in the penultimate step of the major pathway (Unterlinner et al., 1999). COR also reduces morphinone to morphine in the final step of the minor pathway. The recent identification of codeine-*O*-demethylase (CODM) and thebaine 6-*O*-demethylase (T6ODM) which catalyse unique *O*-demethylation steps of morphine biosynthesis completed the biochemical pathway (Hagel and Facchini, 2010a). Kinetic data using recombinant enzymes suggested the biosynthetic pathway through codeine is preferred. Virus Induced Gene Silencing (VIGS) showed large increases in the accumulation of thebaine and codeine upon silencing of T6ODM and CODM, respectively. Thus these genes are excellent targets for mutagenesis to obtain alleles with impaired functions. These alleles may provide stable and heritable alkaloid phenotypes and eventually lead to novel bespoke varieties of opium poppy.

1.4.5 Enzymes involved in BIA metabolism

A limited number of enzyme families are involved in BIA metabolism. The most prominent enzyme classes are the CYPs and S-adenosylmethionine-dependent *O*- and *N*-methyltransferases (Hagel and Facchini, 2013; Ziegler and Facchini, 2008). NCS, already mentioned above, is a member of the pathogenesis-related (PR10) and Bet v 1 protein families. FAD-linked oxidoreductases, and 2-oxoglutarate/Fe^{2+}-dependent dioxygenases and NADPH-dependent short-chain dehydrogenase/reductases are the other enzyme families with more than one participant in BIA metabolism. Codeinone reductase (COR) is a NADPH-dependent aldo-keto reductase involved in the final steps of morphine biosynthesis (Unterlinner et al., 1999). Finally, an acetyl-CoA dependent acetyltransferase called salutaridine 7-O-acetyltransferase (SalAT) is involved in the formation of thebaine. Figure 4 shows the array of BIAs that can be formed after the initial synthesis of (S)-norcoclaurine using this limited number of enzyme families.

1.4.6 Sites of BIA biosynthesis and storage

All members of the genus *Papaver* produce latex containing BIAs. Latex is the cytoplasm of specialised elongated secretory cells called laticifers (Hagel et al., 2008b). Latex has a positive turgor and exudes when tissues are damaged e.g. by a herbivore. Laticifers in *Papaver somniferum* have an articulated (i.e. composed of a series of cells joined together as seen in opium poppy) morphology (Nessler and Mahlberg, 1977). They are found adjacent to or proximal to the sieve elements of the phloem in all organs and contain vesicles into which alkaloids are sequestered (Hagel et al., 2008b; Nessler and Mahlberg, 1979). Vesicles are thought to derive from dilations of the endoplasmic reticulum (Alcantara et al., 2005). Compartmentalisation may provide protection from potentially toxic intermediates and confer controls over metabolic regulation. Laticifers contain low molecular mass major latex proteins (MLPs) in the cytosol whose precise function remains to be discovered (Nessler et al., 1985).

Results from experiments with *in situ* labelling of mRNA and proteins in opium poppy suggest pathway enzymes are assembled in companion cells and translocated to sieve elements where BIAs are synthesised (Bird et al., 2003; Samanani et al., 2006). Another study localised four BIA biosynthetic enzymes to phloem parenchyma rather than sieve elements and codeineone reducatase (COR) to laticifers (Weid et al., 2004). This discrepancy may have been due to differences in experimental design between the two groups i.e. differences in specificity of antibodies or differences in plant growth conditions.

A shotgun proteomics approach was recently used to determine the relative abundance of proteins in isolated latex and whole stem (including latex) (Onoyovwe et al., 2013). The results demonstrated that nine morphine biosynthetic enzymes were represented in the whole stem, whereas only four of the final five pathway enzymes were detected in the latex. This suggests that the genes catalysing the final steps of morphine biosynthesis, T6ODM, COR and CODM, are primarily located in laticifers. Noscapine synthase (NOS)(Chen and Facchini, 2014) and 7-OMT (Onoyovwe et al., 2013) are also abundant in laticifers, suggesting that synthesis of noscapine and methylated derivatives of reticuline such as laudanosine also

occurs in the latex. Upstream enzymes are spatially separated. They are made in companion cells and transported to sieve elements to catalyse reactions modifying the benzylisioquinoline backbone. It is thought that thebaine and narcotinehemiacetal are the substrates transported into laticifers from sieve elements for the final steps of the morphine and noscapine biosynthesis pathways, respectively.

Spatial regulation of pathway enzymes at the subcellular and cellular levels may be a way of optimising production by providing optimal microenvironments for reactions or reducing interactions with competing pathways. Elicitor treated opium poppy cell cultures produce sanguinarine rather than morphine. Transcripts for enzymes involved in sanguinarine biosynthesis were detected, but some of the enzymes involved in morphine biosynthesis were absent, which offers a transcriptional basis for the lack of morphine production in dedifferentiated cells (Desgagné-Penix et al., 2010).

Following their synthesis, BIAs are ultimately transported to adjacent or proximal laticifers for storage. It is not yet clear whether BIA transport is symplastic involving plasmodesmata or apoplastic implicating as yet unknown transporters or channels. One study has demonstrated a role for an ATP-binding cassette (ABC) transporter in berberine transport from the root to the rhizome in Japanese goldthread (*Coptis japonica*) (Shitan et al., 2003). ABC transporters are involved in the transport of other plant secondary metabolites such as terpenoids and phenolics (Yazaki, 2006, 2005). Additionally, an *N*-methyltransferase catalysing the first step in tropane alkaloid metabolism and a homolog of *BBE* from *Nicotiana benthamiana* were independently fused to GFP, and shown to localise at mesophyll plasmodesmata (Escobar et al., 2003). This result would indicate that alkaloid biosynthetic enzymes may possess targeting information which could facilitate symplastic transport. There has been suggestions that BIA biosynthetic enzymes may be organised into complexes or metabolic channels (Allen et al., 2004). This was offered as an explanation for why RNAi blockage of COR, in opium poppy resulted in elevated levels of (S)-reticuline and not codeinone. The loss of COR through RNAi was postulated to disrupt the normal multienzyme complex and prevent metabolic flux to morphine. It now appears than RNAi silencing was

non-specific and the closely related STORR gene, involved in the epimerization of reticuline, may also have been silenced (Winzer et al., 2015).

1.5 CODM and T6ODM are 2-ODDs involved in O-demethylation in alkaloid metabolism

T6ODM is a 2-oxoglutarate/Fe(II) dependent dioxygenase (2-ODD) identified by differential expression in a morphine and codeine free mutant (T). The morphinan pathway in this mutant was blocked at thebaine (Hagel et al., 2008a). An opium poppy 23,000-element microarray was used to compare the stem transcriptomes of this mutant and three morphine-accumulating varieties (L, 11 and 40). T6ODM, a 2-ODD, had lower transcript levels in T compared to the morphine-accumulating varieties. CODM, which is specific for the 3-O-methyl group of thebaine and codeine, was identified after querying an opium poppy expressed sequence tag (EST) database with T6ODM. Upstream oxidative reactions such as ring hydroxylation and C-C phenol coupling steps are catalysed by P450 family members and so the involvement of 2-ODDs in morphine biosynthesis was unexpected (Hagel and Facchini, 2010a). Demethylation of thebaine and codeine in humans is achieved with cytochrome P450 enzymes (Kramlinger et al., 2015; Zhu, 2008). The demonstration that poppy achieves these reactions with 2-ODDs was interesting because application of the acylcyclohexanedione trinexapac-ethyl (Moddus) increases thebaine content in crops at the expense of morphine (Cotterill, 2010). This chemical also inhibits dioxygenases involved in the synthesis of gibberellin hormones, and so these observations may hint at a possible mechanism of action of Moddus with regard to increasing thebaine content of poppy crops.

T6ODM and CODM are the only members of the 2-ODDs capable of catalysing O-demethylation reactions (Hagel and Facchini, 2010b). This class of enzymes require 2-oxoglutarate and molecular oxygen as substrates and iron Fe(II) as a cofactor to catalyse the oxidation of a substrate along with decarboxylation of 2-oxoglutarate forming succinate and carbon dioxide. Other members of the 2-ODDs are known to catalyse N-demethylation reactions. These include AlkB-type proteins, which catalyse the removal of N-linked methyl

groups of purine and pyrimidine bases in ribonucleic acids (Yi et al., 2009), and histone N-demethylases with Jumonji C domains (Mosammaparast and Shi, 2010). These 2-ODDs achieve demethylation by hydroxylation of the O- or N-linked methyl groups, leading to the departure of formaldehyde (Hagel and Facchini, 2010b; Loenarz and Schofield, 2008).

Kinetic analysis of recombinant T6ODM showed oripavine is actually the preferred substrate of T6ODM, followed by thebaine, while codeine is not accepted as a substrate. Recombinant CODM displayed a stronger preference for codeine than thebaine (Hagel and Facchini, 2010a).

1.5.1 O-demethylation in plant metabolism

Demethylases play a role in a number of biological processes such as epigenetic regulation, DNA repair, toxin degradation and the biosynthesis of specialised metabolites. These activities are achieved by members of enzyme families such as the cytochrome P450s, flavin adenine dinucleotide (FAD)-dependent oxidases, 2-ODDs and Rieske-domain proteins. T6ODM and CODM are examples in a small group of plant enzymes with the ability to catalyse O-demethylation reactions (Hagel and Facchini, 2010b). By contrast there are many examples of plant O-methyltransferases known to be involved in the biosynthesis of secondary metabolites (Lam et al., 2007).

1.5.2 2-ODDs are involved in both primary and secondary metabolism in plants

In plant primary metabolism, 2-ODDs are known to be involved in posttranslational modification by proline hydroxylations (P4H) (Gorres and Raines, 2010), DNA repair by the removal of deleterious methyl lesions from plant genomes (ALKBH2) (Meza et al., 2012), histone demethylation by the "Jumonji" enzymes (Takeuchi et al., 2006) and plant hormone synthesis and catabolism. They catalyse several steps in synthesis gibberellins (Yamaguchi, 2008), plant growth regulators that play roles in all stages of plant development such as germination, stem elongation, leaf expansion, flower and fruit development and seed set (Phillips, 1998). Ethylene, a plant growth regulator with roles in senescence and fruit

ripening, similarly has a 2-ODD (ACCO) involved in its synthesis (Nakajima et al., 1990; Straeten et al., 1990). Recently, a 2-ODD (DAO) was shown to be involved in indole-3-acetic acid (IAA) catabolism in rice (DeLoache et al., 2015; Lee and Facchini, 2010; Liscombe et al., 2005; Samanani et al., 2004). Salicylic acid (SA) is involved in plant defence, stress responses and senescence. An enzyme called S3H in *Arabidopsis thaliana* was shown to be involved in SA inactivation (Zhang et al., 2013).

2-ODDs have also been shown to have roles in the biosynthesis of numerous plant secondary metabolites. Several facilitate the formation of different classes of flavonoids such as flavone synthase (FNS I), flavanone 3 β-hydroxylase (F3H) and flavonol synthase (FLS) and anthocyanidin synthase (ANS), also known as leucoanthocyanidin dioxygenase (LDOX) (Cheng et al., 2014; Gebhardt et al., 2005). The 2-ODDs AOP2, AOP3 and GSL-OH are involved in the biosynthesis of glucosinolates in the *Brassicales*, compounds responsible for bitter flavour (Hansen et al., 2008; Kliebenstein et al., 2001). Hyoscyamine 6β-hydroxylase (H6H) is a 2-ODD involved in hallucinogenic and toxic tropane alkaloid biosynthesis in members of the Solonaceae family (Hashimoto and Yamada, 1986). Desacetoxyvindoline-4-hydroxylase (D4H) is a 2-ODD involved in the biosynthesis of the monoterpene indole alkaloid vindoline in *Catharanthus roseus* (Vazquez-Flota et al., 1997). BX6 is a member of a gene cluster encoding the enzymes necessary for the production of DIBOA and DIMBOA benzoxazinoid defense molecules in the Poaceae family (Jonczyk et al., 2008). The coumarins are another class of defence compounds that have 2-ODDs involved in their synthesis in *Arabidopsis* (Kai et al., 2008), sweet potato (Matsumoto et al., 2012) and common rue (Vialart et al., 2012). Finally, a gene cluster in solanaceous food plants which encodes the enzymes necessary for the production of antinutritional steroidal glycoalkaloids contains a 2-ODD (Itkin et al., 2013).

Crystal structure analysis of leucoanthocyanidin synthase (LDOX) and 1-aminocyclopropane-1-carboxylic acid oxidase (ACCO) have demonstrated components of the active site (Wilmouth et al., 2002; Zhang et al., 2004)-a double-stranded beta helix fold with a conserved sequence motif, $HX(D/E)X_nH$, that is essential for binding Fe(II) (Clifton et al.,

2006). 2-oxoglutarate-binding residues are less well conserved and may be characteristic of different subfamilies, whereas residues linked to substrate binding are variable and may involve mobile elements surrounding the active site (Loenarz and Schofield, 2008).

Until recently, three 2-ODDs were known in opium poppy- T6ODM, CODM and a dioxygenase initially called DIOX2 but later renamed protopine O-dealkylase (PODA) when its function was determined (Farrow and Facchini, 2013). A fourth, papaverine 7-O-demethylase (P7ODM), was shown to be involved in the biosynthesis of pacodine from the vasodilator papaverine (Farrow and Facchini, 2015). Sequence comparisons of known plant 2-ODDs indicated a monophyletic origin of the opium poppy members (Hagel and Facchini, 2010b). The phylogenetic analysis supports different origins of these enzymes and other 2-ODDs involved in monoterpene indole alkaloid biosynthesis (D4H) and tropane alkaloid biosynthesis (H6H). A more recent ancestral link was suggested between these enzymes and a noroclaurine synthase (NCS) involved in BIA metabolism from *Coptis japonica* (Minami et al., 2007). 2-ODDs involved in glucosinolate and flavonoid metabolism indicated a low degree of sequence conservation and suggest these are a highly diversified enzyme classes which evolved to meet the needs of the particular plant in which they are found (Hagel and Facchini, 2010b; Kawai et al., 2014). In contrast, 2-ODDs involved in primary metabolism were found to be conserved, indicating their important roles in core metabolism (Kawai et al., 2014). Figure 5 shows a multiple sequence alignment of opium poppy 2-ODDs, and *Petunia hybrida* ACCO and *Arabidopsis thaliana* LDOX for which crystal structures have been determined.

```
CLUSTAL O(1.2.1) multiple sequence alignment

PhACCO       ------------------------------------------------------MENFPI
AtLDOX       ----------MVAVERVESLAKSGIISIPKEYIRPKEELESINDVFLEEKKEDGPQVPT
PsCODM       METPILIKLGNGLSIPSVQELAKLTLAEIPSRYTCTGESPLNNI---GASVT-DDETVPV
PsPODA       METAKLMKLGNGMSIPSVQELAKLTLAEIPSRYICTVENLQLPV---GASVIDDHETVPV
PsT6ODM      MEKAKLMKLGNGMEIPSVQELAKLTLAEIPSRYVCANENLLLPM---GASVINDHETIPV

PhACCO       ISLDKV---NGVERAATMEMIKDACENWGFFELVNHGIPREVMDTVEKMTKGHYKKCMEQ
AtLDOX       IDLKNIESDDEKIRENCIEELKKASLDWGVMHLINHGIPADLMERVKKAGEEFFSLSVEE
PsCODM       IDLQNLLSPEPVVGKLELDKLHSACKEWGFFQLVNHGVDALLMDNIKSEIKGFFNLPMNE
PsPODA       IDIENLISSEPVTEKLELDRLHSACKEWGFFQVVNHGVDTSLVDNVKSDIQGFFNLSMNE
PsT6ODM      IDIENLLSPEPIIGKLELDRLHFACKEWGFFQVVNHGVDASLVDSVKSEIQGFFNLSMDE

PhACCO       RFKEL--VASKALEG-----VQAEVTDMDWESTFFLKHLPIS--NISEVPDLDEEYREVM
AtLDOX       KEKYANDQATGKIQGYGSKLANNASGQLEWEDYFFHLAYPEEKRDLSIWPKTPSDYIEAT
PsCODM       KTKYG--QQDGDFEGFGQPYIESEDQRLDWTEVFSMLSLPLHLRKPHLFPELPLPFRETL
PsPODA       KIKYG--QKDGDVEGFGQAFVASEDQTLDWADIFMILTLPLHLRKPHLFSKLPLPLRETI
PsT6ODM      KTKYE--QEDGDVEGFGQGFIESEDQTLDWADIFMMFTLPLHLRKPHLFSKLPVPLRETI

PhACCO       RDFAKRLEKLAEELLDLLCENLGLEKGYLKNAFYGSKGPNFGTKVSNYPPCPKPDLIKGL
AtLDOX       SEYAKCLRLLATKVFKALSVGLGLEPDRLEKEVGGLEELLLQMKINYYPKCPQPELALGV
PsCODM       ESYLSKMKKLSTVVFEMLEKSLQ--LVEIKGMTDLFEDGLQTMRMNYYPPCPRPELVLGL
PsPODA       ESYSSEMKKLSMVLFEKMEKALQVQAVEIKEISEVFKDMTQVMRMNYYPPCPQPELAIGL
PsT6ODM      ESYSSEMKKLSMVLFNKMEKALQVQAAEIKGMSEVFIDGTQAMRMNYYPPCPQPNLAIGL
               ◆  ◆                                                    ◆
PhACCO       RAHTDAGGIILLFQDDKVSGLQLLKDGQWIDVPPMRHSIVVNLGDQLEVITNGKYKSVMH
AtLDOX       EAHTDVSALTFIL-HNMVPGLQLFYEGKWVTAKCVPDSIVMHIGDTLEILSNGKYKSILH
PsCODM       TSHSDFSGLTILLQLNEVEGLQIRKEERWISIKPLPDAFIVNVGDILEIMTNGIYRSVEH
PsPODA       TPHSDFGGLTILLQLNEVEGLQIKNEGRWISVKPLPNAFVVNVGDVLEIMTNGMYRSVDH
PsT6ODM      TSHSDFGGLTILLQINEVEGLQIKREGTWISVKPLPNAFVVNVGDILEIMTNGIYHSVDH
                      * *
PhACCO       RVIAQKDGARMSLASFYNPGSD-AVIYPAPALVEKEAEENKQVYPKFVFDD--YMKLYAG
AtLDOX       RGLVNKEKVRISWAVFCEPPKDKIVLKPLPEMVSVESPAK---FPPRTFAQHIEHKLFGK
PsCODM       RAVVNSTKERLSIATFHDSKLE-SEIGPISSLVTPETPAL---FKRGR---------YED
PsPODA       RAVVNSTKERLSIATFHDPNLE-SEIGPISSLITPNTPAL---FRSGST--------YGE
PsT6ODM      RAVVNSTNERLSIATFHDPSLE-SVIGPISSLITPETPAL---FKSGST--------YGD

PhACCO       LKFQAKEPRFEAMKAMETDVKMDPIATV
AtLDOX       EQEELVSEKND-----------------
PsCODM       ILKENLSRKLDG-KSFLDYMRM------
PsPODA       LVEEFHSRKLDG-KSFLDSMRM------
PsT6ODM      LVEECKTRKLDG-KSFLDSMRI------
```

Figure 5. Multiple sequence alignment of opium poppy 2-oxoglutarate/Fe(II)-dependent dioxygenases and 2-oxoglutarate/Fe(II)-dependent dioxygenases from *Petunia hybrida* and *Arabidopsis thaliana* for which crystal structures have been determined. PhACCO, *Petunia hybrida* 1-aminocyclopropane-1-carboxylic acid oxidase (Q08506); AtLDOX, *Arabidopsis thaliana* leucoanthocyanidin dioxygenase (Q96323), PsCODM, *Papaver somniferum* codeine O-demethylase (ADD85331), PsPODA, *Papaver somniferum* protopine O-dealkylase (ADD85330), PsT6ODM, *Papaver somniferum* thebaine 6-O-demethylase (ADD85329). Diamonds indicate the conserved catalytic triad motif essential for binding Fe(II). Asterisks indicate residues thought to be involved in 2-oxoglutarate binding. Clustal Omega at EMBL-EBI was used to create the multiple sequence alignment (Sievers et al., 2011).

1.5.3 Broader roles of T6ODM, CODM and PODA in opium poppy

The three opium poppy 2-ODDs were recently shown to differentially accept various non-morphinan BIA substrates (from a pool of 26) (Farrow and Facchini, 2013). In particular certain protopine and protoberberine alkaloids were effective substrates for catalysis. CODM and PODA were also shown to catalyse the O,O-demethylation of methylenedioxy bridges on protopine alkaloids. Protopines are intermediates in the synthesis of benzophenanthridine and rhoeadine alkaloids. PODA can also catalyse O,O-demethylation on protoberberine alkaloids. The metabolic roles of 2-ODDs in the O-demethylation and O,O-demethylation of protopine compounds and downstream derivatives such as sanguinarine and rhoeadine alkaloids (N-methylporphyroxine and glaudine) was demonstrated with virus induced gene silencing (VIGS). PODA expression was not detected in the cultivar of opium poppy used in the experiment so VIGS was targeted to T6ODM, CODM and both genes simultaneously (Farrow and Facchini, 2013). CODM displayed greater affinity to some of these substrates than towards thebaine and codeine, pointing to the possible origin of recruitment of these genes into the morphinan-branch pathway.

Silencing of T6ODM alone, or both genes together, resulted in a substantial increase in the accumulation of thebaine at the expense of morphine and codeine. By contrast, suppression of CODM transcript levels resulted in a significant increase in the accumulation of codeine and lower levels of morphine. This confirms earlier results showing roles for CODM and T6ODM in morphine biosynthesis (Hagel and Facchini, 2010a; Wijekoon and Facchini, 2012). Sanguinarine in the roots was down in plants with reduced levels of T6ODM and/or CODM transcripts while levels of the protopine alkaloids cryptopine, protopine and allocryptopine increased significantly (Farrow and Facchini, 2013). O-demethylcryptopine increased in abundance in CODM-silenced plants. Similarly, accumulation of the protoberberines N-methylstylopine and N-methylcanadine were increased in T6ODM and/or CODM-silenced plants.

PODA performs an O,O-demethylation on the isoquinoline moiety of the protopine or protoberberine backbone. In contrast CODM targets methylenedioxy bridges on the

isoquinoline or benzyl moieties of protopine alkaloids only. The accumulation of protopine alkaloids in plants with reduced CODM transcript levels was generally in agreement with the catalytic functions and substrate range of recombinant CODM, which displayed high relative activities against protopines. Recombinant T6ODM, however, displayed only trace activities towards protopines and so the accumulation of these compounds in response to T6ODM silencing was unexpected. O,O-demethylated BIAs were not detected *in planta*, suggesting their rapid recycling or redistribution into other products.

T6ODM and PODA share 85% amino acid sequence identity, yet PODA displays only trace levels of 6-O-demethylase activity with thebaine and oripavine (Farrow and Facchini, 2013). Remarkably, the enzymes catalyse different reactions on the same substrate. (R,S)-canadine and allocryptopine were O-demethylated by T6ODM but O,O-demethylated by PODA. These results may offer an insight into the origins of CODM and T6ODM. Morphine biosynthesis in opium poppy may have arisen from duplication and neofunctionalisation of ODDs involved in the formation of protopine, protoberberine, benzophenanthridine and rhoeadine alkaloids. The effects of minor amino acid changes on the substrate specificities and activities of opium poppy ODDs was demonstrated when a four amino acid substitution of CODM resulted in an enzyme capable of O-demethylating codeine, but not thebaine (Runguphan et al., 2012). The mutations are thought to change a region of alpha helix in the wild type CODM to a random coil in the codeine-specific mutant. This change appears to prevent the productive binding of thebaine without altering the efficiency of codeine O-demethylation.

Certain chemotypes of *Papaver somniferum* have a 10-gene cluster enabling diversion of (S)-reticuline away from morphine biosynthesis and towards the synthesis of the anticancer alkaloid noscapine (Winzer et al., 2012). It was therefore reasonable to suggest the plant also contains a similar cluster for the synthesis of morphine. Clusters of genes in plants are thought to arise through gene duplication followed by neo-and sub-functionalisation or translocation (Boycheva et al., 2014). Clustering could lead to improved functionality and heredity. Coordinate regulation or the avoidance of the build up of toxic intermediates may be reasons why clustered genes operate more efficiently. Coinheritance of an entire pathway

leading to an important defence molecule against pathogens, e.g. avenacin in oat (Qi et al., 2004) or herbivores e.g benzoxaxinoids in grasses (Sue et al., 2011), can confer a selective advantage.

1.5.4 Regulation of CODM and T6ODM

Treatment of opium poppy with trinexapac-ethyl (a acylcyclohexanedione inhibitor of dioxygenases involved in the synthesis of gibberellin hormones) increases thebaine and codeine content at the expense of morphine (Cotterill, 2010). However, acyclcyclohexanediones did not inhibit enzyme activities of T6ODM and CODM *in vitro*, suggesting an as yet unknown indirect mode of action (Hagel and Facchini, 2010a). Yields of morphine have been suggested to increase in response to treatment of gibberellic acid (Khan et al., 2007). The developmental and phytohormonal control of BIA biosynthesis in opium poppy is an area that needs more attention (Beaudoin and Facchini, 2014).

WRKY and bBLH transcriptional regulators have been shown to be involved in BIA biosynthesis in *Coptis japonica* (Kato et al., 2007; Yamada et al., 2011). In opium poppy transcriptional regulation of BIA biosynthetic genes in response to wounding showed a possible role for WRKY transcription factors (Mishra et al., 2013). WRKY elements also appear in the promoter regions of components of the noscapine gene cluster (Winzer et al., 2012) and a latex specific dioxygenase (CODM; Raymond, 2004), suggesting a wider role for WRKY transcription factors in the regulation of BIA biosynthesis.

1.6 Supply of opiates

The medicinal properties of *Papaver somniferum* have long been recognised due to its unique ability to synthesize the analgesic morphinans morphine, codeine and thebaine. Opium poppy also produces a variety of perhaps less well known biologically active BIAs such as the antispasmodic 1-benzylisoquinoline papaverine, the antimicrobial benzophenanthridine sanguinarine and the antitussive and anti-tumour phthalideisoquinoline noscapine.

1.6.1 Controlled substances

The poppy industry is highly regulated under international law. The INCB implements the United Nations *Single Convention on Narcotic Drugs 1961* (the Convention), as amended by the *Protocol Amending the Single Convention 1972*. It determines quotas for the growing of narcotic plants based on estimates of worldwide demand. In 1972, Australia restricted the growing of opium poppies to the state of Tasmania. However, in recent years permission has been granted to start growing poppies in Victoria and the Northern Territory (Safi, 2014). Cultivation, possession or refining of opium poppies are all offences in Tasmania unless in the possession of a licence which is granted annually.

1.6.2 Tasmania

Tasmania in Australia is the world's largest producer of opium alkaloids for the pharmaceutical market. GlaxoSmithKline in Latrobe and Tasmanian Alkaloids in Westbury together produce half of the world's concentrate of poppy straw (CPS) from approximately 10% of the production area (Dicker, n.d.). CPS is the concentrated near pure alkaloid powder produced from the chemical extraction of poppy straw (husk). CPS-morphine, CPS-codeine, CPS-thebaine and CPS-oripavine are all produced in Australia. Tasmanian Alkaloids commenced operations in 1975 and it is now owned by Johnson & Johnson. A third producer began operations in the state in 2004. TPI Enterprises is Australian owned and it also has operations in Victoria, the Northern Territory and Portugal.

The poppy industry in Tasmania began in the 1960s after a subsidiary of the multinational Glaxo sought out a replacement for an unreliable British climate. Trials were undertaken in New South Wales, Southern Australia, Western Australia and New Zealand. The combination of a suitable climate (spring rains, lots of sunshine and hot dry summers) and fertile soils (high organic matter and suitable pH) ensured Tasmania remained the location of choice. Today the three producers provide seed to farmers with which 20,000-30,000 hectares of opium poppy are grown each year. The industry is a major contributor to the economy of the state. It uses the extraction of dried poppy straw, based on that first

developed by Janos Kabay in Hungary in 1925, with modified bespoke extraction technologies. The capsules are left to dry in the summer sun and then crushed to straw. Excess seed can go on to be used by the food or poppy oil industries. The remaining husk is used to extract alkaloids by chemical treatment. In Tasmania, Macquarie Oils extracts poppy seed oil from seed collected from thebaine crops. The poppy meal that remains after the oil has been extracted has use as an agricultural fertiliser. The oil is used in bio-diesel. The alkaloid industry grosses in excess of AUD 100m (£52m) annually and farm gate returns to growers are typically AUD 70-90m (£36-47m) (Tasmanian Government, 2012). Tasmania's geographical isolation and climate together with the expertise and best practices developed by both farmers and processors over the past 40 years have ensured that Tasmania is the largest producer of opiate based analgesics in the world.

1.6.3 Other producers

Turkey, Hungary, France and Spain also produce CPS products. India, by contrast, produces traditional opium by making incisions in the epidermal wall of the capsules of unripe plants. When the latex that exudes dries it is called opium. The US and UK are the largest markets for CPS, while the US is the largest importer of opium (Fist, 2001). However, the US has an "80-20" rule whereby it imports 80% of its narcotic raw materials from the traditional producers, India and Turkey. Imports of thebaine are not restricted with the "80-20 rule" and so in recent years Australia has exported much of the thebaine obtained from high thebaine varieties to the US.

1.6.4 Potential alternative BIA production systems

BIAs made by opium poppy are structurally complex and chemical synthesis is not economical. Only sanguinarine has been found in significant levels in opium poppy cell cultures (after treatment with a fungal elicitor) (Alcantara et al., 2005; Facchini and Bird, 1998; Zulak et al., 2006). Specialised lacticifer cells are required for synthesis of morphinans in the plant. Inactivation of developmentally regulated genes or spatial segregation of

pathway intermediates from enzymes involved in morphine biosynthesis may account for the failure of dedifferentiated cell cultures to produce morphinans.

Many groups are working towards reconstituting BIA biosynthesis in microbial systems to provide an alternative to metabolic engineering in the native host, opium poppy (Glenn et al., 2013). Recently, a complex ten gene pathway introduced to yeast resulted in the formation of dihydrosanguinarine and its oxidised derivative sanguinarine from (R,S)-norlaudanosine (Fossati et al., 2014). Commercial production of sanguinarine is achieved with the bloodroot (*Sanguinaria canadensis L.*). Similarly, the ability of engineered yeast to convert simple sugars to an intermediate in BIA biosynthesis, salutaridine, was demonstrated (Hawkins and Smolke, 2008). Recently, a seven enzyme pathway introduced to yeast was shown to enable conversion of L-tyrosine to (S)-reticuline (DeLoache et al., 2015). A second recent paper described how the second half of the pathway to morphinans was introduced to yeast (Fossati et al., 2015) leaving all except one of the chemical reactions required to convert sugar to morphinans achievable in yeast. The last remaining enzyme involved, STORR ((S)-to-(R)-reticuline), was recently discovered (Farrow et al., 2015; Winzer et al., 2015), opening up the prospect of commercial production of morphinans in yeast. Manufacture of the antimalarial terpenoid artemisinin, originally derived from *Artemisia annua*, is an example where a pathway to a plant based medicinal product can be viably produced in yeast (Paddon et al., 2013). Semi synthetic production of artemisinin can be achieved whereby the steps up to artemisinic acid are produced in yeast leaving the final steps involving introduction of the endoperoxide bridge to be performed by an expensive photooxidative process. This final step adds significantly to the overall cost of production of artemisinin by genetic engineering.

Poppy BIA production, however, is a much more challenging proposition with a complex multistep biosynthetic pathway. The yields of morphinans produced by engineered yeast to date are very low, and, until the process is just as efficient and cost effective, commercial production of alkaloids is likely to remain in plant based systems in the short to medium term. Thebaine yields in excess of 70 kg/ha can be achieved in Tasmania and so the target for microbial based production systems is very high. The need increased efficiency of these

systems was recognised in another recent paper describing the production of thebaine in yeast by expressing 21 enzymes from a variety of sources (Galanie et al., 2015). However, one advantage the approach may have is in the production of alkaloids that would not accumulate *in planta*. The limitations of plant metabolic engineering may be overcome with microbes and allow exploration of the potential structural diversity of natural product derivatives. Yeast expressing poppy and *Pseudomonas putida* M10 genes were shown to convert thebaine to valuable final products such as hydrocodone and oxycodone (Galanie et al., 2015; Thodey et al., 2014).

1.7 Poppy Breeding

Opium poppy has been bred for thousands of years. Although there are many species under the genus *Papaver* which produce alkaloids, only *Papaver somniferum* L and *Papaver setigerum* DC produce morphine in significant quantities (Panicker et al., 2007). In the past two decades global consumption of opiates and opioids, as measured by the INCB, has tripled (Figure 6a; INCB, 2015). 'Opiates' is the term used to describe drugs derived from opium poppy and their derivatives, such as the semi-synthetic alkaloids. 'Opioids' is a more general term for both natural and synthetic drugs with morphine-like properties. They are mainly used for their analgesic properties to treat severe pain (fentanyl, hydromorphone, methadone, morphine, pethidine), moderate to severe pain (buprenorphine and oxycodone) and mild to moderate pain (codeine, dihydrocodeine, dextropropoxyphene) as well as in anaesthesia (fentanyl and its analogues such as anfeltanil). They are also used to treat cough (codeine, dihydrocodeine), diarrhoea (codeine and diphenoxylate) and addiction (buprenorphine and methadone). Innovation in poppy breeding is required to keep pace with demand for morphine but also to develop varieties with increased levels of key intermediates that have valuable uses in the preparation of pharmaceuticals. Increases in the numbers of defined daily doses for statistical purposes (S-DDD) are correlated with landmark growths in the use of oxycodone and hydrocodone in 2000-2001 and buprenorphine in 2005. These three narcotics are derived from thebaine because older chemical routes from morphine are highly inefficient.

1.7.1 Requirement for varieties with high yields of codeine

Codeine is an opiate used for its analgesic, antitussive and antidiarrheal properties (American Society of Health-System Pharmacists, 2015.). It is on the WHO Model List of Essential Medicines, a list of the most important medicines required for a basic health system (WHO, 2013). It is a natural alkaloid of the opium poppy plant *Papaver somniferum* L., but normally constitutes only a minor fraction of the total alkaloids. Codeine comes from the Greek word "kodeia" which means "poppy head" (Van Hout et al., 2014). Codeine is now the most widely used opiate in the world and is regarded as an effective analgesic with a wide safety margin. Its strength ranges from 8-12% of morphine in most people. This reduced potency means codeine is a more versatile drug. In our bodies codeine is converted into morphine and codeine-6-glucoronide (Vree et al., 2000). In fact over 80% of a dose of codeine is conjugated to glucuronic acid to codeine-6-glucoronide and this is the primary route to analgesia from codeine.

The majority of the global codeine is consumed as codeine phosphate with small amounts as codeine hydrochloride and codeine base. These salts are then used to manufacture medications for the relief of pain (analgesics) and cough (antitussives). Codeine is often used in the treatment of diarrhoea. Codeine has also been used as the starting material and prototype of a large class of mainly mild to moderately strong opioids such as dihydrocodeine, hydrocodone (Vicodin when combined with acetaminophen) and oxycodone (OxyContin). These opioids are generally made from thebaine in recent times. Codeine is marketed as both a single-ingredient drug and in combination with paracetamol (co-codamol e.g. Panadeine), aspirin (co-codaprin) or ibuprofen (Nurofen Plus). Codeine is also sold in cough syrups in combination with a range of other active ingredients.

Worldwide demand for codeine far exceeds the available supply produced in plants. Therefore most (85-90%; Fist et al., 2010) of the codeine currently being manufactured is obtained from morphine through the process of O-methylation, with the result that the major part of the world's licit morphine production is to support the manufacture of codeine (Figure 6b).

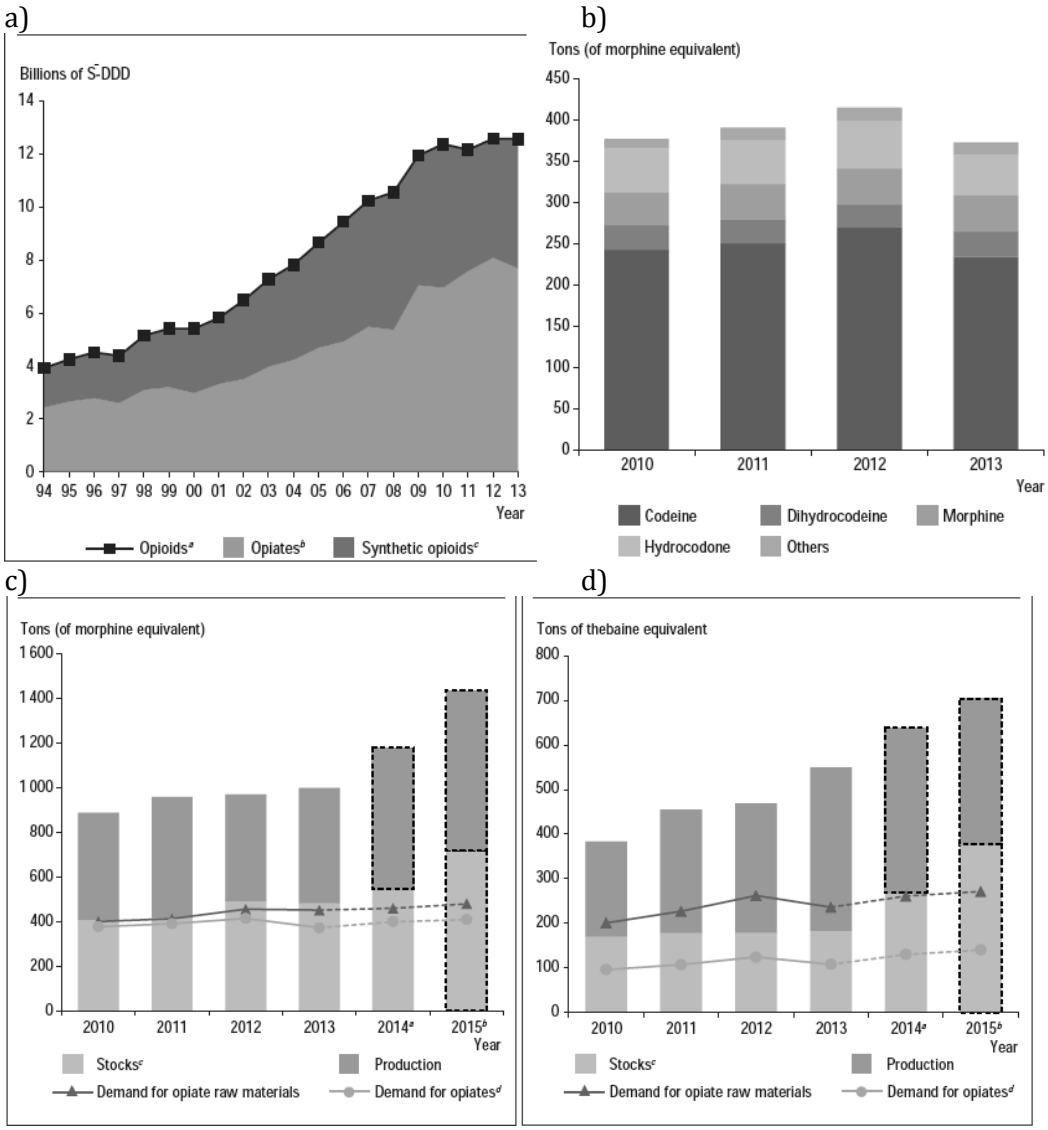


Figure 6. International Narcotics Control Board estimated (dotted lines) world requirements for 2015. a) Increasing consumption of opiates and opioids across the world in the last two decades. To allow for aggregation of data for drugs with different potencies consumption levels are expressed in billions of defined daily doses for statistical purposes (S-DDD). b) Consumption of opiates manufactured from morphine, in tonnes of morphine equivalent, 2010-2013. c) Expected rise in demand for opiate raw materials rich in morphine. d) Expected rise in demand for opiate raw materials rich in thebaine. Source: INCB, 2015.

Quaternary ammonium methylating agents in the presence of various bases are used to methylate morphine at the phenolic hydroxyl group at position 3 (Fist et al., 2010). However this approach generates N,N-dimethylaniline (DMA), a particularly odiferous and toxic byproduct which must be completely removed from the product. This imposes waste disposal, safety and environmental concerns. Solvents such as toluene or xylene are also required in the process, which creates further safety, environmental and cost burdens. Industrial-scale morphine methylations may suffer from yield losses from undermethylation, leaving too much unreacted morphine, or on the other hand overmethylation, leading to

formation of codeine-O (6)-methyl ether ('methylcodeine'). Impurities such as methylcodeine and dimethylpseudomorphine are difficult to remove and can result in significant yield losses.

The first reported commercial codeine crops were reported in Australia in 2009 and France in 2013. Global utilisation of codeine CPS amounted to 24.6 tonnes in 2013, which is still a fraction of the amount of morphine used (INCB, 2015). Over 350 tonnes of codeine are produced worldwide each year, with France and the UK being the biggest producers (approximately 20% each). India is by far the biggest importer of codeine with over 40 tonnes imported and consumed in 2013 (INCB, 2015). Global demand by manufacturers for opiate raw materials rich in morphine is increasing modestly (Figure 6c) and is expected to be above 400 tonnes this year. Much of this is to support the manufacture of codeine. On the whole industrial-scale methylation of morphine for the production of codeine is an expensive and time consuming process, one that could be reduced or avoided with the development of a variety in which the morphinan pathway was blocked at codeine.

1.7.2 Requirement for varieties with high yields of thebaine and oripavine

Demand for thebaine-based opiates is concentrated mainly in the United States and has increased sharply since the late 1990s (for oxycodone in particular). Thebaine reached a production peak of 158 tonnes in 2012 (INCB, 2015). Among the biggest producers the US uses much of its thebaine internally and so Australia and Spain are the main exporters to the rest of the world. The UK is the largest importer of thebaine (25.9 tonnes in 2013). Thebaine and oripavine are themselves not used therapeutically but are used for the synthesis of powerful drugs such as oxycodone (OxyContin), buprenorphine (Subutex) and naloxone (Narcan) which are used as analgesics and in the treatment of addiction (Berenyi et al., 2009). Ethorphine is another semi-synthetic analgesic that is 8000 times more potent than morphine and is used to immobilise large game animals.

Oripavine is an intermediate in the alternative pathway from thebaine to morphine (Brochmann-Hanssen, 1984). In 2013 oripavine was produced by the US (14.9 tonnes),

Australia (2.7 tonnes) and Switzerland (0.8 tonnes) (INCB, 2015). However, the US and Swiss figures actually refer to imported oripavine from Australia. Oripavine is used to make naloxone, naltrexone, methylnaltrexone, hydromorphone and oxymorphone.

Global trends point to continuing increases in demand for thebaine and oripavine (Figure 6d) and so there is a genuine requirement to develop new varieties with increased yields of these important alkaloids. Oripavine, in particular, is in demand, and development of lines with high yields of this alkaloid is a priority. The reason is that the above listed derivatives require O-demethylation using expensive and toxic reagents with concurrent yield loss. Development of varieties with metabolic blocks at thebaine or oripavine would increase yields and therefore lower production costs of both of these valuable opiates.

1.7.3 GSK Australia's opiates division

The poppy industry in Tasmania was set up by GSK in the 1960s and the company still produces about 25% of the world's opiates supply on the island. It does not grow commercial crops itself but rather contracts out the growing to farmers in the fertile areas of the northern coast and central plains. Farmers are paid in terms of yield of specific alkaloid in the harvested crops. Prices are negotiated with the Tasmanian Poppy Growers Association prior the planting period (July-October). Prices fluctuate based on global demand for specific alkaloids and levels of stocks held in vaults. A GSK field officer would then provide written recommendations to the farmer concerning, for example, weed and pest control. Typically, poppies are grown in a field once every 3-4 years as part of a rotation with other crops such as onion or potato. The plants flower in December and are ready to harvest from January to March. Harvesting of dried poppy straw and capsules is automated. Harvested material is initially brought to GSK's site at Latrobe, Tasmania for initial processing. The seed cleaning facility removes the seed for export to the food industry around the world. This accounts for approximately 5% of sales. However, the main value in the crop is in the alkaloids it produces and so the remaining straw and dust is pelleted before being loaded on trucks bound for the ferry across the Bass Strait to Melbourne. Extraction of opiates takes place at GSK's Port

Fairy, Victoria site. The company also has a number of sites in Tasmania for field trials to develop new varieties for future production. The research programme broadly aims to increase the yield of alkaloid per hectare, improve crop reliability and reduce the cost of production. In 2015 GSK Australia's opiates division was acquired by Mumbai based Sun Pharmaceutical Industries (SunPharma).

1.7.4 Chemical sprays used by the poppy industry

The costs of spraying crops are the borne by the grower. As well as regular spraying for herbicides and fungicides, the grower may apply specific chemicals for altering characteristics of the crop. Application of Moddus decreases the proportion of morphine in the capsules of poppy plants and increases proportions of thebaine and oripavine (Cotterill, 2010). Ultimate has the effect of reducing seed weight by up to 30% while increasing stem weight by the same proportion (personal communication). Sumagic is a plant growth regulator which is used for reducing plant height (Dean, 2011). It is thought to achieve this by affecting gibberellin production in sprayed plants. Breeding these desired characteristics into poppy would reduce the cost for both the grower and GSK, while also reducing the amount of chemicals released into the environment for the purpose of crop production.

1.8 Sources of genetic variation for breeding

During crop evolution there has been a continuous reduction in genetic diversity as breeders focussed on "elite" cultivars. This genetic erosion can become a bottleneck. Collection and characterisation of diverse germplasm from distinct geographic regions can be a good starting point for breeding. This is largely the conventional approach to breeding with opium poppy whereby variability and useful traits can be introduced by hybridisation. Alternatives are genetic modification and the artificial increase in variation though (random) mutagenesis. The latter aims to provide maximum genomic variation with a minimum decrease in viability of the plant. It has been successfully applied in recent decades to introduce much needed variation in commercial crops. Records maintained by the joint FAO-

IAEA Division show that 3218 crop cultivars, with useful traits introduced by mutagenesis, have been released (FAO-IAEA, 2015).

1.8.1 Genetically engineered opium poppy

Metabolic engineering often involves the transfer of desired genes into the genome of a plant cell, followed by regeneration of a transgenic plant from undifferentiated tissue. Transformation of poppy has been reported with plants (Chitty et al., 2003; Frick et al., 2004; Park and Facchini, 2000a; Pathak et al., 2012), root cultures (Facchini et al., 2008; Flem-Bonhomme et al., 2004; Park and Facchini, 2000b) and cell suspension cultures (Belny et al., 1997). However, it remains a difficult and inefficient procedure due primarily to recalcitrant regeneration.

To date, poppy transformation protocols have been used to upregulate or silence specific genes for research purposes. Genetically engineered poppy plants expressing antisense-*BBE* have altered latex profiles compared to wild type plants (Frick et al., 2004). RNAi mediated silencing of members of the *COR* gene family resulted in the accumulation of the non-narcotic (S)-reticuline rather than morphine (Allen et al., 2004). This may have resulted in concurrent silencing of the STORR locus involved in the epimerisation of reticuline (Winzer et al., 2015).

Genetically engineered opium poppy crops are not in use commercially. Furthermore, in Tasmania, where over half of the world's opiate based medicines are produced, a moratorium on genetically modified organisms has been in place. There are increasing calls to lift this moratorium from both industry ("GMO Review Submission No. 158 - GSK Australia - Department of Primary Industries, Parks, Water and Environment," 2013) and growers (Williams, 2013). Traditional mutagenesis approaches avoid the hurdles of regulation in the meantime.

1.8.2 Chemical mutagenesis

Chemical mutagenesis with alkylating agents such as ethylmethane sulfonate (EMS) causes a high density of irreversible point mutations (Henikoff and Comai, 2003). Guanine is

alkylated leading to a transition of the original G/C base pair, resulting in the formation of an A/T pair (Greene et al., 2003). In Arabidopsis, it was estimated that 5% of EMS-generated mutations lead to premature termination of the gene product or disruptions in splicing whereas 50% cause missense mutations (Greene et al., 2003; Martín et al., 2009). *N*-methyl-*N*-nitrosourea (MNU) treatment of rice and soybean produced primarily G/C to A/T transitions (similar to EMS) but a combination of MNU and sodium azide resulted in 20% A/T to G/C changes (Cooper et al., 2008; Suzuki et al., 2007; Till et al., 2007).

Chemical mutagenesis can result in four types of mutation:

- Silent: in an intron or non-coding sequence
- Neutral: nucleotide substitution but new codon encodes the same amino acid
- Nonsense: introduction of premature stop codon
- Missense: change of amino acid

Nonsense and missense mutations can render the protein non-functional or they can alter the original function in some way.

A *forward genetics* approach involves the application of chemical mutagen to wild type seeds to acquire plants with desired phenotypes. These treatments cause either nucleotide substitutions or deletions and are transmitted to future generations. Forward screens of poppy mutants may identify a novel phenotype. However, locating the gene (and mutation) responsible would prove difficult due to the absence of genetic maps and genomic sequence information for opium poppy, resources which have been used in a process of chromosome walking in other species. Alternatively, microarray analysis and comparative EST profiling could be used to link a gene with a phenotype.

Alternatively, a *reverse genetics* approach requires mutant populations for the detection of plants with specific genotypes rather than phenotypes. Technologies such as TILLING (Targeting Induced Local Lesions in Genomes; Comai and Henikoff, 2006; Kurowska et al., 2011; Till et al., 2003) are deployed to identify point mutations in genes of interest. Once

mutations are stabilised, any changes in phenotype may be attributed to the induced mutation.

1.8.2.1 Previous mutagenesis of opium poppy

Various groups have used chemical mutagens and ionising radiation on opium poppy in the past (Chauhan and Patra, 1993; Floria and Ghiorghita, 1980; Sharma et al., 1999). This approach led to the production of non-narcotic opium-free 'Sujata' from a narcotic crop (Sharma et al., 1999). It had use as a food crop seeds rich in protein and unsaturated oil. Indian researchers also identified a high thebaine variant in a mutant population and discussed the possibility of using it in tissue culture systems for the production of thebaine (Chaterjee et al., 2010).

Australian researchers found that EMS mutagenesis of opium poppy resulted in plants with increased yields of thebaine and oripavine (Fist et al., 2000; Millgate et al., 2004). M1 plants were selfed and the resulting M2 seed was sown. The latex of 18,000 M2 plants was screened for alkaloid content. One M2 plant, designated "Norman", was identified which was free of morphine and codeine and accumulated increased levels of thebaine and oripavine. The resulting *top1* (meaning thebaine, oripavine poppy 1) poppy mutant caused a change in the industry in Tasmania whereby thebaine and oripavine could be efficiently produced from morphine-free crops. The segregation of 375 F2 individuals was Mendelian and consistent with the involvement of a single gene (Millgate et al., 2004). A 17,000 element microarray was queried for changes in gene expression in the mutant. Ten transcripts were differentially suppressed in *top1*. However, none of the corresponding genes appeared to be involved in morphine biosynthesis. The *top1* phenotype may be caused by an alteration of an as yet unidentified transport or structural component, preventing the correct inter- or intra-cellular compartmentalisation of thebaine and oripavine.

Another Canadian group used EMS mutagenesis on poppy seeds and from 10,000 M1 plants grown to maturity two interesting mutant phenotypes were identified-one

accumulating reticuline and the second curiously accumulating the antimicrobial sanguinarine in the latex (Desgagne-Penix et al., 2009).

Fast neutron mutagenesis (FNM) of thebaine/oripavine TOP1 material has been used to produce 21 thebaine only plants after a screen of approximately 35,000 M2 plants (Fist et al., 2013). Fifty-two high thebaine selections (including 9 of the 21 having substantially no oripavine, morphine or codeine) were cross-pollinated with 3 different morphine-containing poppy lines. The F2 generation of plants from these crosses were then tested, and many were found to contain high levels of codeine, variable levels of thebaine and substantially no oripavine or morphine.

They also produced plants having codeine as the predominant alkaloid and substantially no oripavine, morphine or thebaine. These plants resulted from crossing the thebaine only plant with plants containing morphine. It is unclear why the new plants accumulate codeine, but the authors postulate that the step between codeine and morphine (CODM) has been substantially blocked.

They suggest that the gene controlling the step between thebaine and oripavine (T6ODM) that appears to have been blocked in high thebaine poppy is also responsible for the conversion of codeine to morphine (CODM). Therefore, when the gene responsible for producing thebaine-only poppies in conjunction with the TOP1 or Norman mutation is introduced into plants lacking that mutation, a codeine phenotype is produced.

There is no evidence that TILLING (for detection of SNPs introduced by EMS) or de-TILLING (for detection of deletions caused by FNM) approaches were subsequently used to identify the molecular basis of these phenotypes. However, their characteristic phenotypes have been patented by Tasmanian Alkaloids and have led to their commercial thebaine and codeine crops (Fist et al., 2013, 2000).

1.9 Targeting Induced Local Lesions in Genomes (TILLING)

TILLING is a reverse genetics technique originally developed as a service to the *Arabidopsis* community whereby allelic series of EMS-induced point mutations could be ordered by researchers (*Arabidopsis* TILLING project; ATP; Till et al., 2003). At its simplest, TILLING allows for high throughput screening of pooled gDNA of mutagenised populations to identify mutations in genes of interest. Originally, the TILLING method relied upon denaturing HPLC to detect mutations (McCallum et al., 2000). The identification of single-strand-specific endonucleases that cleave DNA at the site of mismatched nucleotides in otherwise double stranded DNA allowed for much higher throughput and the eventual proliferation of TILLING as a reverse genetics tool (Colbert et al., 2001). Mismatches form a heteroduplex 'bubble' or 'bulge' that is recognised by these enzymes and nicked. The endonuclease CEL1, extracted from celery (*Apium graveolens*), was used in the method published by the ATP (Colbert et al., 2001; Till et al., 2003). In fact celery juice extract (CJE) was often successfully used in place of relatively expensive purified CEL1 enzyme (Till et al., 2004). Figure 9 in Chapter 2 provides an overview of the process used in this project.

1.9.1 Choice of TILLING platforms

CEL1 cleaves at the sites of induced mutations in heteroduplexes formed after PCR with fluorescently labelled primers. For many years, LI-COR DNA analysers were the platform of choice for separating out DNA fragments of 96 pools in a single run (i.e. 384 plants or 768 plants if pooled 4 fold or 8 fold respectively). The forward and reverse primers are labelled with fluorescent dyes that are detected at 700 nm and 800 nm, respectively. Two electronic images files are produced per gel run-one from each of the 700 nm and 800 nm channels (Till et al., 2006). Mutation detection is carried out using a programme called GelBuddy (Zerr and Henikoff, 2005).

Researchers undertaking TILLING with *Medicago truncatula* offered an alternative mutation detection platform when they demonstrated separation of CJE-cut PCR fragments with the Applied Biosystems™ 3730xl (Le Signor et al., 2009). The method also relies on the use of fluorescently labelled primers in PCR. This method was adopted for the RevGenUK

TILLING service offered to the *Medicago truncatula*, *Brassica rapa* and *Lotus japonica* communities (Stephenson et al., 2010).

Avoiding the use of expensive labelled PCR primers, a modified TILLING approach was demonstrated in wheat using non-denaturing polyacrylamide gels with ethidium bromide staining for mutation detection rather than LI-COR systems (Uauy et al., 2009). Low budget TILLING of a rice mutant population was demonstrated with unlabelled PCR primers and a simple agarose gel based detection assay (Raghavan et al., 2007).

Changes in melting temperatures of PCR fragments conferred by induced mutations can be detected using high-resolution melting (HRM;Botticella et al., 2011; Dong et al., 2009; Gady et al., 2009; Li et al., 2010; Ririe et al., 1997; Zhou et al., 2005). The technique is limited to small amplicons of less than 400bp but it can be useful to target specific exons for example.

In recent years, with the advent of more affordable next generation sequencing platforms, TILLING-by-sequencing has been demonstrated successfully. Illumina™ sequencing of target genes amplified from multidimensionally pooled DNA templates allowed the identification of mutants in mutant populations of rice and wheat (Tsai et al., 2011). Pools of PCR products are given molecular barcodes with the addition of DNA adapters allowing for multiple target genes or even amplicons from different species in a single next generation sequencing run. Sophisticated bioinformatics pipelines deconvolute samples after sequencing based on the barcodes (Missirian et al., 2011). Mutations are identified by comparison to a wild type reference. Tilling by sequencing could be advantageous in polyploid crops such as *Brassica napus* where it is often difficult to design primers to amplify specific orthologues (Wells et al., 2013). Conventional TILLING approaches with polyploid crops are often complicated by mixed amplicons. This results in CEL1 cleavage of every mismatch position throughout the amplicon making it more difficult to detect induced mutations. Like HRM, the technique is limited to relatively small amplicons i.e. typical sequence read lengths of the next generation sequencing platform used. Conventional TILLING with gel-based separation allows for analysis for much larger amplicons (~1.5 kb on LI-COR DNA analysers).

1.9.2 Applications of TILLING outside of model plants

TILLING is appropriate to many crop species especially in Europe where work with genetically engineered crops is restricted and in parts of the world without the regulatory infrastructure for work on such crops. Interesting mutant alleles detected by TILLING can be introgressed into different backgrounds. In recent years there has been various examples of the utilisation of TILLING as a breeding tool for agronomically important crops. In bread wheat, *Triticum aestivum*, for example, TILLING mutants of the starch branching enzyme II (SbeIIa) genes resulted in higher amylose content (Botticella et al., 2011; Slade et al., 2012). TILLING mutants of Δ12 and Δ15 desaturase genes of hemp, *Cannabis sativa*, produced major alterations in seed fatty acids. One mutant produces particularly high levels of high oleic hemp oil (Bielecka et al., 2014). A TILLING mutant of the waxy gene of potato, *Solanum tuberosum*, was used to establish elite lines lacking granule-bound starch synthase I activity and producing high-amylopectin starch (Muth et al., 2008). TILLING has also been successfully applied to tomato, *Solanum lycopersicum*, where mutants of an ethylene receptor gene, Sletr1, resulted in delayed fruit ripening and a longer shelf life (Okabe et al., 2011), and, mutants of the DET1 gene resulted in elevated levels of both carotenoid and phenylpropanoid phytonutrients in ripe fruit (Jones et al., 2012).

1.10 Summary and aims

Morphine, codeine and thebaine are part of a large and diverse group of benzylisoquinoline alkaloids made by *Papaver somniferum*. Opiates from this plant are used to manufacture a range of medicines for use in pain relief, vasodilation, antimicrobials, anticancer agents, antitussives and the treatment of addiction. Morphine represents the largest output from the poppy industry but much of morphine produced each year is converted back to codeine, which is less potent and so has more versatile uses. BIA biosynthesis begins with a condensation reaction between two derivatives of the amino acid tyrosine yielding (S)-norcoclaurine. Subsequent 6-O-methylation, N-methylation, 3'-hydroxylation, and 4'-O-methylation results in the formation of the central branch point intermediate, (S)-reticuline. Biosynthesis of morphine proceeds from here via epimerisation

to (R)-reticuline catalysed by STORR (Winzer et al., 2015). Within the morphinan subgroup two 2-ODDs and a codeinone reductase are involved in the conversion of thebaine to morphine in a bifurcated pathway.

Prior to the commencement of this PhD project chemical mutagenesis had been carried out on high morphine commercial poppy variety, designated HM2 (Winzer et al., 2015). The mutant population was generated by colleagues. The chemical mutagen, ethyl methanesulfonate (EMS), was used to induce random point mutations. M1 plants were sown and selfed in York and M2 plants were screened for desirable alkaloid profiles in field trials in Tasmania with the help of our industrial collaborators GSK Australia. However, a potential problem with the forward genetic approach is that mutations in key genes in a heterozygous state will not result in phenotypes in the M2 if mutant alleles are recessive. A reverse genetic approach, that identifies individuals with mutations in key genes involved in alkaloid biosynthesis, could overcome this problem, since mutations could be brought to homozygosity in subsequent generations and assessed for changes in alkaloid profiles. Hence, a reverse genetic approach was adopted as the main strategy in this thesis work. The first target gene would be CODM, with the aim of identifying mutations that would contribute to a high codeine phenotype. The second target would be T6ODM, with the aim of creating high thebaine/oripavine lines. The existing mutant population described above would be screened for CODM and T6ODM mutations. Any mutants found would be assessed for changes in their alkaloid profiles i.e. a reduction in morphine content and increases in codeine, thebaine and oripavine. Any new varieties with desirable alkaloid profiles may reduce the need for chemical processing of opiates and spraying of poppy crops. The approach would also provide molecular markers for breeding.

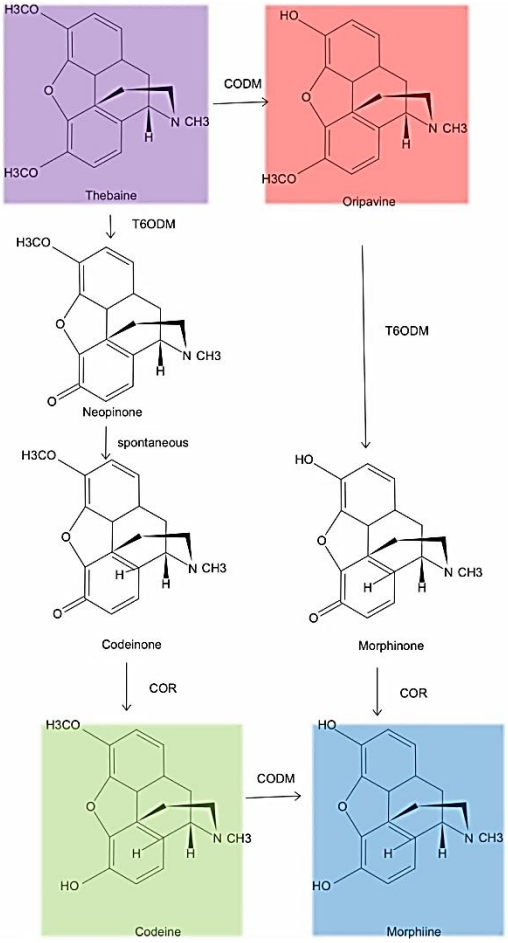


Figure 7. Final steps in the biosynthesis of morphine from thebaine catalysed by the T6ODM, CODM and COR enzymes.

GSK is the largest producer of morphine in the world. However, each year the majority of the morphine crop goes to support the manufacture of codeine (O-methylation of morphine). CODM was selected as a candidate gene for reverse genetics with the aim of developing an improved *Papaver somniferum* line with higher yields of codeine. This gene is involved in the final step of morphine biosynthesis (Figure 7) and silencing of CODM by VIGS has been shown to increase the yields of codeine to morphine (Farrow and Facchini, 2013; Hagel and Facchini, 2010a; Wijekoon and Facchini, 2012). Plants with compromised CODM activity would allow codeine to be produced without the need to grow morphine poppies and the expensive and time-consuming industrial-scale methylation of the resulting morphine. Codeine is more soluble than morphine and easier to purify. Therefore, the yield, quality and cost of natural codeine obtained from a codeine poppy would all be improved relative to synthetic codeine obtained from morphine. Of course this would depend on the ability of the CODM mutant plant to accumulate codeine to levels at least as high as current morphine lines.

Likewise, in the past 15 years there has been increasing worldwide demands for two other precursors of morphine, namely thebaine and oripavine. T6ODM was selected as a candidate gene for reverse genetics with the aim of developing an improved *Papaver somniferum* line with higher yields of thebaine and oripavine. This gene is involved in the final steps of morphine biosynthesis and silencing of T6ODM by VIGS has been shown to increase the yields of thebaine (Farrow and Facchini, 2013; Hagel and Facchini, 2010a; Wijekoon and Facchini, 2012).

The availability of the gene sequences of CODM and T6ODM (Hagel and Facchini, 2010a) provided an opportunity for a reverse genetics approach to try and identify alleles which would result in similar phenotypes which could be stably transmitted. The initial aim of this project then was to screen the already existing HM2 mutant population for mutations in the genes for these 2-ODDs. The starting HM2 seed produces plants which accumulate predominantly morphine in their capsules and only approximately 3% (of total alkaloids) of each of codeine and thebaine. DNA from a subpopulation (~4000 M2 plants) of the larger mutant population grown in Tasmania was available for the reverse genetic screen. Any mutants identified would then be assessed for altered yields of the highly valuable opiates codeine, thebaine and oripavine. Identification of mutations which would terminate biosynthesis of morphine at a valuable precursor would be an ideal output. Morphine-free single alkaloid crops would also lead to more straightforward extractions from raw material and remove the need for several hazardous steps in the chemical processing of opiates. A thebaine-oripavine only variety with high yields would avoid the need to spray commercial crops with chemicals for the enhancement of thebaine production. Bespoke varieties created by reverse genetics would have the associated molecular markers that could be used in future molecular breeding programmes.

Chapter 2- Methods

2.1 Characterisation of gene targets

Complementary DNA sequences for opium poppy CODM (GenBank accession: GQ500141), T6ODM (GenBank accession: GQ500139) and PODA/DIOX2 (GenBank accession: GQ500140) provided the starting point for a characterisation of these genes in GSK commercial lines (Hagel and Facchini, 2010a). Previous work in our group involved 454 pyrosequencing on complementary DNA libraries derived from stem and capsule tissue from two morphine accumulating cultivars, HM2 and HM1, a noscapine accumulating cultivar, HN1, and a thebaine accumulating cultivar, HT1. The resulting expressed sequence tag (EST) database was queried with the cDNA sequences of CODM and T6ODM. Matching ESTs and the published cDNA sequences of CODM and T6ODM were aligned and used to design primers in an effort to obtain the genomic DNA sequence for both T6ODM and CODM. This information would be fundamental for the design of appropriate primers for a reverse genetic screen for mutants.

At the beginning of the project probes for T6ODM and CODM were prepared to screen the same bacterial artificial chromosome (BAC) library in which the noscapine gene cluster was identified and characterised (Winzer et al., 2012). They successfully pulled out a number of CODM- and T6ODM-containing BACs. There was no overlap i.e. probes for each target did not have hits with any one BAC. Sequencing of CODM and T6ODM containing BACs was attempted at GSK and Amplicon Express (a company that offers BAC library preparation and sequencing services; Pullman, WA) using various next generation sequencing platforms. The assembly of short sequence reads into useful contigs was complicated due to the presence of repetitive DNA. While these efforts were ongoing, attempts were made to amplify and clone the genes from gDNA of opium poppy (Chapter 3). Sequencing of inserts would reveal the sequences of introns and enable appropriate primer design for TILLING.

2.1.1 Cloning of full length PCR products and sequencing

PCR products generated with PfuUltraII HS DNA polymerase (Aglient, 600670) were cloned with the StrataClone Blunt PCR Cloning Kit (Aglient, 240207) according to the manufacturer's instructions. Briefly, the Strataclone blunt PCR cloning vector mix contains two blunt-ended DNA arms, each charged with topoisomerase I on one end and containing a *loxP* recognition sequence on the other end. Blunt-ended PCR products are efficiently ligated to these vector arms in a 5min ligation reaction by topoisomerase I-mediated strand ligation. The resulting linear molecule is then transformed into a competent cell line engineered to transiently express *Cre* recombinase. *Cre*-mediated recombination between the vector *loxP* sites creates a circular DNA molecule that is proficient for replication in cells growing on media containing ampicillin or kanamycin.

PCR products generated with GoTaq Flexi DNA polymerase (Promega, M8305) were cloned with the StrataClone PCR Cloning Kit (Aglient, 240205) according to the manufacturer's instructions because of the 3'-A overhangs generated using PCR. The PCR cloning kit is similar to the Blunt PCR cloning kit except that the topisomerase-charged ends have a modified uridine (U*) overhang. *Taq* amplified PCR products, which contain 3'-adenosine overhangs, are efficiently ligated to these vector arms in the 5min ligation reaction, through A-U* base-pairing followed by topoisomerase I-mediated strand ligation. Transformation proceeds in the same manner as the Blunt PCR kit.

M13 colony PCR was performed by preparing 20 µl of mastermix in wells of a 96 well plate. A sterile toothpick was used to take up a colony of bacteria from a petri dish and it was placed in the appropriate well of the 96 well plate (containing PCR mastermix). The toothpick was agitated to remove the bacterial cells. Once all samples were loaded on the plate, PCR took place immediately.

20 µl mastermix for each colony contained:
0.125 µl GoTaq Flexi DNA polymerase (Promega, M8305)
1.2 µl $MgCl_2$ (25 mM)
4 µl 5x GoTaq Flexi green buffer
2 µl dNTPs (2 mM) (New England Biolabs, N0447)
1 µl of each M13 primer (10µM)
10.675 µl HyClone molecular biology grade water (Thermo Scientific, SH3053801).

M13 PCR conditions were as follows:
95°C 2 min,
30 cycles of (95°C for 30 sec, 55°C 30 sec, 72°C 3 min)
72°C 5 min.

PCR products were visualised on a 1.5% w/v agarose gel. Selected colonies were inoculated in 5 ml of fresh LB medium containing kanamycin (50 mg/L) and incubated overnight before plasmid extraction.

2.1.2 Plasmid extraction

Cloned plasmids were extracted using the QIAprep Spin Miniprep Kit (Qiagen, 27104) according to the manufacturer's instructions. Briefly the kit is based on alkaline lysis of bacterial cells followed by adsorption of DNA onto silica in the presence of high salt. Washing and elution steps result in high yields of plasmid DNA. Recovered plasmids were quantified with the nanodrop (Thermo Scientific) and sequencing reactions were set up using the procedure indicated below.

2.1.3 PCR product purification

PCR products were purified using the QiaQuick PCR purification kit (Qiagen, 28104) according to the manufacturer's instructions. Briefly, the kit makes use of spin column technology for the purification of PCR products. DNA adsorbs to the silica membrane in the presence of high concentrations of salt while contaminants pass through the column. Impurities are efficiently washed away and DNA is eluted with water. 30 µl PCR product was added to 150 µl of buffer PB as a starting point and purified PCR products were eluted with 30 µl of HyClone molecular biology grade water (Thermo Scientific) by adding the water to the centre of the Qiaquick membrane, letting the column stand for 1min, and then centrifuging at 13,000 rpm for 1 min.

2.1.4 Preparation of samples for sequencing

2.1.4.1 University of York sequencing service

1µl primer (3.2µM) and purified PCR product were deposited premixed in a volume of 7.2µl for Sanger sequencing. Enzyme and buffer were added by the Genomics lab to take the final volume to 10µl. The quantity of template required was calculated using the following guidelines:

PCR product	Final concentration per reaction
500-1000 bp	10-20 ng
1000bp-2000 bp	20-40 ng
>2000 bp	30-50 ng
Plasmids	200-300 ng

Samples were submitted in 0.2 ml domed capped PCR tubes (Axygen, PCR-02D-C) or alternatively in 96-well plate format (Starlabs, I1402-9800).

2.1.4.2 Eurofins sequencing service

17 µl of premix was prepared containing 2 µl of 10 µM primer and 15 µl of purified PCR products. If PCR products were between 300-1000 bp 75 ng of DNA was added to premix. For PCR products between 1000-3000bp 150 ng of DNA was added to premix.

2.1.5 BAC isolation

Primer sequences were passed to Amplicon Express to allow screening of a bacterial artificial chromosome (BAC) library made from a GSK noscapine variety, HN1, for CODM and T6ODM containing BACs. Two of the hits were chosen for further study here-namely PS_158A11 and PS_33D07 (both hits for CODM). A tip was placed in the appropriate wells of the 384 well plates and streaked across a LB agar plate containing 12.5 µg/ml chloramphenicol. Plates were incubated at 37°C overnight. One colony was picked and streaked across another plate and incubated overnight again at 37°C. A single colony was then placed in a starter culture of 2.5 ml LB medium (12.5 µg/ml) chloramphenicol) and incubated for 16 h at 37°C at 300 rpm. Cells were harvested at 6000 x *g* for 15 min at 4°C. The Qiagen large construct kit (Qiagen, 12462) was then used for BAC isolation according to the

manufacturer's instructions. The ATP required for the exonuclease digestion step was obtained from Sigma (A3377). BAC DNA was quantified with the nanodrop.

The work displayed in Figure 15 and Appendix J took place with BAC DNA isolated by Amplicon Express.

2.1.6 LongAmp PCR

The LongAmp *Taq* PCR kit (New England Biolabs, E5200S) was used to investigate whether DIOX genes were physically linked to one another. Successful long range PCR between primers designed with different DIOX cDNA sequences would indicate linkage between the genes and an estimation of how close they are.

50 µl LongAmp PCR reactions were set up as follows:
20 µl HM2 gDNA template (10 ng/µl)
10 µl 5x LongAmp Buffer
1.5 µl dNTPs
2 µl LongAmp Taq
12.5 µl dH_2O
2 µl forward primer (10 µM)
2 µl reverse primer (10 µM).

The PCR programme used was 94°C 30 sec, 30 cycles of (94°C 30 sec, 65°C 13 min 30 sec), 65°C 10 min. PCR products were visualised on 0.6% agarose gel run at 100 V for 3 h 30 min.

2.1.7 RNA extraction

Each frozen stem sample was ground under liquid nitrogen in a mortar. Approximately 200 mg of powderised material was collected in a 2 ml eppendorf and weighed. RNA was extracted according to a method developed for extracting RNA from pine trees with slight modifications (Chang et al., 1993).

Extraction buffer consisted of:
2% CTAB (hexadecyltrimethylammonium bromide) (Sigma, M52365)
2% PVP (polyvinylpyrrolidinone K30) (VWR, A2259)
100 mM Tris-HCl (pH 8.0) (Sigma, 93363)
25 mM EDTA (Sigma, E8008)
2.0 M NaCl
0.5 g/l spermidine (Sigma, 49761)
(all mixed and autoclaved before use).

450 µl of extraction buffer and 9 µl of 2% beta-mercaptoethanol (Sigma, M3148) was prepared for each 100 mg of sample and heated to 65°C. 450 µl was added to each 100 mg of sample and vortexed to fully homogenise. Samples were then incubated at 65°C for 10 min. The first extraction occurred with equal volumes of chloroform:isoamyl alcohol (24:1; $CHCl_3$:IAA). After it was added the samples were vortexed for 2 min to mix and then spun at 10,000 rpm/10 min/RT. The supernatant was transferred to a fresh 2 ml Eppendorf and extracted with equal volumes $CHCl_3$:IAA at 10,000 rpm/10 min/RT. This was repeated for a third time before 600-640 µl of supernatant was collected. One quarter volume of 10 M LiCl was added and it was precipitated overnight at -20°C.

Following LiCl precipitation RNA was centrifuged at 12,000 rpm/20 min/4°C and the supernatant was discarded. The pellet was dissolved in 500 µl warmed SSTE. SSTE was made from 1.0 M NaCl, 0.5% SDS, 10 mM Tris-HCl (pH 8.0) and 1 mM EDTA (pH 8.0).

After another extraction with one volume of 500 µl $CHCl_3$:IAA the supernatant was transferred to a fresh 1.5 ml Eppendorf tube and two volumes of 100% ethanol was added and mixed gently. The RNA precipitated at -20°C for 2 h.

To harvest the precipitated RNA, the sample was spun at 10,000 rpm/20 min/4°C. The supernatant was carefully and completely removed, the pellet washed with 75% ethanol and then air dried for 2-3 min at RT. The pellet was then dissolved in 50 µl of nuclease-free dH_2O and tapped gently for 5 min. It was then incubated at 65°C for 5 min. Crude RNA was then quantified with a nanodrop (Thermo Scientific).

RNA was further purified with the RNeasy Plus Micro Kit (Qiagen, 74034). 45 µl of crude RNA was added to 5 µl of water and 350 µl Buffer RLT plus and mixed well by pipetting. The sample is then passed through a gDNA Eliminator spin column. This, together with the high-salt buffer, ensures removal of gDNA. Ethanol is added to the flow-through to provide appropriate binding conditions for RNA, and the sample is then applied to the RNeasy MinElute spin column, where total RNA binds to the membrane and contaminants are

efficiently washed away. With this procedure RNA molecules larger than 200 nucleotides are purified. Purified RNA was then quantified on the nanodrop (Thermo Scientific).

2.1.8 cDNA preparation

RNA was reverse transcribed with the Quantitect Reverse Transcription Kit (Qiagen, 205310) according to the manufacturer's instructions. Briefly, the Quantitect Reverse Transcriptase is a blend of omniscript and sensiscript Reverse Transcriptases, which has a high affinity for RNA and is capable of cDNA synthesis from a wide range of RNA amounts in 20min. Purified RNA is added to a master mix prepared from Quantitect Reverse Transcriptase, Quantitect RT buffer and RT Primer Mix. RT Primer Mix ensures cDNA synthesis from all regions of RNA transcripts, even from 5' regions. The reaction takes place at 42°C and is inactivated at 95°C.

2.2 Sample tracking

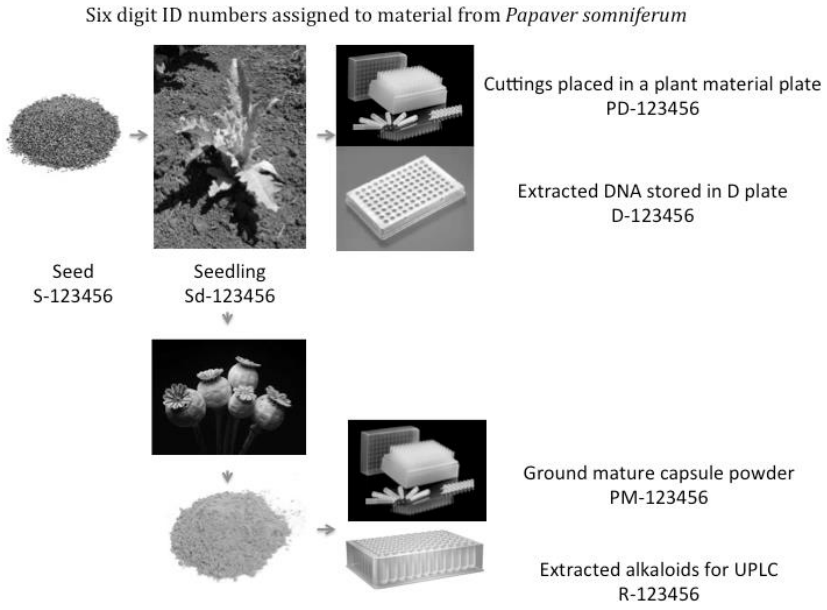


Figure 8. Overview of the sample tracking process used in the project. Seed batches are assigned an S number. The resulting seedlings are given Sd numbers. At 4 weeks old, a small section of leaf is placed into a well of a PM collection plate. DNA is extracted and placed into the corresponding well of a D plate. Mature capsules split open. The seed is assigned an S number for the next generation and the remaining capsule material is ground into a fine powder and approximately 10mg is placed in a PM plate for alkaloid extraction. Running plates for UPLC analysis are given an R number. Links at each stage are tracked using apps on a barcode reading PDA.

Figure 8 provides a graphical overview of the sample tracking procedures in the project. For the HM2 TILLING population each M2 seedling had the same "Sd" number as the M2 seed batch from which it was derived i.e. any mutation identified in Sd-123456 could be recovered

by sowing siblings from seed batch id S-123456. For subsequent generations "S" numbers were assigned to collected seed in an unordered fashion but tracked in a database.

2.3 Reverse Genetic Screen

The original EMS mutagenesis of HM2 seed and extraction of DNA from the M2 seedlings was carried out by Thilo Winzer, Marcelo Kern and Lynda Sainty. During the project the method was used again to mutagenise seed derived from one of the M2s in the original TILLING population which carried a CODM Q254* knockout mutation.

2.3.1 Ethyl methanesulfonate (EMS) mutagenesis

Seeds for mutagenesis were weighed out. As a guide 1000 Australian Blue seeds were taken to weigh approximately 0.5325 g. Seeds were soaked in 0.1% (w/v) potassium chloride solution (Sigma, 60131) for 24 h on a rocking platform using a 10ml universal tube for 1000 seeds and a 50 ml Falcon tube for 2000 or more seeds. The moving solution ensured all seeds in the tube were agitated.

On the day of mutagenesis, the potassium chloride was poured off from the soaking seeds. A disposable fine pipette was used to remove the last of the liquid. In the fume hood 105 µl of EMS (Sigma, M0880) and 500 µl of DMSO (Sigma, 276855) were added to 5 ml of 0.2M $NaH_2PO_4.2H_2O$ (Sigma, 71505, pH to 5.0 with NaOH) and 4.395 ml of H_2O to produce the 10 ml 100mM solution needed for every 1000 seeds. This could be adjusted for different total volumes and/or concentrations of EMS. The mixture was shaken well to form an emulsion. It was added to the seeds using a disposable pipette which was then placed with any other solid waste in a container containing 2 l of deactivating solution (10% sodium thiosulphate; Sigma 72049). A control mixture lacking EMS was also prepared. Tubes were placed on the rocking platform to be agitated by the moving solution thus ensuring adequate EMS treatment.

Enough sodium thiosulphate to produce a 10% solution with all the liquid waste that would be produced was weighed out and placed in a large conical flask. After EMS treatment,

the EMS solution was poured off into the waste flask. A disposable fine pipette was used to remove the last of the liquid from around the seeds.

The seeds were washed twice in deactivating 100mM sodium thiosulphate solution by placing the tubes back on the rocking platform for 15 min. The seeds were then washed twice with distilled water in the same manner before drying them overnight on 3 mm paper in large Petri dishes.

2.3.2 TILLING populations

The letters assigned to TILLING subpopulations refer to the dose of and length of exposure to EMS: B=3 h, 200 mM EMS and C=5 h, 200 mM EMS. HM2 received the B and C treatments by group members prior to the commencement of this PhD. The lowest germination rates and M1 survival rates were observed in the C subpopulations due to the longer exposure time to EMS. M1 seedlings were sown in the glasshouse in York. All M2 seedlings were grown in the field in Tasmania (forward screen) and a subsection of the B and C populations (~4000) were grown in the glasshouse in York which allowed extraction of M2 DNA (Table 1). These M2 seedlings were sacrificed after DNA sampling. Resowing and identifying the mutation in siblings would recover any mutations identified.

Each M2 seed family was assigned a unique S number and the same Sd number was assigned to the resulting seedling from which DNA was extracted. DNA was normalized to 10 ng/µl for all 45 DNA plates. DNA was pooled fourfold for PCR by mixing equal amounts of four DNA samples in new pool plates. The pooling strategy is outlined in Appendix A.

M5 CODMb Q254* mutant seed also received the B and C treatments. M1 seedlings were grown in the glasshouse in York. M2 seedlings were grown in the field in Tasmania (forward and reverse screens) and M3 seed was collected (Table 2).

Population	M1 pop size (% germination freq.)	No of M2 seed lines	M2 seed batch ids	No of germinating M2 seed lines (% germination)
B2	3670 (20%)	3197	S-100001-S-102800, S-102831-S-103227	2826 (88%)
C1	750 (12.5%)	419	S-103228-S-103646	339 (56%)
B3	5600 (31%)	1582	S-103647-S-105228	981 (62%)

Table 1. Subpopulations of the original HM2 mutant population that were screened for mutations by TILLING

Population	M1 pop size (% germination freq.)	No of M2 seed lines	M2 seed batch ids
B 5931 seeds (200mM EMS, 3h)	1200 (20.2%)	459	S-186857-S-187316
C 5086 seeds (200mM EMS, 5h)	245 (4.8%)	178	S-187317-S-187494
Control 375 seeds (EMS-free solution, 3h)	111 (29.6%)	45 (17 sown)	S-186811-S-186856

Table 2. Subpopulations of the new CODMb Q254* mutant population. M2s were assessed in a forward screen in Tasmania in the 2014/2015 growing season.

2.3.3 Plant growth conditions

Following EMS treatment dried seeds were sown immediately on poppy mix (John Innes no 2 compost, vermiculite and perlite (4:2:1)) in Maxi (Fleet) Rootrainers™ (Haxnicks, Mere, UK) and then covered with a thin layer of compost. Seeds were kept moist by covering the trays with a lid with the vents open. The lid was removed once shoots appeared. Plants were grown under glass in 16 h days at the University of York horticulture facilities.

For analysis of subsequent generations and crosses plants were grown under glass using the same Rootrainers and conditions described above for M2 seedling emergence. Plant lines assessed in field trials was grown at the GlaxoSmithKline Australia field-trial site, Latrobe, Tasmania from September 2014 to February 2015.

2.3.4 DNA extraction

Plant material for DNA extraction was taken from young leaves (approximately 4 weeks). 30-50 mg was placed in the bottom of a collection microtubule (Qiagen, 19560). Samples were stored at -20°C until processing. DNA was extracted in a 96-well plate format using the Qiagen Biosprint 96 Plant kit (Qiagen, 941558) and the Tissue Lyser (Qiagen). Briefly, with this automated system DNA is extracted from lysate and transferred between wash plates by

means of magnetic DNA-binding beads. In the elution plate, DNA dissociates from the beads and dissolves in TE.

One 3 mm tungsten-carbide bead was added to each well of the microtubule collection plate containing plant material. 300 µl of buffer RLT (lysis buffer) was added and the tubes were sealed with collection microtube caps (Qiagen, 19566). Each rack of collection microtubules was placed between the adapter plates and fixed in to the Tissue Lyser clamps. Samples were homogenised for 90 sec at 25 Hz, allowed to rest for 60 sec, then homogenised again for 90sec. The adapter set was then disassembled before reversing rotating the rack 180° and reassembling. The adapter sets were placed back into the Tissue Lyser for two more homogenisations. Each rack was then centrifuged at 4350 rpm for 15 min while the wash, elution and lysate plates were prepared.

Wash 1	Buffer RPW (reconstituted)	500 µl
Wash 2	100% ethanol	500 µl
Wash 3	100% ethanol	500 µl
Wash 4	0.02% Tween 20	500 µl
Elution	TE	200 µl
Lysate	100% isopropanol + MagAttract Suspension G	200 µl + 20 µl respectively

200 µl of lysate from each sample was transferred to the S-block containing isopropanol and magnetic beads taking care not to carry over any plant material. The 96 well format and sample order was retained during all procedures and transfers. All plates were loaded on to the Biosprint 96 machine for DNA extraction. The BS96 DNA plant protocol was used. The elution plate was covered with a self-adhesive film after processing.

DNA from the material grown in Tasmanian field trials in the 2014/2015 growing season was extracted by the Australian Genome Research Facility in Adelaide, SA.

2.3.5 Normalisation of DNA plates

Following DNA quantification on the nanodrop (Thermo Scientific) DNA samples were normalised to 10 ng/µl concentration using HyClone molecular biology grade water (Thermo Scientific, SH3053801).

The DNA extractions for the original HM2 mutagenised population was carried out by group members prior to the commencement of this PhD. However, twelve new pool plates were created from the DNA storage plates prior to TILLING CODM and T6ODM. Normalised DNA plates were pooled 4 fold for PCR amplification. Refer to Appendix A for the pooling strategy.

DNA from plants in subsequent generations was also normalised to 10 ng/µl prior to genotyping. DNA concentrations of samples from the Tasmanian field trials were quantified at the Australian Genome Research Facility. Working DNA plates with 10 ng/µl concentrations of each sample were prepared in York prior to genotyping.

2.3.6 Primer design for TILLING

Primers for TILLING were designed using web-based Primer 3 Plus (Untergasser et al., 2007) making use of intron sequences and identified polymorphic sites obtained in initial sequencing work to target specific copies of CODM and T6ODM. Optimum annealing temperatures for each primer pair were identified in temperature gradient PCR. Touchdown PCR was used to ensure specificity.

For CODM it was decided to screen exons 2, 3 and 4 in all three identified copies of CODM simultaneously using primer pair F9/R5 (Figure 16). For T6ODM, where more numerous copies were detected, information obtained from sequencing was used to exploit a polymorphic site to design primers (F14/R27) that would specifically amplify the copy of T6ODM with an exact match to the published cDNA sequence (Figure 32).

2.3.7 Targeting Induced Local Lesions IN Genomes (TILLING) method

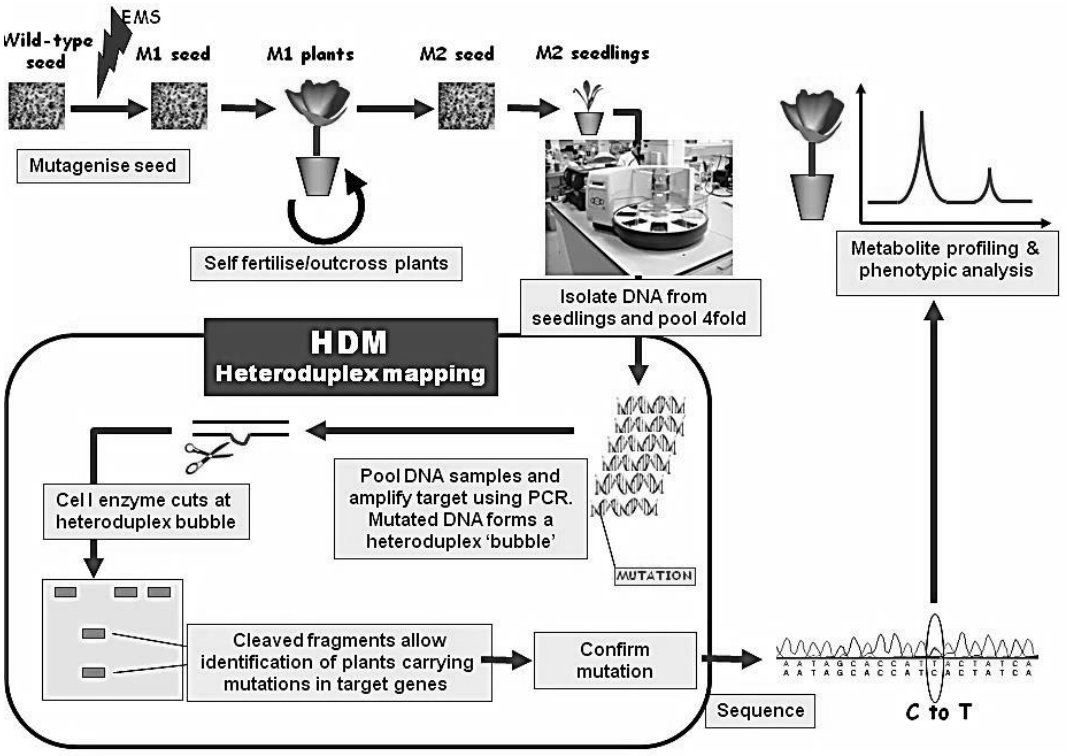


Figure 9. Overview of the TILLING method. Wild type seed is mutagenized with EMS. M1 plants are allowed to self. DNA is collected from M2 seedlings and pooled four fold for PCR with primers for the target gene. PCR products are heated to separate fragments and left to reanneal with any complementary strand in the sample i.e. to form heteroduplexes. Differences in gene sequence between members of the pool would result in heteroduplex bubbles at the point of the mismatch. The sample is then treated with an enzyme that recognises and cleaves at the sites of heteroduplexes resulting in two fragments the sum of the sizes of which is equal to the size of the full length PCR product. Separation of fragments out on an agarose gel enables mutation detection. Prospective mutations are confirmed by sequencing after steps are taken to identify the individual in the pool of 4 that is carrying the mutation. The final step is to assess whether the mutation has the desired effect on the phenotype.

2.3.7.1 Amplification and heteroduplex formation of BBE PCR products for validation of Fragment Analyser as a mutation detection platform

BBE was amplified from gDNA pools in nested PCR using the primers listed in Appendix K. In the first round 10 µl PCR reactions were set up with 0.1 µl GoTaq Flexi DNA polymerase (Promega, M8305), 0.6 µl $MgCl_2$ (25 mM), 2 µl GoTaq Flexi colourless buffer, 0.1 µl dNTPs (20 mM; New England Biolabs, N0447), 0.2 µl of each primer (10 µM), 4.3 µl HyClone molecular biology grade water (Thermo Scientific, SH3053801) and 2.5 µl (25ng) of template DNA. PCR conditions were as follows:

94°C 2 min,
20 cycles of (94°C for 20 sec, T_{ann} of 65°C (-0.5°C per cycle) 30 sec , 72°C 2 min)
15 cycles of (94°C for 20 sec, 55°C 30 sec , 72°C 2 min
72°C 7 min.

20 µl of was added to HyClone molecular biology grade water was added to each PCR product and 1 µl was used for the second round of PCR set up with 0.1 µl GoTaq Flexi DNA polymerase (Promega, M8305), 0.6 µl $MgCl_2$ (25 mM), 2 µl GoTaq Flexi colourless buffer, 0.1 µl dNTPs (20 mM; New England Biolabs, N0447), 0.2 µl of each primer (10 µM), 5.8 µl HyClone molecular biology grade water (Thermo Scientific, SH3053801).

PCR conditions were as follows:
94°C 1 min,
20 cycles of (94°C for 20 sec, 55°C 40 sec, 72°C 2 min)
72°C 7min

This was immediately followed by heteroduplex formation:
95°C for 3 min
Ramp to 55°C at 0.1°C/sec
55°C 20 min
Ramp to 25°C at 0.1°C/sec

2.3.7.2 Amplification of CODM and T6ODM for TILLING

CODM and T6ODMa were amplified from pool plates using CODM F9/R5 and T6ODM F14/R27, respectively. 10 µl PCR reactions were set up in wells of a semi skirted PCR plate (Starlabs, I1402-9800) with 0.1 µl GoTaq Flexi DNA polymerase (Promega, M8305), 0.6 µl $MgCl_2$ (25 mM), 2 µl GoTaq Flexi colourless buffer, 0.2 µl dNTPs (New England Biolabs, N0447), 0.5 µl of each primer (10 µM), 4.1 µl HyClone molecular biology grade water (Thermo Scientific, SH3053801) and 2 µl (20 ng) of template DNA.

CODM PCR conditions were as follows:

95°C 2 min,
10 cycles of (94°C for 20 sec, T_{ann} of 67-62°C (-0.5°C per cycle) 30 sec , 72°C 75sec)
37 cycles of (94°C for 20 sec, 62°C 30 sec , 72°C 75 sec
72°C 5 min.

(T6ODM F14/R27 temperature gradient was 73-68°C, with a final T_{ann} of 68°C)

This was followed immediately by heteroduplex formation:
99°C 10 min
70°C 20 sec (-0.3°C per cycle) x 70
hold at 4°C

PCR products were diluted 1_in_5 with sterile dH_2O. 2 µl of this diluted PCR product was treated with 2 µl of working concentration dsDNA Cleavage Enzyme (AATI, FS-CLKENZ) in

the heteroduplex digestion reaction. In brief, an intercalating dye in the gel binds to dsDNA and allows detection of fragments of DNA as the run through the gel. The system allows the separation of fragments between 35 bp and 5000 bp in 80 min and will accurately measure the length of (cleaved) PCR products. Mutations were detected using the Mutation Discovery Kit (AATI, DNF-920-K0500T) on the Fragment Analyser machine (AATI) according to the manufacturer's instructions for the Mutation Detection Kit. The machine was fitted with a 96-Capillary Array Cartridge (Fluorescence), 55 cm effective/80 cm total length, 50 μm ID (AATI, A2300-9650-5580). Running conditions were as follows:

Full condition
Prerun: 12 kV 30sec
Marker voltage injection (35 bp and 5000 bp): 7.5 kV 15 sec
Sample voltage injection: 7.5 kV 45 sec
Separation: 12 kV 75 min

A mutation in a pool resulted in heteroduplex formation following PCR and subsequent cleavage. Mutations were detected if two small fragments of DNA whose total sizes corresponded to the full length PCR product size were observed in a lane.

2.3.7.3 'EcoTILLING' for detection of CODM and T6ODM alleles from other sources

1μl of a 10ng/μl DNA from new varieties was combined with 1 μl of 10 ng/μl HM2 gDNA and amplified as before with the CODM F9/R5 primer combination. Heteroduplexes were allowed to form in each well and treatment with the cleavage enzyme precluded polymorphism detection on the Fragment Analyser. If there were differences in sequence unique bands would be observed on the TILLING gel.

2.3.7.4 Validation of Fragment Analyser based mutation detection

Previously the group used LICOR DNA analysers to screen for the HM2 TILLING population for mutations. Advanced Analytical's Fragment Analyser is a high-throughput capillary electrophoresis system that counts TILLING as one of its applications. Advantages over the LICOR system include automatic sample loading from a 96 well plate and the fact that expensive fluorescently labelled primers are not required. Instead a dye in the separation gel gives a signal in the presence of dsDNA. The TILLING kit provided includes an

endonuclease for the recognition and nicking of heteroduplexes. Previously the group identified mutations in the gene for the berberine bridge enzyme (BBE) which catalyses the conversion of (S)-reticuline to (S)-scoulerine. BBE mutants were found in wells B2, B4, D5, A8, C8 and F12 of pool plate 2. BBE was amplified from these six pools once more and the Fragment Analyser system was used to detect these positive controls. In addition, BBE PCR products were digested with Surveyor, the endonuclease previously used for heteroduplex cleavage, and also the dsDNA cleavage enzyme supplied with the Fragment Analyzer TILLING kit. The new enzyme proved just as efficient as Surveyor (data not shown). Furthermore, four different heteroduplex formation methods were tested in the validation (Table 3).

Both enzymes worked equally well cleaving heteroduplexes caused by the presence of BBE mutations in the pools analysed. It was decided to use the AATI enzyme supplied with the Fragment Analyser kit in future experiments rather than purchase additional Surveyor.

Method	Description	References
1	95°C 3 min Ramp to 55°C at 0.1°C per sec 55°C 20 min Ramp to 25°C at 0.1°C per sec 25°C 1 sec	Previously used in the group prior to Surveyor digestion of heteroduplexes and Licor visualisation
2	99°C 10 min 70°C 20 sec x 70 cycles (-0.3°C per cycle)	(Le Signor et al., 2009)
3	Plate placed in waterbath at 95°C for 10 min. Waterbath switched off and allowed to cool overnight.	Novel
4	96°C 10 min Ramp -0.5°C per sec to 85°C 85°C 20 sec x 150 (-0.3°C per cycle)	(Suzuki et al., 2007)

Table 3. Heteroduplex formation methods trialled in Fragment Analyser validation experiment

Likewise all four heteroduplex formation methods worked well. Mutations most easily stood out from background fluorescence when method 2 was used. In addition diluting PCR products by a factor of five prior to digestion with cleavage enzyme ensured the mutations were most easily observed (data not shown). Following this experiment cleavage of diluted PCR product subjected to hetroduplex formation method 2 with the AATI enzyme was chosen as the optimum method for screening. This was the combination in which mutations were most easily determined. Figure 10 shows an example of the peaks in flouresence generated by dsDNA running through the gel (a dye in the running gel binds to dsDNA and excites a

detector which provides a measurement in RFU). A mutation in BBE in one of the four plants in the pool caused the 977 bp PCR product to be cut into two fragments of 446 bp and 564 bp.

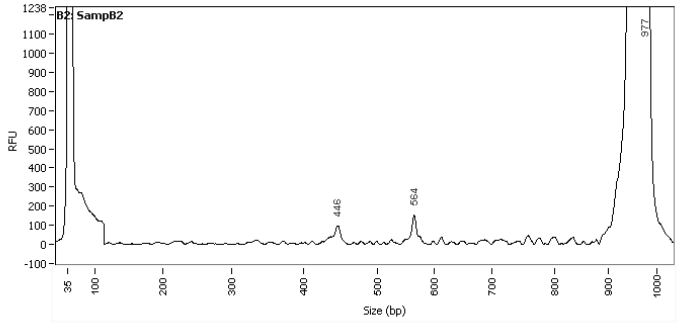


Figure 10. Example of peaks generated by AATI enzyme cleavage of 977 bp BBE PCR product. These peak sizes correspond to those observed on the Licor system where the mutation was originally discovered.

Figure 11 shows the 'gel image' from the Fragment Analyser depicting recovery of known BBE mutations with method 2 and dsDNA cleavage enzyme. In each case the ~980 bp BBE PCR product was cut into two fragments by the cleavage enzyme. The expected fragment sizes (based on Licor measurements) were from left to right- well B2 (431 bp & 551 bp), well B4 (422 bp & 556 bp), D5 (360 bp & 623 bp), A8 (233 bp & 753 bp), C8 (288 bp & 705 bp), F12 (343 bp & 644 bp).

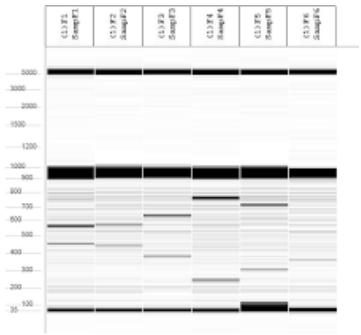


Figure 11. Recovery of six known BBE mutations on the Fragment Analyser. The brightest bands are for the 35 bp and 5000 bp markers and the full length PCR product (~980 bp). Pairs of bands with combined sizes of ~980 bp indicate the presence of a BBE mutation in the pool.

The results demonstrate the sensitivity and reproducibility of the system. It was decided to use the Fragment Analyser with these sample preparation criteria for the detection of CODM and T6ODM mutations in the HM2 TILLING population. One advantage of the Licor system is that mutation detection with labeled primers indicates where to look for the mutation in subsequent sequencing i.e. if a 300 bp fragment was observed with one

particular dye you know the mutation is approximately 300 bp from that end of the PCR product. With the Fragment Analyser the mutation could be 300 bp from either end of the PCR product. The entire PCR product was sequenced to ensure coverage of both ends. However, the superior size estimations of the Fragment Analyser meant locating the base affected by the mutation was easily accomplished i.e. it would be close to 300 bp from one of ends.

2.3.8 Rescreen & mutation confirmation

De novo DNA plates containing pooled DNA samples in which mutations were identified and the four individuals which made up the pool were created with concentrations of 10 ng/µl. These were used as template in PCR and TILLING to validate the mutation signals previously observed and identify which of the four individuals in the pool carried the mutation.

For T6ODM mutation rescreens, PCR reactions involving individuals needed to be spiked with wild type HM2 DNA to ensure heteroduplex formation and mutation detection e.g. PCR was set up with 1 µl individual DNA + 1 µl wild type DNA + 8 µl mastermix. This was not required for CODM mutation rescreens as multiple copies of CODM were amplified in PCR ensuring heteroduplex formation even in the presence of homozygous mutations in a specific copy. Once the individual plant carrying the mutation was identified the relevant gene(s) were amplified by PCR, purified and sequenced as described above.

2.4 Breeding using identified CODM and T6ODM polymorphisms

In order to track mutations in crosses, allele specific primers were designed to distinguish mutant from wild type alleles in PCR. The mutations were then used as molecular markers in a marker assisted breeding programme.

2.4.1 Set up of crosses

Sowing of forward screen lines, TILLING mutants and GSK commercial lines was staggered over a number of weeks to ensure enough parent plants flowered together to

enable crosses to be carried out. The high codeine forward screen plants in particular were known to be shorter and to flower quite early relative to controls. Each week 20 plants from each seed batch were sown.

Preferably, TILLING mutants were used as males in a cross with commercial lines. This would enable concurrent selfing of the mutant with the remaining pollen and the production of seed for the next generation if required. At the hook stage, would be female plants were emasculated and bagged. Two days after emasculation, pollen was transferred to the stigma from a suitable male plant with tweezers. The bag was replaced to ensure fertilisation by the transferred pollen only.

2.4.2 Design of primers for Allele-specific PCR (AS-PCR)

The online tools WebSNAPER (Drenkard et al., 2000) and WASP (Wangkumhang et al., 2007) were used to design primers for allele specific PCR (AS-PCR). The WASP system makes use of different destabilising effects by introducing one deliberate 'mismatch' at the penultimate base of the AS primers to improve the specificity of the resulting AS primers. The tools returned a number of possible primer pairs for each SNP. Primer pairs were assessed to determine the optimum annealing temperatures and to ensure specificity. Positive controls were M2 and M3 DNA samples with which mutations were detected by TILLING while negative controls were siblings that lacked the relevant mutation and wild type HM2 DNA. WebSNAPER and WASP outputs together with the primers ultimately used are detailed in the Appendix D.

2.4.3 AS-PCR

10 µl AS-PCR reactions were set up in wells of a 96 well PCR plate (Starlabs, I1402-9800) with 0.1 µl GoTaq Flexi DNA polymerase (Promega, M8305), 0.6 µl $MgCl_2$ (25mM), 2 µl GoTaq Flexi colourless buffer, 0.2 µl dNTPs (New England Biolabs, N0447), 0.5 µl of each primer (10 µM), 4.1 µl HyClone molecular biology grade water (Thermo Scientific, SH3053801) and 2 µl (20 ng) of template DNA.

PCR conditions were as follows:
95°C 2 min,
35 cycles of (95°C 30 sec, optimum T_{ann} 30 sec , 72°C 30 sec)
72°C 5 min.

PCR products were visualised on 1.5% w/v agarose gel. Alternatively, if large numbers of samples were to be screened for different mutations PCR products generated with different AS-PCR primers were pooled four fold before loading on the Fragment Analyser i.e. 2.5 µl of PCR product from each mutation screen plate were pooled together to create 10 µl in each well of a new PCR product pool plate. If an AS-PCR was not performed for a certain well (because it was not relevant) or there was no DNA target in a certain well, 2.5 µl of sterile dH_2O was added instead to ensure the total volume in each well was 10 µl. The PCR product pool plate was vortexed to mix and spun down briefly. 4 µl was added to 20 µl of Dilution Buffer 1X TE (AATI, DNF-495-0125).

Mutations were detected using agarose gel electrophoresis or the dsDNA Reagent Kit (AATI, DNF-900-K0500T) on the Fragment Analyser machine (AATI) according to the manufacturer's instructions. In brief, an intercalating dye in the gel binds to dsDNA and allows detection of fragments of DNA as the run through the gel. The system allows the separation of fragments between 35 bp and 500 bp and will accurately measure the length of PCR products. The machine was fitted with a 96-Capillary Array Cartridge (Fluorescence), 55cm effective/80cm total length, 50 µm ID (AATI, A2300-9650-5580).

Running conditions were as follows:
Full condition
Prerun: 8 kV 30 sec
Marker voltage injection (35 bp+500 bp): 7.5 kV 10 sec
Sample voltage injection: 7.5 kV 10 sec
Separation: 8 kV 80 min

2.5 Phenotyping

Both latex and capsules were analysed for alkaloid content using protocols established previously in the group. These procedures are detailed below. Material assessed in field trials in Tasmania in the final year were phenotyped at GSK's facilities in Port Fairy, Victoria.

2.5.1 Capsule harvest and seed cleaning

In the greenhouse, capsules were removed from mature dry plants together with the "Sd" number and placed in a medium sized negative bag. (If the plant carried more than one seed capsule and all the capsules stem from self-pollination events they were put into the same bag. If the plant was used as a mother in a cross and contained more than one capsule, the capsule used in the cross was put in a separate bag).

In the lab, the stigmatic disc was removed with a scissors and the seed was poured on to a piece of filter paper. Any large pieces of debris were removed with a forceps. The seed was then poured in to a pre labelled seed bag. The seed was assigned a seed batch id at this point by using the PDA to scan the new "S" number and the "Sd" numbers of the parent(s) used to create the seed (Figure 8). The "Sd" label of the mother plant was stored with the capsule.

2.5.2 Alkaloid extraction from latex

This method was adapted from an opium poppy mutant screening method (Fist et al., 2005). Latex was typically sampled from plants between the 10-leaf stage and the "running-up" stage (about 15 cm high). 500 µl of 10% acetic acid (Sigma, 695092) was distributed into each collection microtube (Qiagen, 19560). The Qiagen 96 well rack of collection microtubes was assigned a "PM" number and a Waters plate (Waters, 186002481) was assigned a corresponding "M" number. The Qiagen rack was taken to the glasshouse for sampling. A leaf was detached from the plant and approximately 0.5 µl of latex was collected from the wound using a 200 µl yellow pipette tip (Starlabs, S1111-0706). The yellow tip with latex was mixed in the tube until most of the latex was suspended in the acetic acid solution. The PDA was then used to link the position on the "PM" plate to the seedling ("Sd") id. The tip was left in the collection microtube while the other samples were collected. This process was repeated until all samples were collected. Tips were then removed from the tubes by tapping gently against the sides to ensure that no latex is left on the tip. The tubes were sealed with collection microtube caps (Qiagen, 19566), the rack cover was replaced and the rack was vortexed vigorously for one minute. The rack was then centrifuged at 4000 x g for 2 min. 50 µl of supernatant was transferred to the "M" plate containing 450 µl 1% acetic acid. The "M"

plate was then sealed (Waters, 186002789) with a Combi Thermo-Sealer (Waters, SP-0669/240). The extraction plate was stored for up to 2 weeks at 4°C for short term storage (-80°C for long term storage).

2.5.3 Alkaloid extraction from capsules

Capsules were thoroughly dry before processing. Safelock 2 ml microcentrifuge tubes (VWR, 211-2165) were assigned "Sd" numbers. Each capsule was placed in a stainless steel grinding jar (Qiagen, 69985) together with a steel bead and the jar was closed. A Tissue Lyser bead mill (Qiagen) was used to grind the capsules using a frequency of 25 Hz for a duration of 50 sec. The jars were allowed to stand for one minute before opening to allow the powder to settle. Powder from the grinding jar was transferred to an Eppendorf using a plastic funnel (Simport, F490-3). Each sample had its own funnel and funnels were placed in a beaker of warm water while the other samples were processed. Using the PDA the eppendorf tube ids ("Sd") were linked to the actual plant ids ("Sd").

A single microtubule was placed in a 10 ml glass beaker. Approximately 10 mg of powder (one levelled scoop of a micro-spatula scoop/round end (Corning, 3013) was weighed out in the microtube. Up to 88 samples could be weighed out and placed on each "PM" plate. Column 12 was reserved for standards-morphine, codeine, oripavine, noscapine, thebaine, blank and 10% acetic acid. Standards for morphine, codeine, oripavine and thebaine were gifts from the Bruce group (CNAP, University of York). Noscapine was obtained from Sigma (363960).

Extractions were performed on the same day as UPLC runs by specialist staff at CNAP (University of York). 500 µl of 10% acetic acid was pipetted over each powdered sample. The tubes were closed with plastic caps and the rack cover replaced. The rack was then vortexed thoroughly. It was then shaken using the bead mill at 5 Hz for 10 min before centrifugation at 4000 x g for 2 min. 450 µl of 1% acetic acid was placed in each well of a Waters plate (Waters, 186002481) and then 50 µl of the was added to each well taking care not to disturb the pellet. The extraction plate was stored for no longer than 2 weeks at 4°C (-80°C for longer storage).

2.5.4 UPLC-MS analysis of major alkaloids

This work was carried out by specialist staff at CNAP (University of York). Running plates ("R" number) were Waters 1 ml round well collection plate (Waters, 186002481). 50 µl of extract from the extraction plate ("M" number) was added to 450 µl 1% acetic acid in the running plate, respecting the plate layout. The running plate was then heat sealed. Samples were run on the UPLC with the Waters Acquity BEH C18 100x2.1 mm column maintained at 60°C. Mobile phase A was 10 mM ammonium bicarbonate, pH 10.2, with ammonium hydroxide. Mobile phase B was methanol. The weak wash solvent was 1% acetic acid (2000 µl) and the strong wash solvent was ethanol (2000 µl). The seal wash solvent was methanol. Injection loop size was 2 µl (with a 15 µl metal needle) and the injection mode was full loop with load ahead at 0.1 min.

UPLC settings: Waters Acquity, Gradient program: 0-0.2 min hold at 98% A 2% B, 0.2-0.5 min linear to 60% A 40% B, 0.5-4.0 min linear to 20%A 80%B, 4.0-4.5 min hold at to 20%A 80%B, 4.5-4.6 min linear to 98%A 2%B, 4.6-5.0 min hold at 98%A 2%B.

Mass spectrometer settings: LTQ-Orbitrap with APCI source. Source conditions: source 5 µA, vaporizer temperature 250°C, capillary temperature 250°C, capillary voltage 9 V, tube lens voltage 60 V, sheath flow (N_2) 40 units, aux flow (N_2) 20 units.

LTQ-Orbitrap conditions: positive ion mode, scan range 150-900 m/z, resolution 7500 FWHM @ 400, 1 µscan, max ion time 100 ms, injection waveforms on.

2.6 Analysis of the impact of detected mutations on alkaloid accumulation in roots and shoots

To assess the impact of CODM and T6ODM mutations on the accumulation of alkaloids in the plant an experiment was set up whereby plants from five mutant lines and wild type HM2 were sampled for alkaloids during their development.

2.6.1 Growth conditions

Seeds were sown on a 50:50 mix of silica sand (Sibelco Minerals and Chemicals, MS 01 0001) and Terragreen to allow the efficient collection of roots for sampled plants. It was anticipated that this would be easier to wash off than soil and allow the collection of minimally damaged root material for analysis. The mix was placed in root trainers (LBS Worldwide, RTTR117 & RTTR105) and was treated once with Intercept 70WG (Levington Horticulture Ltd). The plants were fed once a week with Phostrogen beginning two weeks after germination. The feeding increased to twice a week once the plants were well established (before the plants had started to show flower buds). Feeding continued through to harvesting of mature dried capsules.

For each of the lines used in the experiment (Table 18) stem and root material was collected from early seedlings (es), plants at the hook stage prior to flowering (hk), flowering plants (flw) and maturing plants 10-14 days post flowering after petal fall (mat). The dates at which each plant flowered was recorded to enable sampling of mat material.

2.6.2 Sampling of stem and root material

On the day plants were sampled, the vermiculite was washed off the roots under a running tap. The roots were then dried thoroughly with a paper towel. A small section of root was (50-100 mg) set aside for RNA extraction by placing it in a 2 ml eppendorf. Both the roots for metabolite analysis and RNA extraction were flash frozen in liquid nitrogen before storage at -80°C. When all the required roots for metabolite analysis were collected they were lyophilised for 48 h and ground up for metabolite profiling in the same manner that mature capsules are processed (see section 2.5.3).

A section of stem (approximately 1 inch in length) was taken immediately below the flower/capsule was used for RNA extraction. Exudated latex was also taken for metabolite analysis. Because sampling involved sacrificing a plant, results for different growth stages are for siblings rather than a timecourse of a single plant. Latex from each plant was stored in wells of a 96 well collection plate containing 10% acetic acid. When latex was collected from

all samples the latex was analysed for alkaloids using the method described previously (section 2.5.2).

Sampling at each stage was done in triplicate i.e. three HM2 plants each of es, hk, flw and mat stages. The dried capsule, latex and root metabolite results displayed are the average of three technical replicates.

2.6.3 RNA extraction and cDNA synthesis

Both root and stem tissue was ground in liquid nitrogen and weighed. RNA was extracted with the RNeasy Plant Mini Kit (Qiagen, 74904) according to the manufacturer's instructions. No more than 100 mg of frozen tissue was used in each extraction. Briefly, the tissue was ground thoroughly in liquid nitrogen with a mortar and pestle and the powder was decanted into a pre-cooled 2 ml eppendorf. A lysis buffer (RLT) was then added before the lysate was run through a QIAshredder spin column to remove the cell debris. After a number of cleaning steps with an RNeasy spin column a 30 μl volume of RNA from each sample was recovered by elution with RNase-free water.

Crude RNA was further cleaned up and concentrated using the RNeasy Plus Micro Kit (Qiagen, 74034). This ensured effective removal of contaminants, such as genomic DNA (gDNA) and phenol, which can interfere with transcript quantification. The procedure outlined in appendix C of the handbook was followed. Briefly, 350 μl of Buffer RLT Plus was added to the crude RNA and mixed well. The sample was run through a gDNA Eliminator spin column and transferred to an RNeasy MiniElute Spin column for a series of washes. RNase-free waster was used to elute the RNA (final volume of 12 μl).

RNA concentration was determined with the nanodrop (Thermo Scientific) and purity was assessed with the 2200 TapeStation System (Aglient) with an RNA ScreenTape (Aglient, 5067-5576).

First strand cDNA was synthesised from total RNA using Superscript II reverse transcriptase (Invitrogen, 18064-022) as described previously (Wijekoon and Facchini,

2012). Briefly, an initial 13 μl reaction was set up (to have final volume of 20 μl) with 1 μl (100-400 ng) of total RNA, 2.5 mM anchored oligo dT primer (Thermo Scientific, AB-1247), 0.5 mM dNTPs (New England Biolabs, N0447) and HyClone molecular biology grade water (Thermo Scientific, SH3053801). The mixture was heated to 65°C for 5 min followed by a quick chill on ice. Superscript II kit components, 1x RT buffer and 5 mM DTT were added before incubation at 42°C for 2 min. Finally 1 μl (200 units) of the reverse transcriptase was added and the reaction was incubated at 42°C for 50 min followed by an inactivation step of 70°C for 15 min.

2.6.4 QPCR

Real time quantitative PCR (RT-qPCR) was carried out with SYBR Green detection on triplicate technical and triplicate biological replicates for each tissue sampled representing different development points in the 6 lines sown for the experiment. 0.5 μl of cDNA was added to 8.3 μl HyClone molecular biology grade water (Thermo Scientific, SH3053801), 1.2 μl of 300nM primer mix and 10 μl of 2x Fast SYBR Green Master Mix (Applied Biosystems, 4385612). The StepOne™ Plus Real Time PCR system (Applied Biosystems) was used for real-time quantitative detection of target nucleic acid sequences. Primers for the relative quantification of target genes were used previously (Hagel and Facchini, 2010a; Wijekoon and Facchini, 2012). PCR conditions were 95°C 20 sec, followed by 40 cycles of 95°C 3 sec, 60°C 30 sec. Fluorescent signal intensities were recorded and analysed using the system software. The threshold (Ct) value for each gene was normalised against the Ct for β-actin, which was used as the constitutive reference transcript. Relative levels of morphinan gene transcripts were compared using the $2^{-\Delta\Delta_T}$ method (Livak and Schmittgen, 2001).

Chapter 3- Reverse genetic screens

3.1 HM2 mutant population

A mutagenised population of 10000 lines has been established from the GSK morphine variety HM2. The mutagenised population has been used both for the isolation of mutants in specific target genes (reverse genetics) in York and for field based screening for altered alkaloid content in Tasmania (forward genetics). The B2, C1 and B3 subpopulations make up the TILLING population (Table 1) with 4146 members.

3.2 Multiple copies of CODM in the opium poppy genome

3.2.1 Sequencing of CODM from genomic DNA template

The cDNA sequence for CODM (GenBank accession: GQ500141) was used to probe our in house EST database for contigs with similar sequence. ps1bContig1267, an EST built from HM2 ESTs, was used to design primer CODM_F1f. Situated before the transcription initiation site, this primer in conjunction with primer CODM_R1, situated downstream of the termination site, would allow amplification of a large portion of the CODM gene from a HM2 gDNA pool. The resulting ~2kb PCR products were cloned with the Strataclone Blunt PCR cloning kit. Colony PCR confirmed the presence of the ~2 kb inserts (Figure 12).

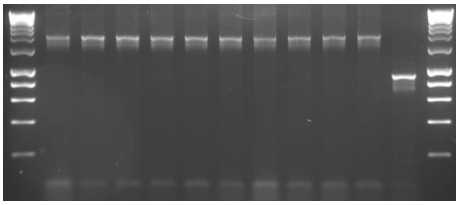


Figure 12. M13 colony PCR with ten positive CODM transformations. Lane 11 contained the control insert supplied with the cloning kit.

Plasmids from 6 of these colonies were purified and sequenced. This led to the identification of three introns in CODM when the cDNA and gDNA sequences were compared in a sequence alignment (Figure 16).

3.2.2 Polymorphic sites identified in CODM sequence

Closer examination of the alignment revealed the presence of two sites of conflict. The first is a silent single nucleotide polymorphism in exon 4. TCG and TCT both encode serine. A SNP was also noted 16bp downstream of the stop site in the 3'UTR. A 'C' or a 'T' is possible at this site. One clone was an exact match to the published cDNA sequence (GQ500141) with a 'G' at the exon 4 site and a 'T' at the 3'UTR site (Table 6). PCR clones containing the containing the alternative combinations of bases at the two polymorphic sites were also sequenced.

Matches for CODM in the EST database were probed for the occurrence of these polymorphic sites to determine if this observation was merely the consequence of different alleles of CODM in HN1 or if it may point to the expression of multiple copies of CODM in GSK commercial cultivars. CODM contigs from the different cultivars (HM2, HM1, HN1 and HT1) often were built up using sequence reads with different bases at the exon 4 and 3'UTR sites (Table 4). For example, ps1bContig1872 is a CODM contig made from HM2 ESTs and at the exon 4 site 6 of the 9 EST reads have a 'G', while the remainder have a 'T' (Figure 13). The ESTs in this contig support the role of multiple copies of CODM in HM2. This is likely replicated across the four cultivars (Table 4).

Exon 4 SNP

'G'	'T'
GQ500141	
ps1bContig1872* (6of9 EST reads)	ps1bContig1872* (3of9 EST reads)
ps1cContig5406* (1of2 EST reads)	ps1cContig5406* (1of2 EST reads)
ps2aContig1105 (3of5 EST reads)	ps2aContig1105 (2of5 EST reads)
ps3cContig12033 (all)	ps3bContig7406 (3of5 EST reads)
ps3dContig9594 (1of2 EST reads)	ps3dContig9594 (1of2 EST reads)
ps4aContig4210 (all)	ps4cFQURZIE02KUPOU

HM2 HM1 HN1 HT1

3'UTR SNP

'T'	'C'
GQ500141	
ps1bContig1872* (4of6 ESTs)	ps1bContig1872* (2of6 ESTs)
ps1cContig5406* (4of7 ESTs)	ps1cContig5406* (3of7 ESTs)
ps2aContig1105 (7of8 ESTs)	ps2aContig1105 (1of8 ESTs)
ps2bContig6192, ps2cContig2179, (all)	
ps3aContig10572 (2of4 ESTs)	ps3aContig10572 (2of4 ESTs)
ps3bContig7406 (4of8 ESTs)	ps3bContig7406 (4of8 ESTs)
ps3cContig12033 (1of4 ESTs)	ps3cContig12033 (3of4 ESTs)
ps3dContig9594 (2of 3 ESTs)	ps3dContig9594 (1of3 ESTs)
ps4aContig4210, ps4bContig6687, ps4cContig1507	

HM2 HM1 HN1 HT1

Table 4. Recording the presence of different bases at two polymorphic sites in CODM ESTs. ps1 to ps4 contigs were built from morphine cultivars HM2 & HM1, noscapine cultivar HN1 and thebaine cultivar HT1, respectively. GQ500141 is the published CODM cDNA sequence (Hagel and Facchini, 2010a). *The EST assemblies for contigs ps1bContig1872 and ps1cContig5406 are displayed in Figure 13 showing the number of times each bases was observed at the two polymorphic sites e.g. at site 1 in ps1bContig1872 6 of the 9 ESTs in this contig contain a G.

ps1bContig1872

Exon 4 variation

```
ps1bFPZG94Y02LD5LU+   TAGGCCCAA--TTTCG-AGCTTGGTCAC
ps1bFPZG94Y01ASC5R+   TAGGCCCAA--TTTCG-AGCTTGGTCAC
ps1bFPZG94Y02LBBVH+   TAGGACCAA--TAGCG-AGCTTGATCACACC-AGAGACACCTG---CTTTGTTCAAAAG
ps1bFPZG94Y02NWZ57+   TAGGCCCAA--TTTCT-AGCTTGGTCACACC-AGAGACACCTG---CTTTGTTCAAAAGA
ps1bFPZG94Y01EFH8P+   TAGGCCCAA--TTTCG-AGCTTGGTCACACC-AGAGACACCTG---CTTTGTTCAGAAGA
ps1bFPZG94Y02KRRRF+   TAGGCCCAA--TTTCT-AGCTTGGTCACACC-AGAGACACCTG---CTTTGTTCAAAAGA
ps1bFPZG94Y01AE996+   TAGGCCCCAATTTTCTTAGCTTGGTCACACCCAGAGACACCTG---CTTTGTTC
ps1bFPZG94Y02O21TK+   TAGGCCCAA--TTTCT-AGCTTGGTCACACC-AGAGACACCTG---CTTTGTTCAAAAGA
ps1bFPZG94Y02LVTVN+   TAGGCCCAA--TTTCG-AGCTTGGTCACACC-AGAGACACCTG---CTTTGTTCAAAAGA
ps1bFPZG94Y01DR7IC+      GCCCAA--TTTCG-AGCTTGGTCACACC-AGAGACACCTG---CTTTGTTCAAAAGA
ps1bFPZG94Y02M2MJD+                            CACACC-AGAGACACCTGTTTCTTTGTTCAAAAGA
                      ____________________________________________________________
consensus             TAGGCCCAA--TTTCG-AGCTTGGTCACACC-AGAGACACCTG---CTTTGTTCAAAAGA
```

3'UTR variation

```
ps1bFPZG94Y01EFH8P+   TTTCTCGACTAC
ps1bFPZG94Y02KRRRF+   TTTCTCGACTACATGAGGATGTGAGAAAGTGTG-AACA
ps1bFPZG94Y02O21TK+   TTTCTCGACTACATGAGGATGTGAGAAAGTGTGTAACATATTATTATACTCCACATTGTGT
ps1bFPZG94Y02LVTVN+   TTTCTCGACTACATGAGGATGTGAGAAAGTGTG-AACATATTATTATACTCCACATTGTGT
ps1bFPZG94Y01DR7IC+   TTTCTCGACTACATGAGGATGTGAGAAAGTGTG-AACATACATTATACTCCACATTGTGT
ps1bFPZG94Y02M2MJD+   TTTCTCGACTACATGAGGATGTGAGAAAGTGTGAACCATATATTATACTCCACATTGTGT
ps1bFPZG94Y02NEY0J+   TTTCTCGACTACATGAGGATGTGAGAAAGTGTG-AACATACATTATACTCCACATTGTGT
ps1bFPZG94Y01CKEMT+   TTTCTCGACTACATGAGGATGTGAGAAAGTGTG-AACATATTATTATACTCCACATTGTGT
ps1bFPZG94Y01DDA1Y-                                                  CCACATTGTGT
ps1bFPZG94Y02NOC71+                                                            T
                      ____________________________________________________________
consensus             TTTCTCGACTACATGAGGATGTGAGAAAGTGTG-AACATATATTATACTCCACATTGTGT
```

ps1cContig5406

Exon 4 variation

```
ps1cFPZG94Y02K61LC+   TCAATCGCGACATTTCATGACTCTAAACTAGAGTCAGAAATAGGCCCAATTTCGAGCTTG
ps1cFPZG94Y01EU29I-                                  CAGAAATAGGCCCAATTTCTAGCTTG
                      ____________________________________________________________
consensus             TCAATCGCGACATTTCATGACTCTAAACTAGAGTCAGAAATAGGCCCAATTTCGAGCTTG
```

3'UTR variation

```
ps1cFPZG94Y01EU29I-   TGTGAACATATTATTATACTCCACATTGTGTTTAATATATGATGAAATAAGTTGCTTTTGA
ps1cFPZG94Y01C510H+   TGTGAACATACATTATACTCCACATTGTGTTTAATATATGATGAATTAAGTTGCTTTTGA
ps1cFPZG94Y01D1274+   TGTGAACATACATTATACCCCACATTGTGTTTAATATATGATGAAATAAGTTGCTTTTGA
ps1cFPZG94Y01C0QQS+   TGTGAACATACATTATACTCCACATTGTGTTTAATATATGATGAAATAAGTTGCTTTTGA
ps1cFPZG94Y02LKF2D+   TGTGAACATATATTATACTCCACATTGTGTTTAATATATGATGAAATAAGTTGCTTTTGA
ps1cFPZG94Y02OJVTU+    GTGAACATATATTATACTCCACATTGTGTTTAATATATGATGAAATAAGTTGCTTTTGA
ps1cFPZG94Y02LTH3C+    GTGAACATATATTATACTCCACATTGTGTTTAATATATGATGAAATAAGTTGCTTTTGA
                      ____________________________________________________________
consensus             TGTGAACATATATTATACTCCACATTGTGTTTAATATATGATGAAATAAGTTGCTTTTGA
```

Figure 13. Alignment around the sites of CODM sequence variation in two HM2 EST contigs. The polymorphic sites are highlighted.

In order to determine whether the different copies of CODM are present in all plants or represent alleles of CODM in the lines analysed, DNA from an M2 plant (derived from EMS mutagenesis of HM2) was taken together with the DNA of numerous M3 plants derived from the seed collected from the self-pollinated M2 plant (Sd-110071). CODM was amplified from

each DNA sample. PCR products were purified and sequenced. Double peaks were observed in the chromatograms for the M2 and all M3 plants indicating at least two, if not all three, versions of CODM are in each plant (Table 5).

Description	Seedling_id	Exon 4	3'UTR
M2	Sd-110071	G/T	C/T
M3	Sd-110994	G/T	C/T
M3	Sd-110996	G/T	C/T
M3	Sd-110997	G/T	C/T
M3	Sd-110998	G/T	C/T
M3	Sd-110999	G/T	C/T
M3	Sd-111000	G/T	C/T
M3	Sd-111001	G/T	C/T
M3	Sd-111003	G/T	C/T
M3	Sd-111004	G/T	C/T
M3	Sd-111006	G/T	C/T
M3	Sd-111008	G/T	C/T
M3	Sd-111012	G/T	C/T
M3	Sd-111013	G/T	C/T

Table 5. Signals observed in the chromatograms at the two CODM polymorphic sites in an M2 individual and M3 plants derived from selfed seed of the M2 plant.

If the M2 plant, Sd-110071, was heterozygous for CODM, and the versions observed were alleles of CODM, a quarter of M3s would have a 'C' at the 3'UTR polymorphic site, half would have 'C/T' and the remaining quarter would have 'T'. However, all M3 plants were genotyped 'C/T' (Table 5). Likewise all M3 plants were genotyped 'G/C' at the exon 4 polymorphic site. Therefore, these do not represent alleles of CODM. Rather, multiple copies of CODM are present in each plant. These contain nucleic acid differences at two polymorphic sites but the amino acid sequence of the encoded CODM would remain unchanged.

3.2.3 Three copies of CODM in the genome of opium poppy

Further cloning and sequencing of CODM from HN1 gDNA and cDNA templates revealed the presence of three copies of CODM which we designate CODMa, CODMb and CODMc (Table 6). CODMa is identical to the characterised gene (GenBank accession: GQ500141; Hagel and Facchini, 2010a).

	Exon 4 site	3'UTR site
CODMa	G	T
CODMb	T	T
CODMc	G	C

Table 6. Nucleotides at the exon 4 and 3'UTR polymorphic sites in copies of CODM

Large cultures of two of the BACs which the CODM probe pulled out of the library, PS_33D07 and PS_158A11, were prepared for BAC purification. CODM was amplified using each of these purified BACs as template in PCR. Products were again cloned and sequenced as before with gDNA and cDNA template.

Template_Primers used (Colony number)	Exon 4 site	3' UTR site	Copy of CODM
33D07_CODM F9/R5 (3)	T	T	CODMb
33D07_CODM F9/R5 (6)	T	T	CODMb
33D07_CODM F9/R5 (7)	T	T	CODMb
33D07_CODM F1f/R1 (2)	T	T	CODMb
33D07_CODM F1f/R1 (5)	T	T	CODMb
33D07_CODM F1f/R1 (6)	T	T	CODMb
33D07_CODM F1f/R1 (7)	T	T	CODMb
158A11_CODM F9/R5 (6)	G	C	CODMc
158A11_CODM F9/R5 (8)	G	T	CODMa
158A11_CODM F9/R5 (9)	G	T	CODMa
158A11_CODM F1f/R1 (2)	G	C	CODMc
158A11_CODM F1f/R1 (3)	G	T	CODMa
158A11_CODM F1f/R1 (4)	G	C	CODMc
158A11_CODM F1f/R1 (5)	G	T	CODMa
158A11_CODM F1f/R1 (6)	G	C	CODMc
158A11_CODM F1f/R1 (7)	G	C	CODMc
158A11_CODM F1f/R1 (10)	G	T	CODMa

Table 7. Bases occurring at the CODM polymorphic sites in CODM clones obtained when using BAC DNA as a template in PCR.

Table 7 shows the nucleic acids at the exon 4 and 3' UTR polymorphic sites detected in each CODM clone sequenced. CODMb was cloned and sequenced seven times from the PS_33D07 template, suggesting that only CODMb resides on this BAC. By contrast CODMa and CODMc were cloned and sequenced using PS_158A11 as a template. The results suggest close linkage of copies of CODM in the genome. BACs typically contain 150-350 kb of sequence and BACs containing CODMa and CODMc were shown to be within this range. PCRs for T6ODM and other genes involved in morphine biosynthesis such as BBE, 7-OMT, COR, SalAT and SalR failed so CODM is unlikely to form part of a morphinan gene cluster if such a cluster does exist. If all three copies of CODM detected are functional then mutations in all would have to be found to completely block metabolism at codeine.

3.2.4 Two copies of CODMc are within ~150 kb of each other

Amplicon Express sequenced a selection of the CODM-containing BACs with the HiSeq (sequencing-by synthesis; Illumina) and PacBio (single molecule real time sequencing; Pacific Biosciences) sequencing platforms. Assembly of the sequence reads into contigs was carried out by Yi Li. Assembly was a challenge due to repetitive sequence. Scaffolds were assembled from HiSeq data of 48 x 2 paired end (PE) libraries from each BAC and four size (500, 800, 2000 and 5000 bp) mate pair (MP) libraries from all BACs.

PacBio contigs for 8 BAC clones were then compared with the scaffolds. The most reliable contigs/scaffolds built were generated from PacBio sequencing of two BACs-89C05 and 109:H06. CODM is highlighted in the sequences of the three largest contigs in Appendix J.

BAC:Contig	Length (kb)	Gene copy present
PS_109H06:000063	54	CODMb
PS_89C05:000086	41	CODMc
PS_89C05:000087	11.5	CODMc

Table 8. Copies of CODM detected on the three largest CODM containing contigs generated using the sequencing results obtained from Amplicon Express.

89C05 contains two copies of CODMc and 109H06 contains CODMb (Table 8, Figure 14). BACs typically have a capacity of 150 kb. Even though the two copies of CODMc were not on the same contig, the contigs were generated with the same template DNA and must be within 150 kb of one another.

DNA from seven of the BACs in which the CODM probe hybridised was used in PCR in an attempt to amplify CODMa/c and CODMb. Reverse primers exploiting the variation at the exon 4 polymorphic site (Figure 13) would enable specific amplification of CODMa/c or CODMb. If independent PCR products were obtained using both reverse primers, it would indicate that at least two copies of CODM reside on the given BAC.

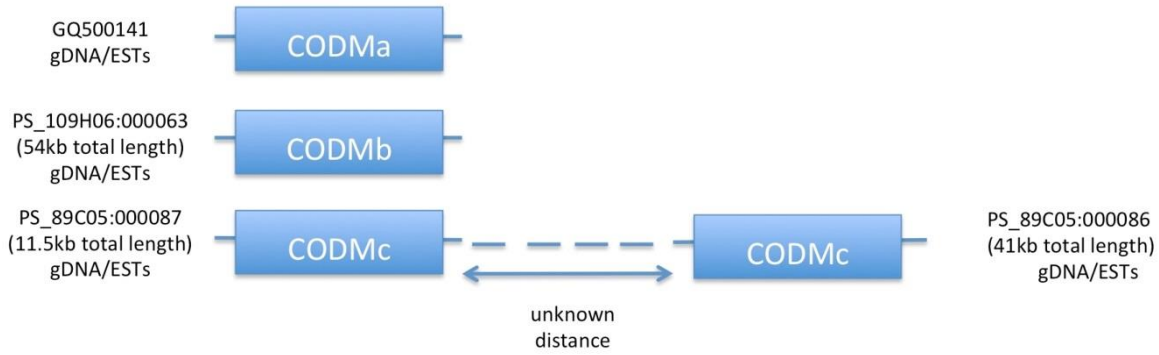


Figure 14. Schematic showing copies of CODM sequenced and their source. CODMa/b/c were originally detected in ESTs (Figure 13) and from gDNA sources (Table 5 and data not shown). Sequencing of two BACs detected with a CODM probe confirmed the presence of copies of CODM. Two copies of CODMc were observed on BAC 89C05. Although they were observed in one continuous contig the fact that they reside on the same insert in this BAC means the two copies of CODM must reside within a range in line with the capacity of a BAC. Overall the results point to the presence of a multigene family possibly originating from duplications.

PCR products were visualised on an agarose gel (Appendix J). The results indicate that 89C05 may contain a copy of CODMb as well as two copies of CODMc described above. CODMb was found to be present on all the BACs tested. CODMb is the only version present on 109H06 (Appendix J).

3.2.5 Sequence analysis of contigs from 89C05 and 109H06

The large 54 kb contig obtained using sequence data from BAC 109H06 contains CODMb. The promoter region agrees almost exactly with a CODM promoter sequenced and characterised by US researchers over a decade ago (Raymond, 2004). There are only 3 conflicts in a 2.5 kb overlap. This validates the PacBio sequencing and contig assembly and means a full length promoter, coding regions, introns and downstream region is available for CODMb on BAC 109:H06.

The 2.5 kb promoter sequence also assembles with the two contigs generated from 89C05 sequencing- PS_89C05:000086 (41 kb) and PS_89C05:000087 (11.5 kb). However, this can only happen gaps in the alignment are allowed. The results suggest that a series of duplications of both coding and promoter regions has taken place.

Primers F14/R10 span a potential duplicated of the CODM promoter in PS_89C05:000086. A tandem duplication of 348 bp was spotted in the assembly and primers spanning this site were designed. Amplification with F14/R10 from the region represented

by PS_89C05:000086 would create a 1099 bp product. In contrast, amplification from the region represented by PS_89C05:00087 would create a 751 bp product. PCR using BAC 89C05 as a template yielded both products (Figure 14). Although the larger product was likely outcompeted in the early rounds of PCR there is a faint band for a 1099 bp PCR product as well as the 751 bp product. This verifies the presence of two copies of CODMc on the region of the poppy genome carried by BAC 89C05. Furthermore, although the coding sequence of CODMc is unchanged, there are a number of duplicated regions both upstream and downstream.

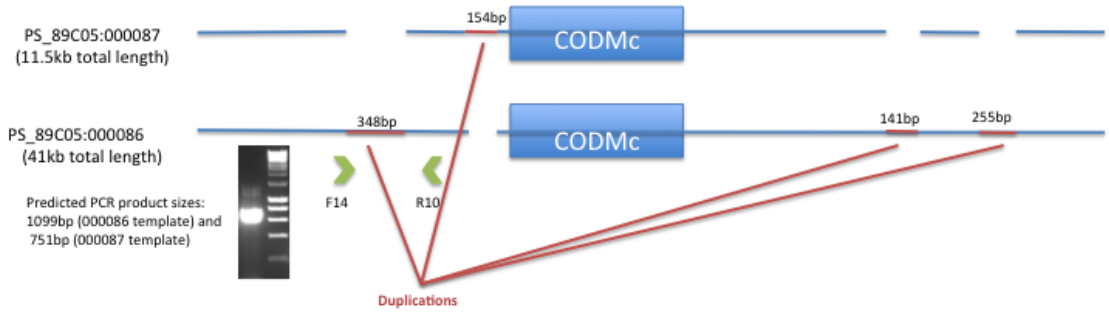


Figure 15. Overview of the duplications in the alignment of contigs generated from PacBio sequencing of BAC 89C05. The blue lines represent sequence upstream and downstream of CODMc. Red lines indicate areas of sequence that have been duplicated e.g. a section of 348 bp upstream of CODMc has been tandemly duplicated in contig PS_89C05:000086. A gap in the other contig is thus present in the alignment. Primers (F14/R10) spanning this site were designed to verify the existence of the duplication in 89C05:000086 by PCR. Different sized PCR products were generated using template represented by contigs 000086 and 000087.

In addition, PS_89C05:000087 contains an 154 bp duplication immediately prior to the CODM coding region out of which 110 bp is an exact match to the beginning of the CODM coding sequence It is not clear yet whether the duplications alter the ability of the genes to be transcribed or if they contain *cis* elements that might alter the circumstances under which different copies of CODM are expressed. There are also further duplications (3 kb and 3.6 kb downstream of CODM) in PS_89C05:000087. Apart from the duplications the sequences of PS_89C05:000086 and PS_89C05:000087 are remarkably similar suggesting a duplication even of the genomic regions containing CODMc.

3.2.6 ORF Prediction

FGENESH (Solovyev et al., 2006) and NCBI's ORF Finder (available at http://www.ncbi.nlm.nih.gov/projects/gorf/; Marchler-Bauer et al., 2011) were used to

identify potential open reading frames in the two largest contigs PS_89C05:000086 and PS_109H06:000063. Apart from CODM, most of the ORFs identified contained conserved domains for retrotransposons related genes such as retropepsin, gag protein, reverse transcriptase, RNase H1 and RVE integrase. No other genes involved in morphine biosynthesis were found. However, a DIOX_N conserved domain was identified on PS_89C05000086 within 7 kb of CODMc. Other ORFs identified on this contig were for a protein with 60-70% identities to predicted nucleobase-ascorbate transporter from *Vitis vinifera* and *Phoenix canariensis* and a protein with 60% identity to a predicted transcription factor, DPB-like from *Fragaria resca*. Predicted ORFs on PS_109H06:000063 had identities to peptidase, M14-carboxypeptidases and chromosome partition protein, Smc.

3.2.7 CENSOR analysis for repeats

A common feature of plant gene clusters is the presence of neighbouring transposable elements and retrotransposons such as those described in *Arabidopsis thaliana*, *Lotus japonica* and *Papaver somniferum* (Field et al., 2011; Krokida et al., 2013; Winzer et al., 2012). The CENSOR programme was used to scan the two largest contigs for repeat elements (Kohany et al., 2006).

Most of the repetitive DNA on both contigs (Table 9) is due to Long Terminal Repeat (LTR) Retrotransposons, specifically of the Ty3/Gypsy family. These elements are similar to retroviruses but lack an ORF for an envelope (ENV) proteins required for intercellular transmission of retroviruses.

LTR and non-LTR retrotransposons (Class I transposable elements) operate via a "copy-and-paste" mechanism whereby mRNA transcribed from the element by RNA polymerase II undergoes reverse transcription and then integrates into a new position in the genome by an integrase (Lisch, 2013). hAT and CACTA are class II elements which operate via a "cut-and-paste" mechanism in which the elements are physically excised from the chromosome and reintegrated at a new location by a transposase enzyme. Helitrons are the third class of transposable elements found in plant genomes and these are thought to transpose via a

"rolling circle" mechanism. Nicking at the Helitron terminus is followed by DNA synthesis, strand displacement and resolution of a heteroduplex by DNA replication. Many Helitrons have short sequences at their termini and "filler" sequences composed of fragments of captured genes.

PS_89C05:000086

Repeat Class	Fragments	Length
Integrated Virus	1	91
Caulimoviridae	1	91
Transposable Element	34	5161
DNA transposon	13	1565
Helitron	2	144
MuDR	4	774
hAT	3	173
LTR Retrotransposon	21	3596
Copia	5	611
Gypsy	15	2927
Total	**35**	**5252**

PS_109H06:000063

Repeat Class	Fragments	Length
Transposable Element	44	9332
DNA transposon	15	1688
EnSpm/CACTA	4	437
Helitron	4	336
Mariner/Tc1	1	181
MuDR	3	355
hAT	2	239
LTR Retrotransposon	28	7587
Copia	4	317
Gypsy	24	7270
Non-LTR Retrotransposon	1	57
Total	**44**	**9332**

Table 9. Summary of repeats identified on the two largest CODM containing contigs

3.2.8 Section summary

Cloning and sequencing of CODM from genomic DNA of GSK commercial lines of opium poppy has revealed that there are multiple copies. Initial efforts determined differences between copies at just two polymorphic sites. Three different versions of CODM were detected, differing only with the bases at these two sites. The differences do not lead to amino acid changes. Subsequent sequencing of BAC 89C05 showed that two copies of CODMc are physically linked to each other. Therefore, there are at least four copies of CODM in the genome. BAC derived sequence information will enable further study of these copies. It is likely that they arose due to gene duplications. No other genes involved in the morphinan pathway were observed on the CODM containing BACs but the genes were in a context of much LTR Retrotransposon sequence.

3.3 TILLING for induced mutations in CODM

3.3.1 Design of appropriate primers to maximise mutation recovery

Using the sequence information obtained from sequencing clones of CODM the Primer 3 Plus webtool (Untergasser et al., 2007) was used to design primers for TILLING. At this stage it was known that there were multiple copies of CODM in the genome and although CODM-containing BACs were not yet sequenced the decision was made to screen all copies of CODM simultaneously using primer pair F9/R5 (Figure 15). PCR with this primer pair would generate a PCR fragment of 1335 bp allowing mutation detection in exons 2, 3 and 4. The 'window' to be screened also spanned numerous splice sites which could be interrupted by mutations introduced by EMS. Mutations at these sites could lead to aberrant RNA splicing and the subsequent translation of altered or truncated proteins. If deemed necessary in future, mutations in exon 1 could be detected with primer pair F10/R6.

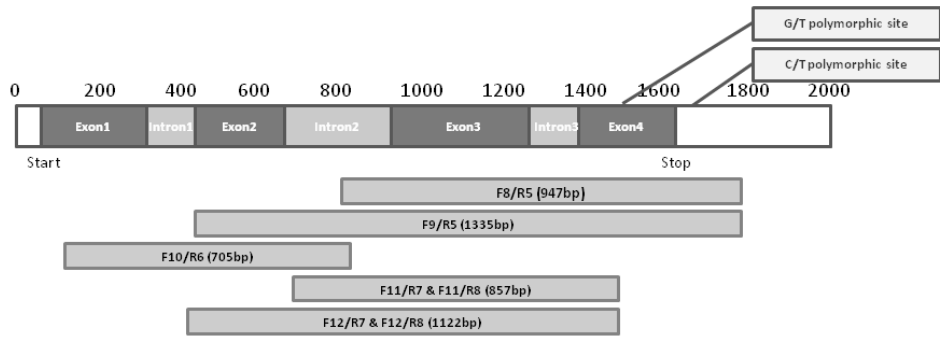


Figure 16. Regions of CODM amplified using different CODM primer combinations. Primer pair F9/R5 was used for TILLING. Primer pairs F12/R7 and F12/R8 were used to amplify CODMa/c and CODM, respectively.

Differences between CODM copies towards the 3' end of the genes were anticipated to result in common banding patterns on TILLING gels across all samples i.e. heteroduplex bubbles would form at these points in all samples and result in cleavage (Figure 17). Cleavage at these sites would generate four fragments with sizes of 1227 bp, 1091 bp, 244 bp and 108 bp.

Primers R7 and R8 make use of the polymorphism in exon 4 in that utilisation of primer R7 results in amplification of CODMa and CODMc. Primer R8, in contrast, was designed to selectively amplify CODMb. These primers would be useful for amplifying CODM for

sequencing reactions in order to determine which copy of CODM any detected mutation resides on.

3.3.2 Mutation screen

Figure 17a shows the Fragment Analyser gel when digested pool plate 2 CODM PCR products were loaded. Due to differences among copies of CODM in the genome there are common bands in the TILLING gel as expected. The task is to identify any unique pairs of fragments. An example of a CODM mutation is present in Lane 34 (well C10). Figure 17b displays the fluorescent signals observed in Lane 34. The software estimated the sizes of the cut fragments unique to this lane to be 456 bp and 841 bp. The sum of the sizes of these fragments is the roughly the size of the full length PCR product. Figure 17c shows the rescreen of well C10 of pool plate 2 plus the four individual DNA samples that made up the pool. These DNA samples were located using the information in Appendix A. The same fragments were seen in the pooled DNA sample and the third individual (Sd-100571). This individual is the carrier of the CODM mutation.

This process was repeated for all twelve pool plates (Appendix B). Typically when doing TILLING with individuals the DNA would be spiked with wild type DNA. This would allow for heteroduplexes to form after PCR between wild type and mutant alleles. For CODM, multiple copies were amplified in the PCR and so heteroduplexes formed between mutant alleles and wild type alleles from an additional copy. In many cases only one unique band was observed in a lane in a TILLING gel. Nevertheless this was brought forward for rescreen to see if it was reproducible and if a clearer signal was observed in an individual.

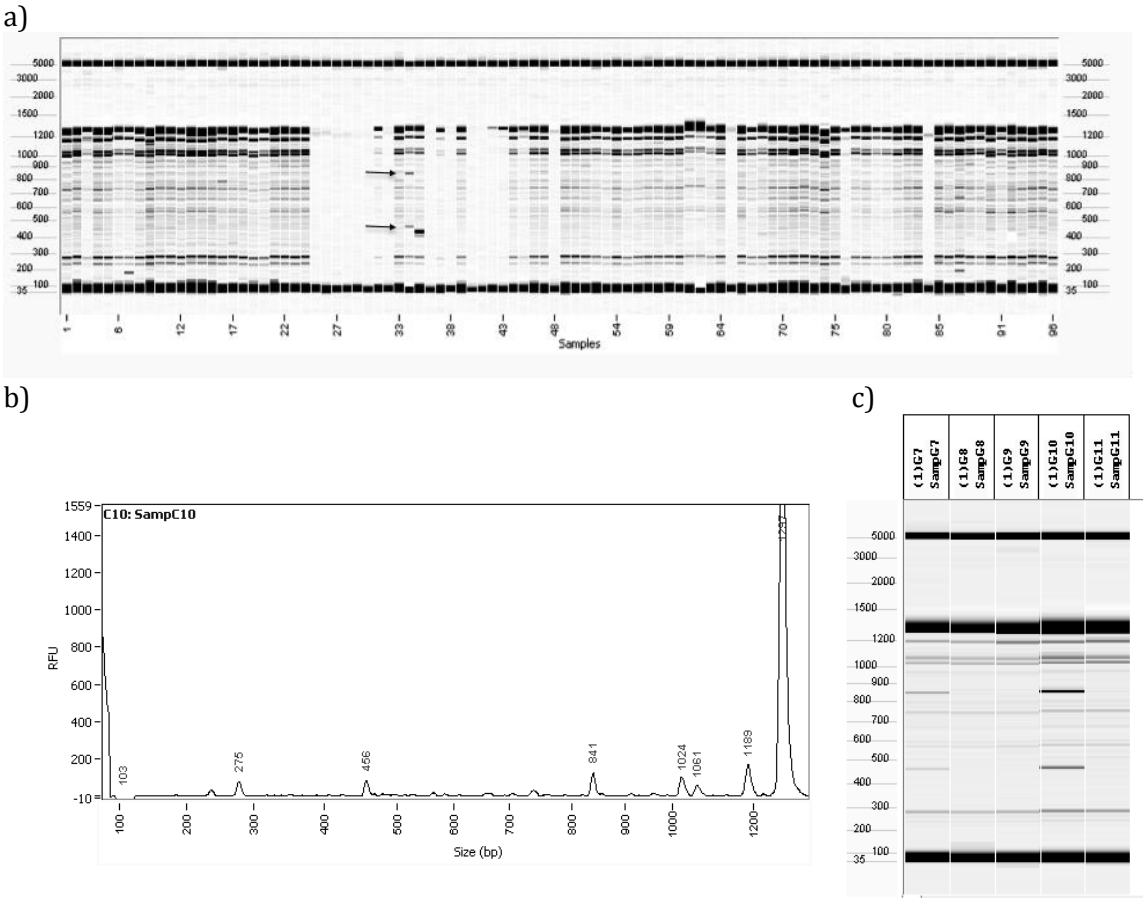


Figure 17. Identification of a CODM mutation (using primer pair F9/R5 for PCR). a) Fragment Analyser gel image for pool plate 2 CODM PCR products after digestion with the AATI enzyme. Lanes 1-12 = Wells A1-A12, Lanes 13-24=Wells B1-B12 etc. A CODM mutation is present in well C10 (Lane 34)-unique bands are indicated with arrows. b) Fluorescent signals detected in well C10. Unique 456 bp and 841 bp fragments indicate a CODM mutation. c) Rerun of (from left to right) well C10 and the four components of the pool. The mutation is in the third individual.

3.3.2.1 Determining which copy of CODM affected by EMS induced mutation

When all individuals carrying CODM were identified, DNA from these individuals was once again used in PCR to amplify CODMa/c and CODMb. It was not possible using the sequence information available at the time to design primers to distinguish between CODMa and CODMc. Both CODMa/c and CODMb PCR products were purified and sequenced and the fragment sizes observed on the TILLING gels were used to inform where the mutations could potentially be.

3.3.2.2 Summary of CODM mutations detected in reverse genetic screen

Table 10 provides a summary of the CODM mutations found in the reverse genetic screen of 4146 M2 plants. Appendix B has the full record of the fragments detected in the initial screen and the follow up rescreen as well as the actual TILLING gel images. In total 44 mutations in CODM were confirmed in 42 plants. Five were silent i.e. the nucleotide change

will not cause an amino acid substitution. Many of the mutations identified were in introns 2 (16) and 3 (12). This left 9 missense mutations in the coding region of CODM (Figure 18). Most importantly, 2 nonsense mutations were identified-W261* in CODMa/c and Q254* in CODMb.

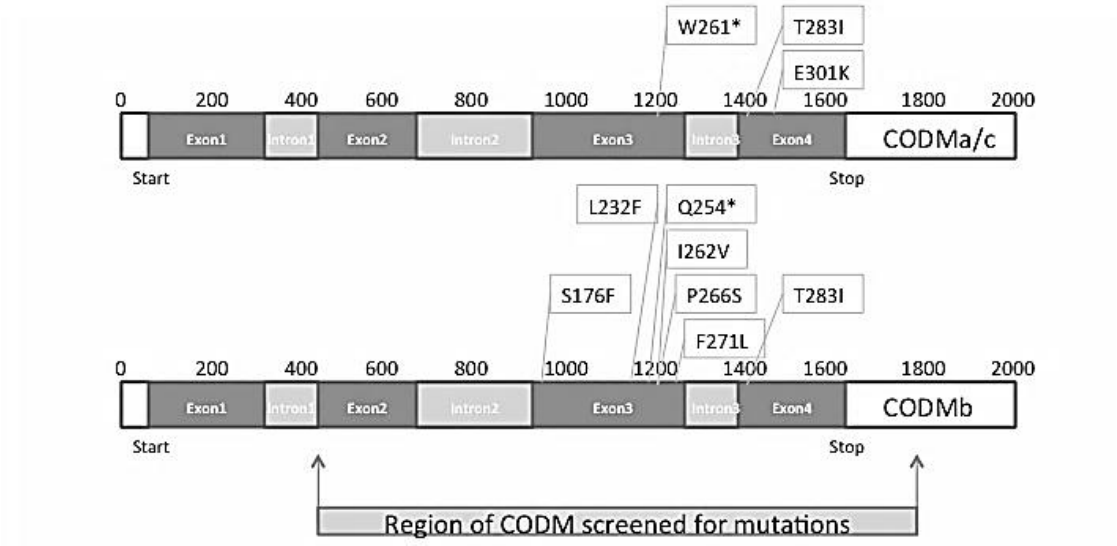


Figure 18. Locations of mutations identified in CODMa/c and CODMb. The region of the genes amplified using primer pair F9/R5 are indicated.

Pool plate	Well	Screening		Rescreening		Plant Sd numbers in pool (Mutation carrier highlighted)				DNA Sources	Sequencing Result	CODM copy affected	Seed batch number of mutant	Population
		Fragment 1 (bp)	Fragment 2 (bp)	Fragment 1 (bp)	Fragment 2 (bp)									
One	A8	315	-	881	-	100057	100058	100059	100060	D-10001 A8-D8	Intron 2	-	-	-
	C4	390	914	386	911	100105	100106	100107	100108	D-10002 A4-D4	Intron 3	-	-	-
Two	B4	778	-	795	-	100429	100430	100431	101234	D-10005 E4-H4	Y177Y silent	-	-	-
	C10	456	841	462	849	100569	100570	100571	100572	D-10006 A10-D10	Intron 2	-	-	-
Three	D8	383	941	373	919	101195	100942	100943	100944	D-10010 E8-H8	Intron 3	-	-	-
Four	E10	912	-	406	896	101712	101714	101720	101721	D-10015 A10-D10	Intron 2	-	-	-
	F3	364	966	349	940	101493	101494	101495	101497	D-10015 E3-H3	T283I	CODMb	S-101497	B2
	H7	446	849	442	844	101626	101627	101628	101629	D-10016 E7-H7	Intron 2	-	-	-
Five	F2	802	-	528	786	101981	101982	101983	101984	D-10019 E2-H2	S176F	CODMb	S-101983	B2
	H1	493	826	483a 370b	813a 831b	101946	101947	101948	101949	D-10020 E1-H1	Intron 2, a Intron 3,b	-	-	-
Six	G7	530	788	536	784	102571	102572	102573	102574	D-10024 A7-D7	W261*	CODMa/c	S-102574	B2
	H3	785	-	535	783	102503	102504	102505	102506	D-10024 E3-H3	Y177Y silent	-	-	-
Seven	G11	495	813	500	807	103225	103226	103227	103228	D-10028 A11-D11	F271L	CODMb	S-103228	C1
Eight	B8	336	-	426	878	103059	103060	103061	103062	D-10029 E8-H8	Intron 2	-	-	-
	D4	317	1050	313	994	103123	103124	103125	103127	D-10030 E4-H4	K300K silent E301K	CODMb CODMa/c	S-103127	B2
	E6	372	955	336	957	103279	103282	103283	103284	D-10031 A6-D6	T283I	CODMa/c	S-103284	C1
	F2	442	939	586a 427b	745a 892b	103248a	103249	103250b	103251	D-10031 E2-H2	Intron 3,b	-	-	-
	F9	337	-	386	935	103312	103314	103315	103316	D-10031 E9-H9	Intron 3	-	-	-
	G2	523	847	505	813	103357	103358	103359	103360	D-10032 A2-D2	Intron 2	-	-	-
	G3	385	984	373	933	103366	103367	103369	103370	D-10032 A3-D3	Intron 3	-	-	-
	G4	554	814	528	769	103378	103379	103381	103382	D-10032 A4-D4	S176S silent	-	-	-
	G8	467	904	452	863	103419	103420	103421	103422	D-10032 A8-D8	Intron 2	-	-	-
	G12	556	823	533	785	103456	103458	103459	103460	D-10032 A12-D12	I262V	CODMb	S-103458	C1
	H2	604	833	489a 519b	823a 796b	103361	103362	103364	103365	D-10032 E2-H2	Intron 2, a Not found, b	-	-	-
	H4	552	817	538	788	103383	103385	103386	103387	D-10032 E4-H4	S176F	CODMb	S-103383	C1

Pool plate	Well	Screening		Rescreening		Plant Sd numbers (Mutation carrier highlighted)				DNA Sources	Sequencing Result	CODM copy affected	Seed batch number of mutant	Population
		Fragment 1 (bp)	Fragment 2 (bp)	Fragment 1 (bp)	Fragment 2 (bp)									
Nine	A1	444	-	422	900	103466	103467	103468	103469	D-10033 A1-D1	Intron 3	-	-	-
	A7	537	817	520	786	103529	103530	103532	103534	D-10033 A7-D7	P266S	CODMb	S-103534	C1
	B12	192	1166	506	816	103647	103648	103626	103627	D-10033 E12-H12	Intron 2	-	-	-
	C3	450	902	436	873	103583	103584	103585	103586	D-10034 A3-D3	Intron 2	-	-	-
	C6	424	934	400	875	103618	103619	103620	103621	D-10034 A6-D6	Intron 3	-	-	-
Ten	B4	467	876	454	840	103781	103784	103785	103789	D-10035 E4-H4	Intron 2	-	-	-
	C5	222	-	422	859	103891	103892	103893	103894	D-10036 A5-D5	Intron 3	-	-	-
	F4	425	935	417	881	103934	103935	103936	103937	D-10037 E4-H4	Intron 3	-	-	-
	G11	644	731	622	705	104359	104281	104282	104283	D-10038 A11-D11	L232F	CODMb	S-104281	B3
	H9	400	936	383	889	104331	104340	104321	104322	D-10038 E9-H9	Intron 3	-	-	-
Eleven	A1	483	849	472	821	104191	104192	104193	104194	D-10039 A1-D1	Intron 2	-	-	-
	A10	573	787	548	754	104529	104530	104531	104532	D-10039 A10-D10	Q254*	CODMb	S-104530	B3
	A12	370	971	362	939	104549	104550	104551	104553	D-10039 A12-D12	Intron 2	-	-	-
	B12	434	919	421	884	104554	104555	104556	104557	D-10039 E12-H12	Intron 2	-	-	-
	F2	393	971	378	930	104737	104738	104739	104741	D-10041 E2-H2	Intron 3	-	-	-
	F6	404	-	346	952	104665	104666	104667	104668	D-10041 E6-H6	G285G silent	-	-	-
Twelve	C5	467	895	451	863	105065	105066	105068	105069	D-10044 A5-D5	Intron 2	-	-	-

Table 10. Summary of identified CODM mutations in the HM2 EMS mutagenised population. Individuals in the pool of four carrying the mutation are shaded. The nature of each mutation and the copy of CODM in which it resides are documented. The shaded mutations were prioritised for further work.

3.3.2.3 Selection of best candidates for producing heritable phenotypic changes

An online tool called PARSESNP (Project Aligned Related Sequences and Evaluate SNPs) was used to assess the potential impact of the CODM mutations obtained (Table 11). Mutation severity was predicted using the PARSESNP, SIFT and SNAP programmes (Bromberg and Rost, 2007; Ng and Henikoff, 2003; Taylor and Greene, 2003). The two nonsense mutants are predicted to be deleterious as the resulting protein would miss approximately 100 amino acids from its C terminus including residues important for substrate recognition and 2-oxoglutarate binding (Hagel and Facchini, 2010a; Runguphan et al., 2012). Appendix B contains shows the nucleotides affected and the corresponding change in amino acid (in the case of missense mutations). The analysis predicts that the T283I mutations, obtained in each copy of CODM (Table 10; Pool plate 4 F3 & Pool plate 8 E6), would have a severe effect on function. The E301K mutation returned borderline PSSM and SIFT scores but it was interesting in that it lies next to an arginine involved in 2-oxoglutarate binding (Hagel and Facchini, 2010a; Runguphan et al., 2012). These two missense mutations as well the two nonsense mutations (Q254* and W261*) were chosen to be the focus of efforts to recover mutations in M2 siblings.

Nucleotide Change	Amino acid change	Copy affected	PSSM Difference	SIFT Score	SNAP Prediction (Expected accuracy)
C1428T	S176F	CODMb			Neutral (69%)
C1595T	L232F	CODMb			Neutral (89%)
A1685G	I262V	CODMb			Neutral (94%)
C1697T	P266S	CODMb			Neutral (78%)
C1714G	F271L	CODMb	6.8	0.42	Neutral (85%)
C1858T	T283I	CODMa/c	17.0	0.00	Non-neutral (63%)
C1858T	T283I	CODMb	17.0	0.00	Non-neutral (63%)
G1911A	E301K	CODMa/c	10.5	0.16	Non-neutral (58%)

Table 11. Predicted severity of discovered CODM missense mutations. Mutation severity is predicted using the PARSESNP and SIFT programmes (Ng and Henikoff, 2003; Taylor and Greene, 2003). Mutations are predicted to have a severe effect on protein function if PSSM scores are >10 and the SIFT scores are <0.05. SNAP is a neural-network based method that uses *in silico* derived protein information (e.g. secondary structure, conservation, solvent accessibility, etc.) in order to make predictions regarding functionality of mutated proteins (Bromberg and Rost, 2007). SNAP outputs indicate whether SNPs are either "neutral" in the sense that the resulting point-mutated protein is not functionally discernible from the wild-type, or they are "non-neutral" in that the mutant and wild-type differ in function.

3.3.3 Detection of alleles from other sources

3.3.3.1 Alleles with slight differences to HM2 CODM are present in other opium poppy varieties

A 96 well DNA plate with 3 replicates of DNA from 32 different poppy lines was created. These ranged from additional GSK varieties to historic germplasm and varieties sourced externally.

GSK noscapine lines	GSK thebaine lines	GSK morphine lines	Externally sourced lines
HN1	HT1	HM1	"Afghan Princess"
HN2	HT2	HM2	"Venus"
HN3	HT3	HM3	"Double Raspberry"
	HT5	HM5	"Applegreen"
		additional four historic lines	"Burgundy Frills"
			"Violetta Blush"
			"Shyama"
			"The Giant"
			"White"

Table 12. Lines used in EcoTILLING experiments for mining of additional CODM and T6ODM alleles. Noscapine lines possess the noscapine gene cluster. The thebaine lines have not been genetically characterised and any key polymorphisms leading to the thebaine phenotype are not known.

Sample DNA from additional varieties was mixed with HM2 DNA and amplified as before with the CODM F9/R5 primer combination. Heteroduplexes were allowed to form in each well and treatment with the cleavage enzyme precluded polymorphism detection on the Fragment Analyser. If there were differences in sequence unique bands would be observed on the TILLING gel (Figure 19). This approach, whereby the objective is to identify alleles in different varieties as opposed to mutant populations is often referred to as 'EcoTILLING'. Unique bands of approximately 700 bp and 400 bp were observed in a number of lanes.

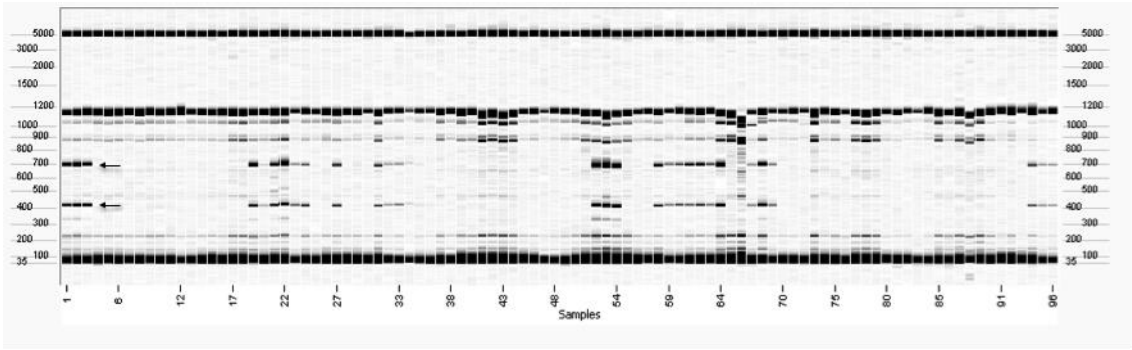


Figure 19. EcoTILLING gel to identify additional CODM alleles. Lanes 1-12 = A1-A12, 13-24 = B1-B12 etc. Unique bands of ~400 bp and ~700 bp were observed in some lanes. Those in the first three lanes are indicated with arrows.

Two polymorphic sites were found when analysing the sequencing results. They are only two bases apart which explains why only one pair of unique bands were seen on the gel in those lines that have both. Both polymorphisms would result in amino acid changes (first letter is amino acid in HM2 at this position).

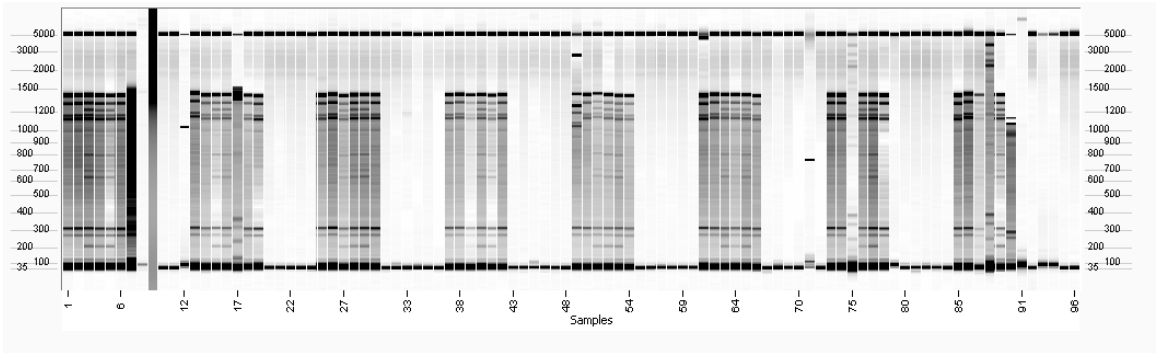
Variety	Differences from HM2 CODM
Violetta Blush	F271L
HT3	F271L
HT1	F271L
HT5	F271L (one individual)
Giant	F271L
Applegreen	F271L & I272T
Double Raspberry	F271L & I272T
Venus	F271L & I272T
Shyama	F271L & I272T
Afghan Princess	F271L & I272T
White	F271L & I272T

Table 13. Differences in CODM sequence detected among certain varieties used and HM2.

CODM in noscapine lines and other morphine lines was identical to HM2. However the thebaine lines contained an F271L polymorphism that was also present in some of the new externally sourced lines analysed (Table 12). Interestingly, HT2 (the oldest thebaine line analysed) did not have the F271L polymorphism. The fact that it is present in other lines suggests it is a natural allele of CODM in opium poppy. Curiously, the polymorphism was also introduced to HM2 during EMS mutagenesis (Table 10; Pool plate 7 G11, Sd-103228) but it is not predicted to be deleterious (Table 11; section 3.3.4 to come).

3.3.3.2 Two alleles of CODM are common to all high codeine forward screen lines

The same EcoTILLING strategy used previously (Figure 19) was employed to check whether there are any differences in the nucleotide sequence of CODM in five high codeine forward screen lines and WT HM2 from which they were derived. These five lines had undergone numerous rounds of selfing in the field in Tasmania. Capsules of these lines accumulate substantially more codeine and thebaine than HM2- codeine comprises typically 30% of the total alkaloid content. The five seed batches were sent to York for DNA extraction and analysis. DNA from the resulting seedlings was spiked with equal amounts of HM2 wild type DNA prior to PCR. Differences in sequence between wild type and mutant CODM would result in heteroduplexes and cleavage. Figure 20 shows the discovery of two new CODM polymorphisms in three of the high codeine lines S-122066 (MAS 1983 C), S-122067 (MAS 1833 B) and S-122068 (MAS 1835 D). Sequencing CODMa/c and CODMb in these and individuals of S-114247 (MAS 2235 D) and S-114248 (MAS 2237 D) revealed the nature of these polymorphisms to be an E193K substitution in CODMa/c and R158K substitution in CODMb (Figure 20). In addition there is also a silent substitution in CODMb-S321S. The E193K mutation was predicted to be non-neutral by the SNAP programme (Bromberg and Rost, 2007). The three nucleotide substitutions are G/C to A/T transitions and so it is likely that they were EMS induced (Greene et al., 2003). All individuals genotyped were homozygous for each of the polymorphisms, indicating that they are fixed and likely contribute to the high codeine phenotype. The two remaining high codeine forward screen lines, S-114247 (MAS 2235 D) and S-114248 (MAS 2237 C), were subsequently shown to possess both of the CODM polymorphisms by AS-PCR rather than TILLING and sequencing (Appendix B). The five high codeine lines had similar phenotypes in the glasshouse in York as they did in the field in Tasmania, demonstrating the stability of the phenotype.



CNAP seed batch id	GSK seed batch id	Original M2 seed batch (Population)	GSK capsule data Grown in field in Tasmania (%DW)				CNAP capsule data Grown in glasshouse in York (%DW)					Polymorphisms identified with Ecotilling	Gene affected
			M	O	C	T	M	O	C	T	CD dimer		
S-114247	MAS 2235 D	S-103234 (C1)	47.5	0.6	31.2	20.6	61.6	0.2	24.1	6.8	5.5	**E193K R158K**	CODMa/c CODMb
S-114248	MAS 2237 C	S-103234 (C1)	50.9	0.5	34.7	13.9	60.0	0.2	26.5	5.8	5.9	**E193K R158K**	CODMa/c CODMb
S-122068	MAS 1835 D	S-103234 (C1)	45.3	0.6	33.1	21.0	47.8	0.3	27.4	15.7	6.7	**E193K R158K**	CODMa/c CODMb
S-122066	MAS 1983 C	S-103234 (C1)	51.5	1.0	28.6	18.9	44.1	0.4	30.3	15.7	6.5	**E193K R158K**	CODMa/c CODMb
S-122067	MAS 1833 B	S-103234 (C1)	43.8	1.6	19.7	34.9	54.5	1.8	13.8	26.6	1.7	**E193K R158K**	CODMa/c CODMb

Figure 20. 'EcoTILLING' high codeine lines for CODM polymorphisms. Two pairs of fragments are visible in various lanes-200 bp +1200 bp & 600 bp +800 bp. In addition a single band of 700 bp is visible. These were caused by CODM E193K, R158K and S321S polymorphisms present in all five high codeine forward screen lines. These polymorphisms may be associated with the high codeine phenotypes (expressed as a percentage of total alkaloids in mature capsules) observed both in the field in Tasmania and in the glasshouse in York.

3.3.3.3 All five high codeine lines are derived from the same M2 plant

All five high codeine lines derived from the same original M2 seed batch, S-103234 of the highly mutagenised C1 subpopulation. This seed batch was actually included in the reverse genetic screen but there was a PCR failure with well E1 of pool plate 8 which contained DNA from Sd-103234 (Appendix B; Pool plate 8, lane 49).

Carol Walker of GSK kindly provided genealogy records for the five high codeine forward screen lines demonstrating how they all derived from the M2 seed batch S-103234. The records show a doubling in the codeine content of capsules between C1M3 RMUUS 314 and C1M4 MUVS 107 (Figure 21; Appendix B). The CODM polymorphisms may have been in segregating in RMUUS 314. Seed was then collected from RMUUS 314 E which gave rise to a number of MUV 107 plants homozygous for the CODM polymorphisms including MUVS 107 A, whose codeine content was four times higher than the mean codeine content of plants in the previous generation.

MAS 1983 C

S-122066

popn	season	old ID	seed ID		plot no	M	C	O	T
				OP row	**HM2**	**3.1117**			
C1M2	07_08	**25c**	**S-103234**	**OP row**	MUT 437	2.9664	0.2557	0.0150	0.2019
				SP caps	MUTS 437 G	4.4067	0.5943	0.0060	0.0769
C1M3	08_09			**OP row**	**HM3**	**3.67**			
				OP row	RMUU 314	3.7286	0.6138	0.0011	0.2387
				SP caps	RMUUS 314 E	4.2614	0.6006	0.0136	0.0813
C1M4	09_10			**OP row**	**HM1**	**2.918**	**0.332**	**0.027**	**0.121**
				OP row	MUV 107	2.352	1.391	0.051	1.105
				SP caps	MUVS 107 A	2.315	2.316	0.024	1.209
C1M5	**2010**	**Seed from MUVS 107 A sent to Tom Davies to grow over Australian winter. Seed harvested and returned labelled UKMUVS 107 A**							
					Seed sown into 5 replicated rows in TAS				
						M+pM	**C**	**O**	**T**	**M**
C1M6	10_11	**REP 4/5**		**OP row**	**HM4**	**2.569**	**0.144**	**0.015**	**0.105**	**2.352**
		ex ukmuvs 107a		**OP row**	**HT4**	**1.361**	**0.541**	**0.23**	**1.862**	**1.163**
		(C1M5 2010)		**OP row**	MUWT 1964	2.356	0.709	0.017	0.429	2.073
				SP caps	MUWTS 1964 C	2.266	2.001	0.059	0.696	2.154
C1M7	11_12					**M**	**C**	**O**	**T**	**Total alk**
		ex muwts 1964 C		**OP row**	**HM5**	**3.373**	**0.129**	**0.017**	**0.116**	**3.634**
		(C1M6 10-11)		**OP row**	**HT5**	**1.001**	**0.189**	**0.709**	**2.473**	**4.372**
				OP row	MA 1983	3.005	1.691	0.040	1.316	6.052
				SP caps	**MAS 1983 C**	3.766	2.089	0.076	1.379	7.309

Figure 21. Development of the high codeine forward screen line S-122066 (MAS 1983 C) from S-103234. In each season the phenotypes are compared to the current morphine (HMx) and thebaine (HTx) lines. OP= open pollinated. SP= self pollinated. Phenotypes expressed as percentage dry weight of morphine (M), codeine (C), oripavine (O), thebaine (T) and pseudomorphine (pM). The best individuals in each row are highlighted. The genealogies for the four other high codeine lines are in Appendix B. They all share the same origin in S-103234.

3.3.4 Section summary

Although it was clear that there were multiple copies of CODM in the genome it was decided to screen all copies simultaneously for EMS-induced mutations. PCR amplification and sequencing would then determine whether any detected mutation would reside on CODMa/c or CODMb. Nonsense mutations were found in the TILLING screen which became the focus of efforts to recover mutations among M2 siblings. Furthermore, two CODM

polymorphisms were detected in high codeine forward screen lines. These may cause or be linked to the cause of the high codeine phenotype.

3.4 Recovery of mutations in M2 siblings and assessment for phenotypic change

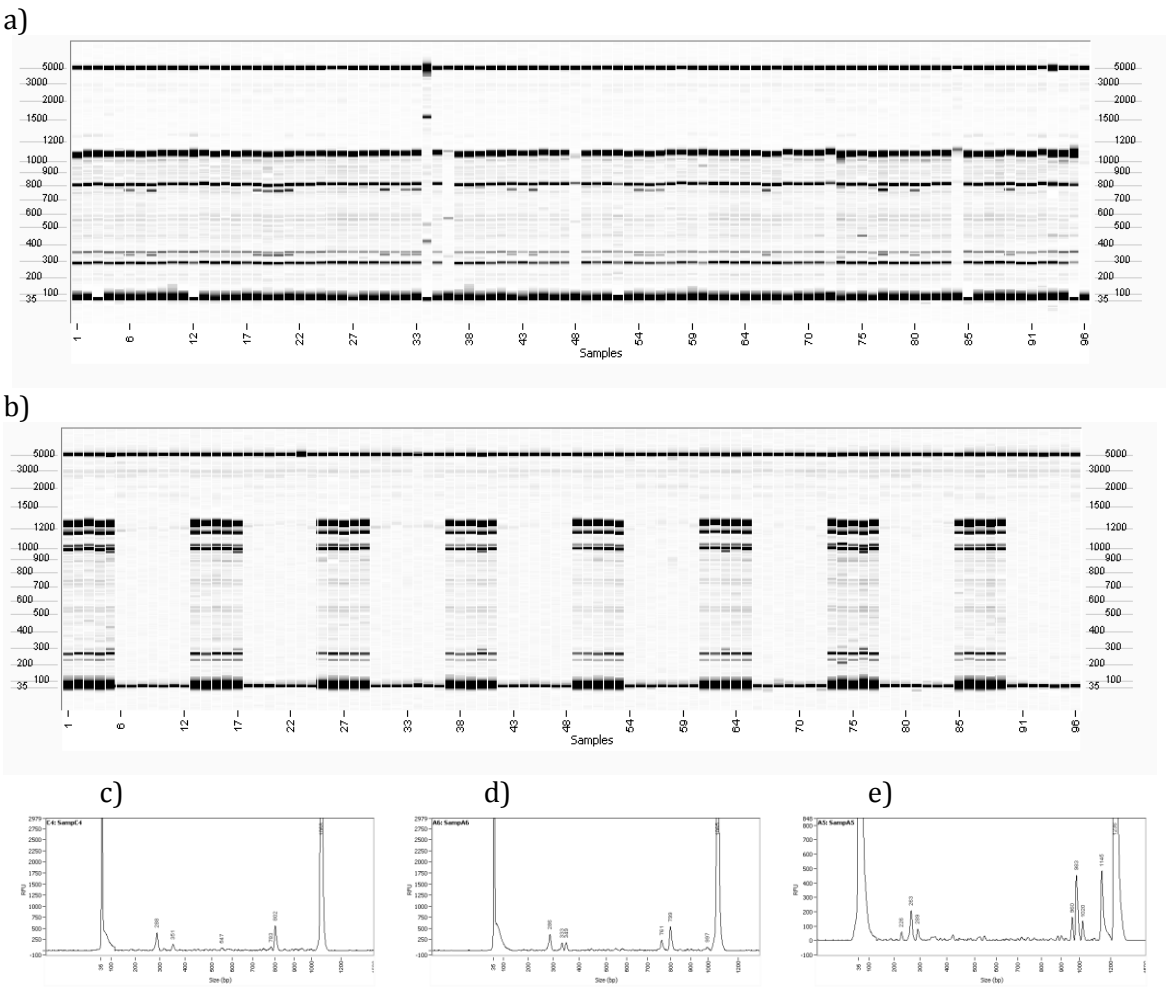


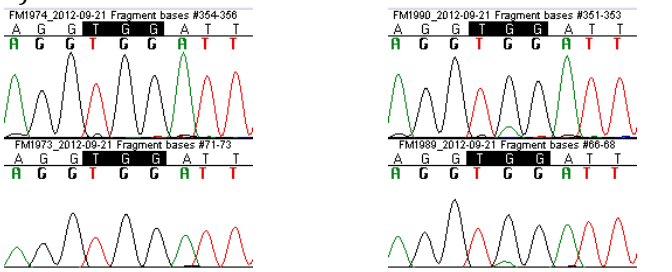
Figure 22. Recovery of CODM mutations
a) Q254*, W261* and T283I mutation recovery plate. Common fragments were produced due to differences between copies of CODM. However sets of fragments in a number of lanes indicate the presence of mutations.
b) E301K mutation recovery plate.
c) Example of W261* mutation detection in well C4 (lane 28 on gel a). The W261* mutation generated fragments of 783 bp and 351 bp.
d) Example of Q254* mutation detection in well A6 (lane 6 on gel a). The Q254* mutation generated fragments of 761 bp and 333 bp.
e) Example of E301K mutation detection in well A5 (lane 5 on gel b). The E301K mutation generated fragments of 960 bp and 289 bp.

The M2 seed for each missense and nonsense mutant was located and sown. DNA was extracted from 4-week-old seedlings derived from each seed batch. These DNA samples were then screened for the CODM mutations initially found in their M2 sibling. The W261* and Q254* mutations were recovered (Figure 21a,c,d). The seed from the CODMb and CODMa/c T283I mutants gave rise to just 2 and 18 plants, respectively. Unfortunately, the mutation

was not detected in any of these M2 siblings (Figure 21a). For this reason, DNA from siblings of the next most interesting CODMa mutant, E301K, was extracted and screened. The E301K mutation was recovered in 5 seedlings (lanes 5, 17, 40, 63 and 76) of the 34 screened (Figure 21b,e).

3.4.1 Sequencing to confirm W261 and E301K CODMa/c mutations*

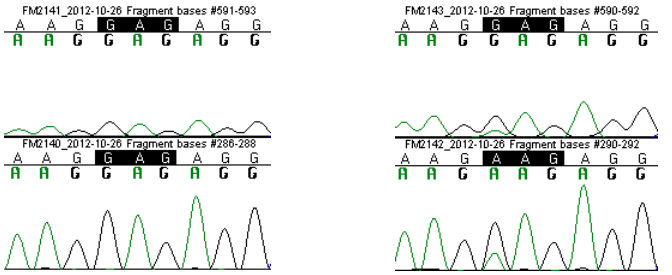
a)



Sd-828044 example of WT

Sd-828047 example of 'carrier'

b)



Sd-828062 example of WT

Sd-828068 example of 'carrier'

Figure 23. Sequencing CODM in CODMa/c mutants

a) Examples of sequencing traces of the codons affected by the W261* mutation. WT CODM has TGG encoding tryptophan. The mutant has TAG stop codon.

b) Examples of the sequencing traces of the codons affected by the E301K mutation. WT CODM was a GAG encoding glutamic acid. The mutant has a AAG encoding lysine.

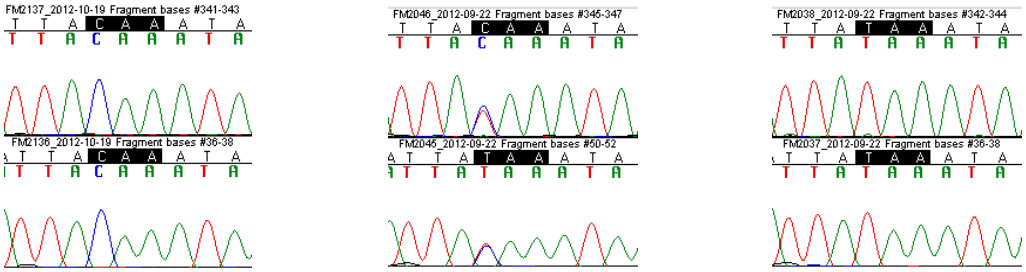
S-102574 seed was sown and 38 seedlings were screened for the W261* mutation. 9 carriers were found. It is not appropriate to designate them heterozygotes as in each case there are not two equally sized peaks in the chromatogram at the mutation position (Figure 23a). In fact the green peak (A) is always a quarter of the size of the black peak (G). This results from amplification of both CODMa and CODMc in PCR with primers F12/R7. This mixed PCR product was the template in the sequencing reaction. In a heterozygote there was one mutant template for every three wild type templates. It is likely that these 9 plants are

heterozygous mutants of either CODMa or CODMc. Likewise S-103127 seed was sown and 34 seedlings were screened for the E301K mutation. 5 carriers were recovered (Figure 23b).

3.4.2 Sequencing to determine zygosity of carriers of the Q254 mutation in CODMb*

S-104530 seed was sown and 32 seedlings were screened for the Q254* mutation. 8 homozygotes and 8 heterozygotes were found.

Figure 24 shows examples of the sequencing traces from a wild type plant, a heterozygote and a homozygote. In this case only CODMb was amplified in PCR (using primer pair F12/R8) prior to sequencing. Heterozygotes had WT and mutant signals of equal magnitude at the position of the mutation. Homozygotes displayed a complete change of base from C to T at this position. This confirms that the primers used selectively amplifies CODMb and indicates that there is only one copy of CODMb in the genome.



Sd-828227 example of WT of heterozygote

Sd-828202 example of homozygote

Sd-828201 example

Figure 24. Sequencing CODM in CODMb mutants. Examples of sequencing traces of the codons affected by the Q254* mutation. WT CODM has CAA encoding glutamine. The mutant has a TAA stop codon.

3.4.3 Analysis of M2 capsules-CODMb Q254 mutants in particular have elevated codeine levels*

Capsules of M2 siblings were ground and their alkaloid contents were analysed by UPLC. All capsules analysed had morphine as the predominant alkaloid, much like the wild type HM2 from which the mutant population was derived. Introduction of CODM mutations, therefore, did not result in a complete block of metabolism at codeine. This is not surprising since even with one copy knocked out, having two remaining functional copies of CODM appears to be sufficient to demethylate codeine to morphine at levels similar to that of HM2. The mean (±SD) codeine yields among CODMa/c W261* carriers (homo & het) and CODMb Q254* homozygotes were 4.0± 1.6 and 3.0+2.1 µg/mg dw, respectively. This compares with a

HM2 average of 1.5±0.8 µg/mg dw. In terms of percentage of total alkaloids, the capsules of CODMa/c W261* carriers (homo & het) and CODMb Q254* homozygotes had 6.4% and 6.7% codeine, respectively, while those of HM2 had 2.8%. Likewise, thebaine yields in CODMa/c W261* carriers (homo & het) and CODMb Q254* homozygotes were 3.1± 3.0 and 1.7+2.6 µg/mg dw, respectively, while that of HM2 was 1.0±0.6 µg/mg. The high standard deviations are likely to reflect the variety of background mutations among siblings. Nevertheless, the codeine (Figure 25) contents of CODMa/c W261* mutants and codeine and thebaine contents in the CODMb Q254* mutants suggested partial blocks of metabolism due to the induced mutations. M3 seed collected from the capsules prior to phenotypic analysis was sown for further analysis. Crosses to combine CODM mutations would be attempted with the resulting M3 plants. Double or triple mutants would be required to produce lines that accumulate much higher levels of codeine.

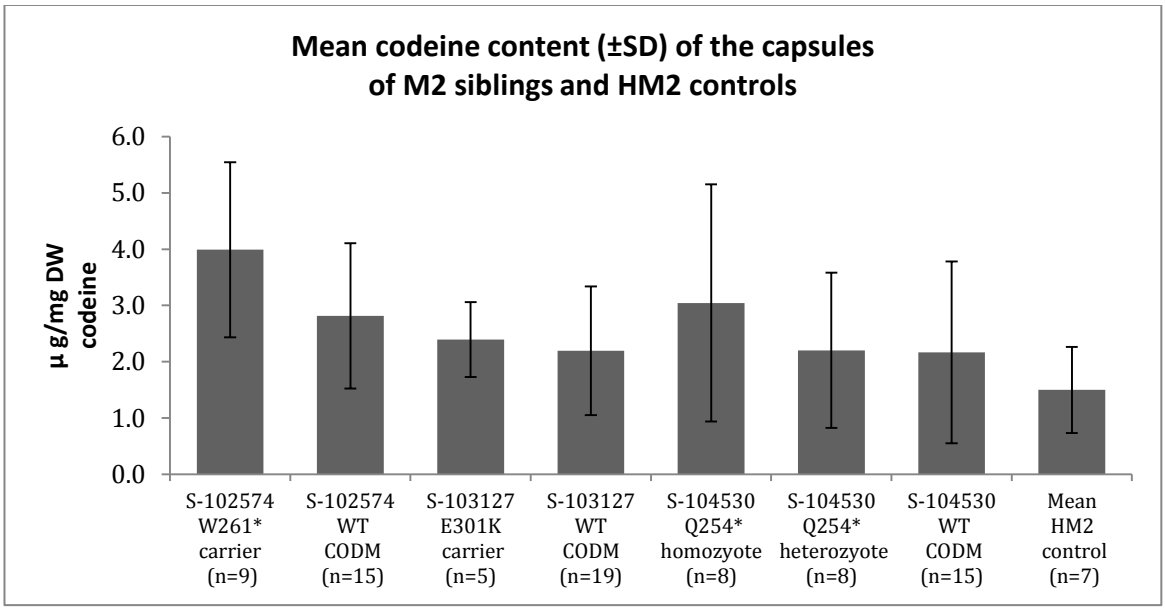


Figure 25. Mean codeine content (±SD) of the capsules of M2 siblings and wild type HM2.

3.4.4 Analysis of M3 capsules-Q254 mutants 3 fold more codeine than HM2*

M3 seedlings were genotyped using allele specific PCR to determine whether the plants carried a CODM mutation prior to being used as a parent in a cross (section 3.5.5 below). CODMa/c W261* and E301K mutations were still segregating in M3 as the M2 parents were heterozygous for these mutations. This led to some variation in the phenotypes as

homozygotes, heterozygotes and wild types were represented among siblings derived from seed batches S-114020, S-114037 and S-114110 (Figure 26).

The two M3 CODMb Q254* seed batches sown were produced by M2s known to be homozygous for the mutation. Therefore, all M3s were also homozygous. In M3 plants derived from S-114055, the mean codeine content of the capsules (5.8±0.6 µg/mg; 10.5% of total alkaloids) was up to three times (Figure 26 and Appendix B) higher than wild type HM2 (1.8±1.1 µg/mg; 3.6% of total alkaloids). Likewise, thebaine content in the capsules (6.2±3.5 µg/mg; 11.2% of total alkaloids) was up to four times (Appendix B) higher than wild type HM2 (1.5±1.4 µg/mg; 3.0% of total alkaloids). Full alkaloid profiles for each line are displayed in Appendix B.

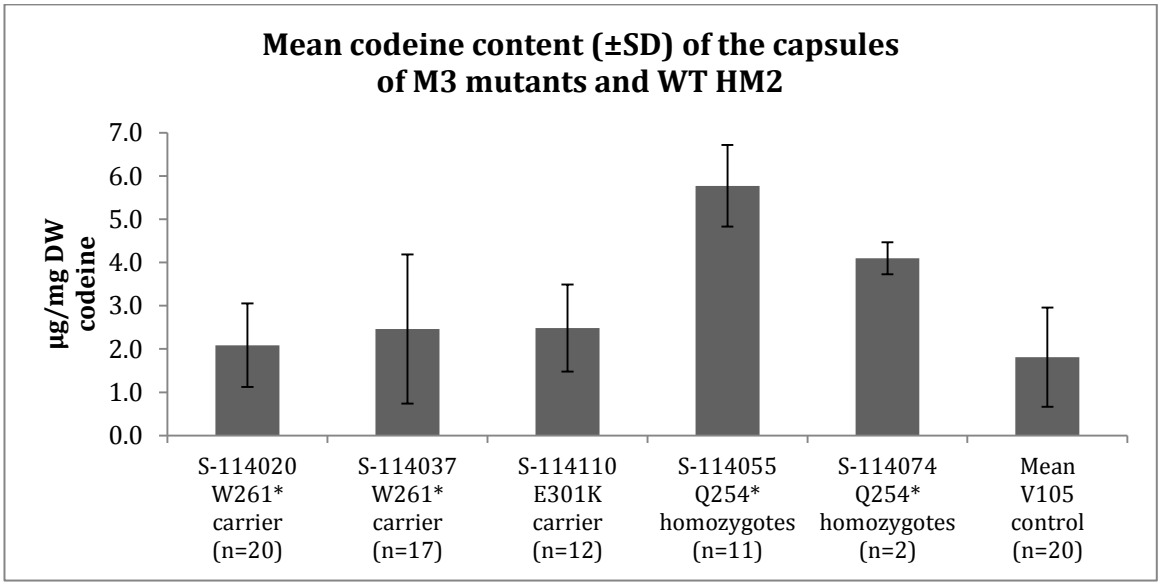
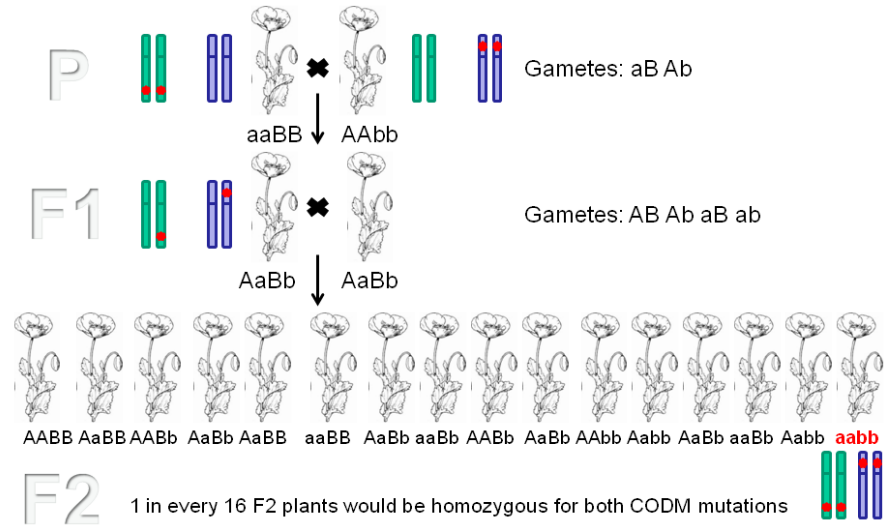


Figure 26. Mean codeine content (±SD) of the capsules in M3 CODM mutants and wild type HM2. W261* and E301K mutations were still segregating among M3s. 25% would be expected to be homozygous, 50% heterozygous and the remaining 25% WT. All Q254* M3s were homozygous for the CODMb mutation.

3.4.5 Attempts to create CODM double mutants to further impact on disrupting O-demethylation to morphine

T6ODMa and PODA display no activity against codeine (Hagel and Facchini, 2010a). Therefore, disrupting all copies of CODM in a plant should be sufficient to block metabolism at codeine. Crosses were set up to combine the CODM mutant alleles identified in the various copies in the TILLING screen. A CODM double or triple mutant may be enough to achieve a codeine poppy or at least to substantially increase the yields of codeine relative to morphine.

a)



b)

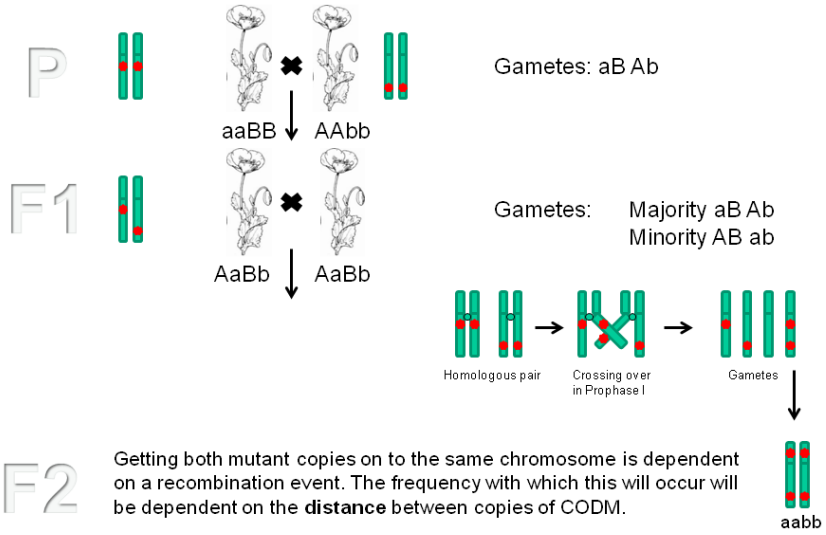


Figure 27. Possible scenarios involved with the recovery of CODM double mutants. a) Copies of CODM are unlinked i.e. on different chromosomes and 1 in every 16 F2s should be homozygous for two CODM mutations. b) Copies of CODM are linked and the recovery of a double mutant will be dependent on a crossing over event. The frequency with which this will occur is dependent on the distance between copies of CODM.

There is evidence from our BAC library to suggest that at least two copies of CODMc are linked in sequence held on BAC 89C05 (Figure 15). Furthermore, CODMa and CODMc were cloned and sequenced from BAC 158A11 (Table 7). If copies of CODM were in fact linked, crosses among M3s with different CODM mutations would produce a uniform heterozygous population, with each plant containing one CODM-containing chromosome from each parent. During meiosis the homologous chromosomes may or may not 'cross over' and form

recombinants (Figure 27b). The closer the copies are to each other the less chance that recombination bringing two different CODM mutations together will occur. If copies of CODM are unlinked the strategy would involve following the crosses through to the F2 generation and identifying double homozygotes (1 in every 16 F2s; Figure 27a).

Eight E301K x W261*, three E301K x Q254* and five W261* x Q254* crosses were successfully carried out. F1 seed was sown and ten F1 individuals from each seed batch was sown for genotyping to confirm transmission of the CODM mutant alleles (Appendix B). F1 plants confirmed to possess two CODM mutations were prioritised for selfing to produce F2 seed (Appendix B).

Five F2 seed batches resulting from W261* x Q254* crosses: S-123066, S-123067, S-123072, S-123079 and S-123086 were sown. S-123028 represented an F2 from a W261* x E301K cross and S-123055 was the F2 seed sown which resulted from a E301K x Q254* cross. 40 F2s from each of the 7 seed batches were genotyped by AS-PCR and sequencing (Appendix B). No double mutants were recovered. The vast majority of F2s were heterozygous for each CODM mutation introduced by the parents in the original cross. In the small number of cases where an F2 individual was homozygous for one CODM mutation, it lost the other one. This suggests that the CODMa/c and CODMb copies are closely linked. These conclusions were confirmed by sequencing results at the time where copies of CODMa and CODMc were amplified and cloned from BAC 158A11 (Table 7) and later PacBio sequencing of BAC 89C05 where two copies of CODMc were found to reside on the same BAC (Figure 15). A BAC typically has a size of about 150 kb and so at least two copies of CODM must be within this range of each other.

Overall, the results suggest that multiple copies of CODM reside in a relatively small region of the poppy genome. However, knocking out just one copy of those identified produced already elevated levels of codeine. Crosses between CODM mutants may be difficult due to close linkage. It is worth stating that the BAC sequencing results were not available at the time the crosses were being set up (Chapter 5) and so efforts to combine CODM

mutations were initiated. As an alternative strategy to crosses, seed from the CODMb Q254* mutant was bulked up for further EMS mutagenesis. Perhaps multiple rounds of mutagenesis will eventually be required to identify knockouts in all copies or at least those copies which are expressed and most active.

3.4.6 Section summary

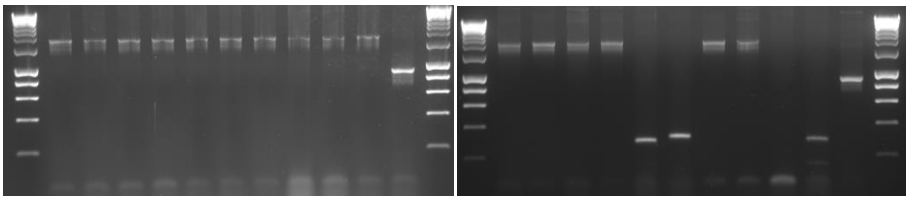
The CODM nonsense mutations (W261* and Q254*) and the E301K polymorphism were recovered in M2 siblings. Alkaloid content of M2 and M3 capsules were assessed which indicated partial increases in codeine content compared to HM2. In an effort to increase the yield of codeine further, crosses were attempted to produce plants with multiple mutated copies of CODM. However, F2s homozygous for two CODM mutations were not identified suggesting close linkage between copies of CODM in the genome.

3.5 Numerous copies of T6ODM in the opium poppy genome

3.5.1 Sequencing of T6ODM from genomic DNA template

The NEB LongAmp PCR kit was used initially to amplify T6ODM from HM2 gDNA using primers designed with the published cDNA sequence (GenBank accession: GQ500139) and matching contigs in the EST database. A ~2 kb PCR product was generated with primers LA F1/R2. This PCR product was purified and sequenced. The results indicated the presence of three introns, similar to CODM. However, intron 2 was much longer than intron 2 in CODM. The sequence information obtained was used to design additional primers for T6ODM amplification.

Primer pairs F2/R2 and LA F4/R3 produced similarly large PCR products with HM2 gDNA. F2/R2 would amplify the 5' region of the gene while LA F4/R3 would amplify the 3' end including the third intron and exon 4. PCR products generated with these primer pairs were then cloned and sequenced with the Strataclone Blunt PCR cloning kit. Colony PCR confirmed the presence of the correct sized inserts (Figure 28).

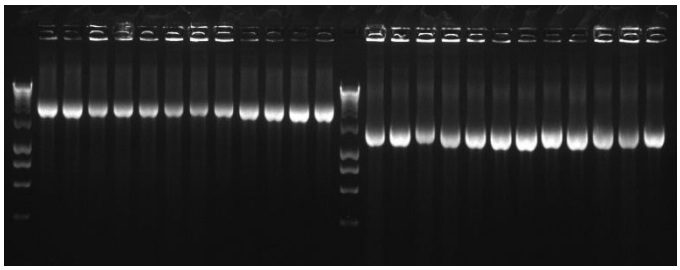


T6ODM F2/R2 T6ODM LA F4/R3

Figure 28. M13 colony PCRs with positive T6ODM F2/R2 and LA F4/R3 transformations. Lane 11 in each case contained the control insert supplied with the Strataclone Blunt PCR cloning kit.

Plasmids from 5 each of the F2/R2 and LA F4/R3-containing colonies were purified and sequenced. This led to the confirmation of three introns in T6ODM when the cDNA and gDNA sequences were compared in a sequence alignment (Table 14, Figure 31). Furthermore, there were numerous polymorphisms between each of the clones and the published T6ODM cDNA sequence, indicating that the primers may have amplified members of a multi gene family. In addition to SNPs there were small differences in the sizes of introns i.e. insertions or deletions evident in the alignment of cloned T6ODM sequence.

Amplification of full length CODM, T6ODM and DIOX2/PODA gene sequence from HN1 gDNA template proved difficult with Phusion polymerase. These difficulties were overcome with GoTaq polymerase. Although this enzyme has a higher reported error rate than Phusion it was decided to use it to amplify as large a portion of these 2-ODD genes as possible. Multiple clones for each would be sequenced so any variation would have to be observed more than once in order to be considered a genuine SNP and distinguish it from possible errors introduced by GoTaq polymerase. T6ODM primer pair F1/R4 generated a 1374 bp PCR product with HN1 cDNA template as expected and a >2kb product with HN1 gDNA template. These PCR products were cloned with the Strataclone PCR cloning kit. Twelve colonies were analysed by colony PCRs (generated from cDNA and gDNA templates) for presence of the inserts (Figure 29). Plasmids from six of each were purified and sequenced.



HN1 gDNA template **HN1 cDNA template**

Figure 29. M13 colony PCR with positive T6ODM F1/R4 transformations.

3.5.2 Numerous copies of T6ODM reside in the genome of opium poppy

The sequences of different clones generated to date were compared

- 10 with HM2 gDNA as template
- 6 with HN1 gDNA as template
- 6 with HN1 cDNA as template
- published T6ODM cDNA sequence

Table 14 lists the bases observed in consensus sequences of clones of T6ODM gDNA and cDNA. Patterns of SNPs allowed grouping together of different clones. The results indicate that there are at least seven different copies of T6ODM in the genome. To date, however, there is only evidence of expression of two of these (cDNA sequences only in the green and blue groups). These are the copies that are expressed. The others appear not to be expressed as suggested by the lack of cDNA and ESTs and in two cases mutations have accumulated at different sites that introduce premature stop codons. If these were expressed they would not be functional. The green group of sequences is an exact match for the published T6ODM cDNA sequence (GenBank accession: GQ500139) so it has been assigned T6ODMa. The blue set is to be called T6ODMb. PODA/DIOX2 (GenBank accession: GQ500140) was not among the amplified clones.

SNP	Exon 1					Intron 1		Exon 2	
	138 A/G	236 A/C	237 A/G	257 A/G	271 T/C	368 G/T	381 A/C	519 A/G	545 A/C
published cDNA	G	A	A	A	T	-	-	G	A
HN1_cDNA(12)	G	A	A	A	T	-	-	G	A
HN1_cDNA(4)	G	A	A	A	T	-	-	G	A
HN1_cDNA(8)	G	A	A	A	T	-	-	G	A
HN1_gDNA(11)	G	A	A	A	T	G	A	G	A
HN1_gDNA(3)	G	A	A	A	T	G	A	G	A
HM2_gDNA(6)	G	A	A	A	T	G	A	G	A
HN1_cDNA(2)	A	C	G	G	C	-	-	A	C
HN1_cDNA(6)	A	C	G	G	C	-	-	A	C
HN1_cDNA(10)	A	C	G	G	C	-	-	A	C
HM2_LA_gDNA_(4)	-	-	-	-	-	-	-	A	C
HN1_gDNA(7)	A	C	G	G	C	G	C	A	C
HM2_gDNA(8)	A	C	G	G	C	G	C	A	C
HM2_gDNA(10)	A	C	G	G	C	G	C	A	C
HM2_LA_gDNA_(8)	-	-	-	-	-	-	-	A	C
HM2_LA_gDNA_(1)	-	-	-	-	-	-	-	A	C
HM2_LA_gDNA_(7)	-	-	-	-	-	-	-	A	C
HN1_gDNA(1)	A	C	A	G	C	T	C	G	C
HN1_gDNA(5)	A	C	A	G	C	T	C	G	C
HM2_gDNA(3)	A	C	A	G	C	T	C	G	C
HM2_LA_gDNA_(2)	-	-	-	-	-	-	-	G	C
HN1_gDNA(9)	A	C	G	G	C	G	C	G	C
HM2_gDNA(1)	A	C	G	G	C	G	C	G	C

SNP	Intron 2																										
	658 T/C	660 A/C	672 G/A	708 G/A/C	719 C/A	743 T/C	883 G/T	894 A/C	947 C/A	969 A/T	986 C/T	1020 C/T	1022 C/T	1030 T/C	1044 G/A	1066 A/G	1098 C/A	1128 T/C	1143 A/G	1167 -/T	1168 C/T/A	1177 C/T	1195 A/G	1199 G/A	1202 C/A	1030 C/T	1044 G/A
published cDNA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HN1_cDNA(12)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HN1_cDNA(4)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HN1_cDNA(8)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HN1_gDNA(11)	T	A	G	G	C	T	G	A	C	T	C	C	C	T	G	A	C	T	A	-	C	C	A	G	C	T	G
HN1_gDNA(3)	T	A	G	G	C	T	G	A	C	T	C	C	C	T	G	A	C	T	A	-	C	C	A	G	C	T	G
HM2_gDNA(6)	T	A	G	G	C	T	G	A	C	T	C	C	C	T	G	A	C	T	A	-	C	C	A	G	C	T	G
HN1_cDNA(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HN1_cDNA(6)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HN1_cDNA(10)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HM2_LA_gDNA_(4)	C	A	G	A	C	C	G	C	C	T	T	C	C	C	G	G	C	C	G	-	T	T	G	G	C	C	G
HN1_gDNA(7)	C	A	G	A	C	C	G	C	C	T	T	C	C	C	G	G	C	C	G	-	T	T	G	G	C	C	G
HM2_gDNA(8)	C	A	G	A	C	C	G	C	C	T	T	C	C	C	G	G	C	C	G	-	T	T	G	G	C	C	G
HM2_gDNA(10)	C	A	G	A	C	C	G	C	C	T	T	C	C	C	G	G	C	C	G	-	T	T	G	G	C	C	G
HM2_LA_gDNA_(8)	C	A	G	A	C	C	G	C	C	T	T	C	C	C	G	G	C	C	G	-	T	T	G	G	C	C	G
HM2_LA_gDNA_(1)	C	A	G	A	C	C	G	C	C	T	T	C	C	C	G	G	C	C	G	-	T	T	G	G	C	C	G
HM2_LA_gDNA_(7)	C	A	G	A	C	C	G	C	C	T	T	C	C	C	G	G	C	C	G	-	T	T	G	G	C	C	G
HN1_gDNA(1)	T	C	A	A	A	T	T	C	C	A	C	T	T	T	A	G	A	T	A	-	T	C	A	A	C	T	A
HN1_gDNA(5)	T	C	A	A	A	T	T	C	C	A	C	T	T	T	A	G	A	T	A	-	T	C	A	A	C	T	A
HM2_gDNA(3)	T	C	A	A	A	T	T	C	C	A	C	T	T	T	A	G	A	T	A	-	T	C	A	A	C	T	A
HM2_LA_gDNA_(2)	T	A	G	C	C	T	G	C	A	T	C	T	C	T	G	G	C	T	A	T	A	C	A	G	A	T	G
HN1_gDNA(9)	T	A	G	C	C	T	G	C	A	T	C	T	C	T	G	G	C	T	A	T	A	C	A	G	A	T	G
HM2_gDNA(1)	T	A	G	A	C	C	G	C	C	T	T	C	C	C	G	G	C	C	G	-	T	T	G	G	C	C	G

SNP	Intron 2 (cont.)																										
	1066 A/G	1098 C/A	1128 C/T	1143 A/G	1167 -/T	1168 C/T/A	1177 C/T	1195 A/G	1199 G/A	1202 C/A	1218 G/T	1223 C/T	1229 C/A	1239 G/A	1285 A/G	1286 T/C	1289 C/A/-	1303 C/T	1350 T/C	1360 T/G	1368 G/T	1371 C/T	1373 T/C	1382 T/G	1386 A/G	1394 C/T	1452 C/T
published cDNA	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HN1_cDNA(12)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HN1_cDNA(4)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HN1_cDNA(8)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HN1_gDNA(11)	A	C	T	A	-	C	C	A	G	C	G	C	C	G	A	T	C	C	T	T	G	C	T	T	A	T	T
HN1_gDNA(3)	A	C	T	A	-	C	C	A	G	C	G	C	C	G	A	T	C	C	T	T	G	C	T	T	A	T	T
HM2_gDNA(6)	A	C	T	A	-	C	C	A	G	C	G	C	C	G	A	T	C	C	T	T	G	C	T	T	A	T	T
HN1_cDNA(2)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HN1_cDNA(6)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HN1_cDNA(10)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HM2_LA_gDNA_(4)	G	C	C	G	-	T	T	G	G	C	G	C	A	G	G	C	A	C	T	T	G	T	C	G	A	C	T
HN1_gDNA(7)	G	C	C	G	-	T	T	G	G	C	G	C	A	G	G	C	A	C	T	T	G	T	C	G	A	**T**	T
HM2_gDNA(8)	G	C	C	G	-	T	T	G	G	C	G	C	A	G	G	C	A	C	T	T	G	T	C	G	A	C	T
HM2_gDNA(10)	G	C	C	G	-	T	T	G	G	C	G	C	A	G	G	C	A	C	T	T	G	T	C	G	A	C	T
HM2_LA_gDNA_(8)	G	C	C	G	-	T	T	G	G	C	G	C	A	G	G	C	A	C	T	T	G	T	C	G	A	C	T
HM2_LA_gDNA_(1)	G	C	C	G	-	T	T	G	G	C	G	C	A	G	G	C	A	C	T	T	G	T	C	G	A	C	T
HM2_LA_gDNA_(7)	G	C	C	G	-	T	T	G	G	C	G	C	A	G	G	C	A	C	T	T	G	T	C	G	A	C	T
HN1_gDNA(1)	G	A	T	A	-	T	C	A	A	C	T	T	A	A	A	T	-	T	C	G	G	T	C	T	A	T	T
HN1_gDNA(5)	G	A	T	A	-	T	C	A	A	C	T	T	A	A	A	T	-	T	C	G	G	T	C	T	A	T	T
HM2_gDNA(3)	G	A	T	A	-	T	C	A	A	C	T	T	A	A	A	T	-	T	C	G	G	T	C	T	A	T	T
HM2_LA_gDNA_(2)	G	C	T	A	T	A	C	A	G	A	G	T	A	A	A	T	-	T	C	T	T	T	C	T	G	T	C
HN1_gDNA(9)	G	C	T	A	T	A	C	A	G	A	G	T	A	A	A	T	-	T	C	T	T	T	C	T	G	T	C
HM2_gDNA(1)	G	C	C	G	-	T	T	G	G	C	G	C	A	G	G	C	A	C	T	T	G	T	C	G	A	C	T

SNP	Exon 3									Intron 3					Exon 4					
	1511 G/T	1522 G/C	1534 T/C	1600 G/A	1605 T/A	1665 C/A	1753 G/T	1766 T/G	1801 T/C	1835 G/A/T	1842 G/C	1865 T/C	1886 G/T	1910 C/T	1939 T/C	2017 G/A	2051 G/T	2079 C/A	2140 T/A	2155 T/G
published cDNA	G	G	T	G	T	C	G	**T**	T	-	-	-	-	-	T	G	G	C	T	T
HN1_cDNA(12)	G	G	T	G	T	C	G	G	T	-	-	-	-	-	T	G	G	C	T	T
HN1_cDNA(4)	G	G	T	G	T	C	G	G	T	-	-	-	-	-	T	G	G	C	T	T
HN1_cDNA(8)	G	G	T	G	T	C	G	G	T	-	G	T	G	C	T	G	G	C	T	T
HN1_gDNA(11)	G	G	T	G	T	C	G	G	T	-	G	T	G	C	T	G	G	C	T	T
HN1_gDNA(3)	G	G	T	G	T	C	G	G	T	-	G	T	G	C	T	G	G	C	T	T
HM2_gDNA(6)	G	G	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HN1_cDNA(2)	G	G	C	G	A	C	G	T	T	-	-	-	-	-	C	A	G	C	T	G
HN1_cDNA(6)	G	G	C	G	A	C	G	T	T	-	-	-	-	-	C	A	G	C	T	G
HN1_cDNA(10)	G	G	C	G	A	C	G	T	T	-	-	-	-	-	C	A	G	C	T	G
HM2_LA_gDNA_(4)	G	G	C	G	A	C	G	T	T	G	C	T	G	C	C	A	G	C	T	G
HN1_gDNA(7)	G	G	C	G	A	C	G	T	T	G	C	T	G	C	C	A	G	C	T	G
HM2_gDNA(8)	G	G	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HM2_gDNA(10)	G	G	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HM2_LA_gDNA_(8)	G	G	C	G	A	C	G	T	T	G	C	T	G	C	T	G	G	C	T	T
HM2_LA_gDNA_(1)	G	G	T	G	T	C	G	G	T	-	G	T	G	C	T	G	G	C	T	T
HM2_LA_gDNA_(7)	G	G	C	G	A	C	T	T	T	-	C	C	T	C	C	G	T	A	A	G
HN1_gDNA(1)	G	C	T	G	A	C	G	T	C	T	C	C	T	T	C	G	G	C	T	G
HN1_gDNA(5)	G	C	T	G	A	C	G	T	C	T	C	C	T	T	C	G	G	C	T	G
HM2_gDNA(3)	G	C	T	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HM2_LA_gDNA_(2)	T	G	T	A	A	A	T	T	T	A	C	C	T	C	C	G	T	A	A	G
HN1_gDNA(9)	T	G	T	A	A	A	T	T	T	A	C	C	T	C	C	G	T	A	A	G
HM2_gDNA(1)	G	G	C	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Table 14. Record of bases observed at different polymorphic sites in sequence alignment of cloned T6ODM sequences. The numbers in the columns represent the positions of conflict in the sequence. This table enabled grouping together of clones. The green group is an 100% match at the nucleotide level to the published T6ODM cDNA sequence. Different primer pairs were used in the initial PCRs. A '-' at the ends indicates portions of the gene not amplified and sequenced. Interior '-' indicate gaps in the alignment due to insertions/deletions at these points and/or spliced out introns in cDNA sequences.

3.5.3 Six differences between T6ODMa and T6ODMb at the amino acid level

The nucleic acid differences between the characterised T6ODMa (Hagel and Facchini, 2010a) and T6ODMb cloned and sequenced from HM2 gDNA and HN1 cDNA and gDNA are summarised in Table 15.

Nucleotide difference T6ODMa-T6ODMb	Effect	PSSM Difference	SIFT Score
G138A	A36T		
A236G	P68=		
A237G	I69V		
A257G	L75=		
T271C	F80S	-5.8	0.56
G519A	V128=		
A545C	E137A		
T1531C	V188=		
T1602A	I212K		
*G1766T	*A266L		
T1914C	S292=		
G1992A	S318=		
T2130G	I364M		

Table 15. Nucleic acid differences between T6ODMa and T6ODMb and their effect on amino acid sequence. For example G138A refers to position 138 in the multiple sequence alignment. T6ODMa has a G at this position while T6ODMb has an A. This results in a corresponding change in amino acid sequence of the protein-A to T. '=' means the same amino acid is encoded by the corresponding codon in T6ODMb. The PSSM difference and SIFT scores are estimations of the impact of the substitutions on protein function. Mutations are predicted to have a severe effect on protein function if PSSM scores are >10 and the SIFT scores are <0.05. * Published T6ODMa has a T at this position which ensures an S amino acid is encoded rather than an A in T6ODMa from GSK HM2 & HN1.

Most of the differences between the genes at the nucleotide level do not correspond to amino acid differences in the resulting proteins. There are 6 amino acid differences between T6ODMa and T6ODMb-A36T, I69V, F80S, E137A, I212K and I364M. The differences/substitutions in T6ODMb were not predicted to have impacts on protein function. The PARSESNP tool previously available at http://www.proweb.org/parsesnp/ was used to analyse the nucleotide polymorphisms (Taylor and Greene, 2003). Position-Specific Scoring Matrix (PSSM) scores, based on homology to related sequences, give an indication of the importance of certain amino acids. Substitutions can affect the structure of function of the resulting protein. PSSM scores of >10 are deleterious. The Sorting Intolerant from Tolerant (SIFT) programme similarly determines the severity of missense mutations but assigning a score between 0 and 1 to each change from the reference sequence. Here a SIFT score of <0.05 is deleterious. None of the differences in T6ODMb were predicted to be deleterious.

The differences do not occur in areas of the protein deemed important for substrate recognition, 2-oxoglutarate binding and coordinating Fe(II). Therefore, T6ODMb appears to be a functional 2-ODD expressed at least in HM2. This has obvious knock on effects for reverse genetics in that mutations in both T6ODMa and T6ODMb would have to be found to completely block metabolism at thebaine.

3.5.4 At least two copies of T6ODM are closely linked

While sequencing and assembly efforts were ongoing long range PCR with genomic DNA from HM2 was used in an attempt to amplify potential intergenic regions of 2-ODDs. LongAmp primers were designed to amplify out from the respective genes (Figure 30). Each primer combination possible was trialled in PCR with HM2 gDNA. It is reported that PCR products of up to 30 kb can be obtained using the LongAmp PCR kit. Successful PCR would indicate close proximity of a pair of 2-ODDs in the genome. Just one ~7 kb PCR product was obtained using primer pair T6ODM_F and DIOX2 (PODA)_R. However, when the PCR product was purified and sequenced from both ends it became clear that the result represents a link between two copies of T6ODM. It was not possible from the sequence information obtained to differentiate which copies of T6ODM were at either end of the PCR product.

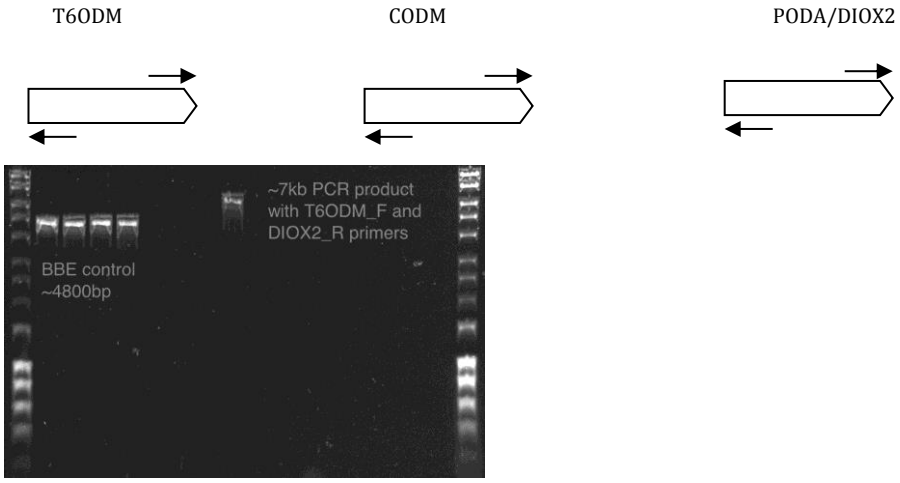


Figure 30. LongAmp experiment to investigate whether there are physical connections between 2-ODD genes.

3.5.5 Section summary

As with CODM, the results indicate there are multiple copies of T6ODM in the genome. However, the nucleic acid and predicted amino acid sequences do differ between copies. Furthermore, long range PCR indicated that at least two copies of T6ODM are closely linked.

3.6 TILLING for induced mutations in T6ODM

3.6.1 Design of appropriate primers to screen the copy of T6ODM with proven activity against thebaine and codeine.

Initial characterisation of T6ODM revealed that there are multiple copies in the genome (Table 14). Likewise, T6ODM-like contigs built from ESTs in our internal poppy EST database contained numerous conflicts between ESTs. Unlike with CODM, where there were just two polymorphic sites between copies, it would not be possible to screen all copies of T6ODM simultaneously. When this was attempted, the numerous polymorphisms created lots of different heteroduplexes after PCR. After treatment with the cleavage enzyme, the large number of bands on the TILLING gels would frustrate efforts to detect genuine mutations (Figure 31). On the other hand only two copies appear to be expressed (Table 11) which means that primer design could theoretically take advantage of the polymorphisms to selectively amplify a specific copy of T6ODM for screening.

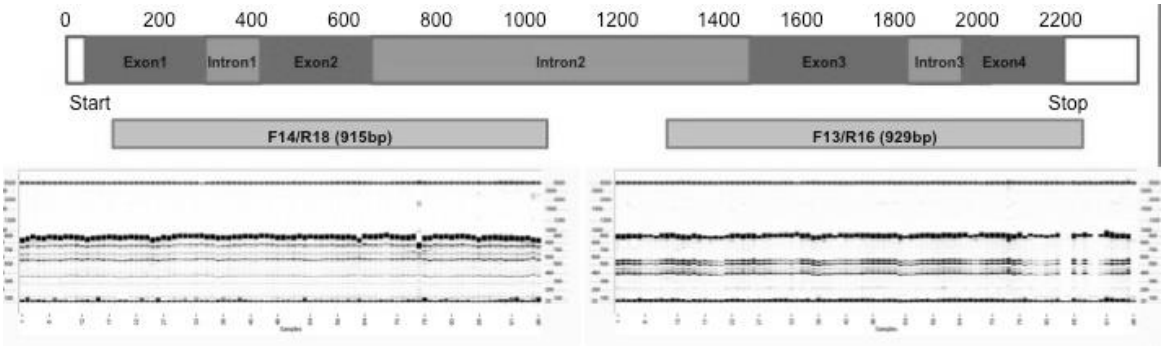


Figure 31. Initial efforts to TILL T6ODM. Differences between copies of T6ODM result in numerous common bands on the TILLING gel.

T6ODMa is an 100% match at the nucleotide level to the previously characterised T6ODM (GenBank accession: GQ500139; Hagel and Facchini, 2010a). Because the function of this gene is known, it was prioritised for targeting by reverse genetics over the others. A

polymorphism at position 708 in the alignment between different copies of T6ODM (Table 14), whereby T6ODMa has a 'G' and T6ODMb and others have either an 'A' or 'C', was exploited to design a primer (R27) for the targeted amplification of T6ODMa only in touchdown PCR. When used in conjunction with primer F14, exons 1 and 2 of T6ODMa would be amplified generating a 626 bp PCR product (Figure 32a). It was important to target exon 2 because it encodes for a number of tryptophans, the codons of which are quite easily changed by EMS to stop codons.

3.6.2 Mutation screen

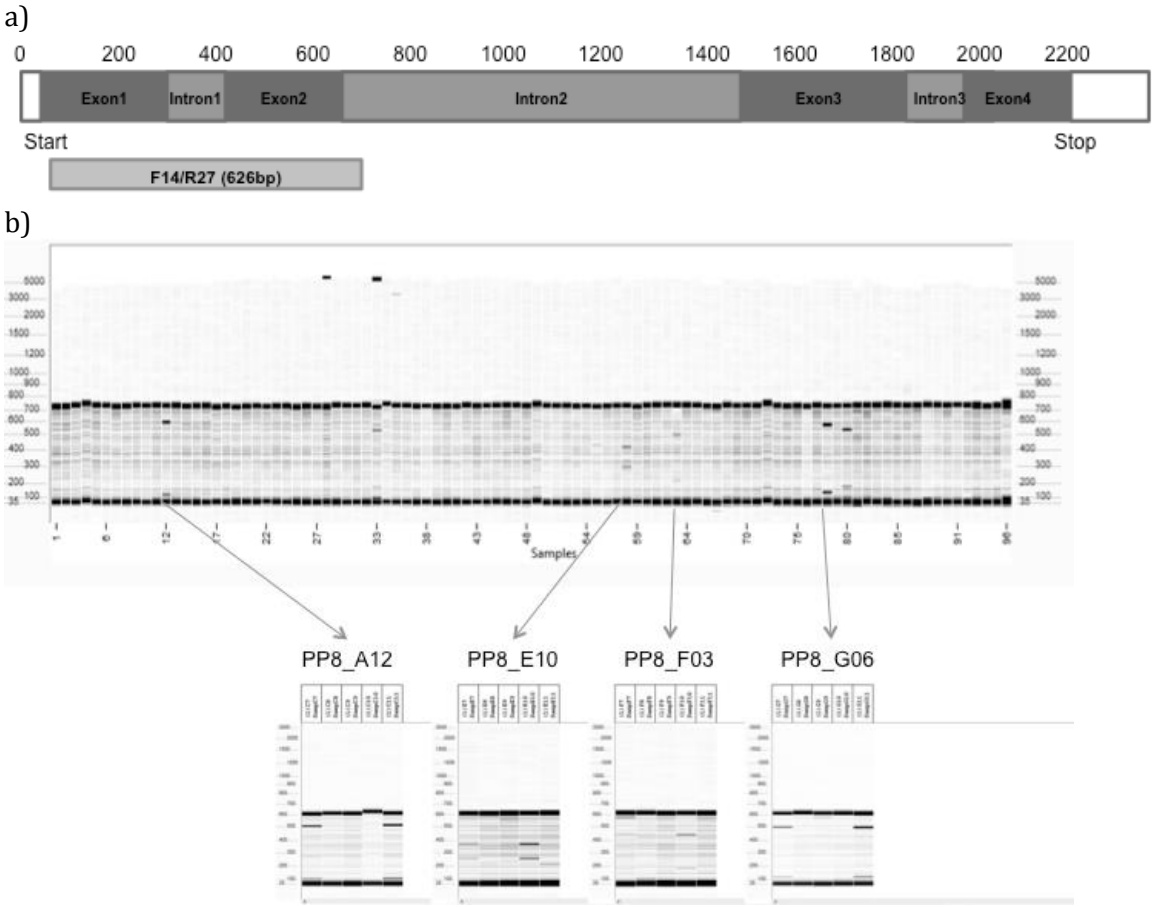


Figure 32. a) Region of T6ODMa screened for mutations using primer pair F14/R27. b) Pool plate 8 T6ODMa screen and the rescreen of some of the mutations found to determine the individual in the pool carrying the T6ODMa mutation.

Figure 32b shows the Fragment Analyser gel for pool plate 8 T6ODMa PCR products. A number of mutations were identified on this pool plate. The PCR product was cut into fragments, the sum of the sizes of which add up to 626 bp. The process was repeated for all twelve pool plates (Appendix C). Any potential mutations found were confirmed in the rescreen e.g. for wells A12, E10, F03 and G06 of pool plate 8 shown here. Unlike with CODM,

where multiple copies were amplified in PCR with the TILLING primers, only T6ODMa was amplified with primer pair F14/R27. For the rescreens then, the PCR needed to be spiked with wild type HM2 DNA to ensure heteroduplex formation potential homozygous mutants. In this way the individual in the pool of four carrying the mutation could be identified.

3.6.2.1 Summary of T6ODM mutations detected in reverse genetic screen

Table 16 provides a summary of the T6ODMa mutations found in the reverse genetic screen of 4146 M2 plants. Appendix C has the full record of the fragments detected in the initial screen and the follow up rescreen as well as the actual TILLING gel images. There were some PCR failures, but these were not repeated as satisfactory results were achieved in the initial screen.

Just nine T6ODMa mutations were recovered. However, importantly two of these were nonsense mutations-Q141* and W145*. Three mutations were silent and one was in an intron leaving three individuals with two different missense mutations. A lower number of mutations was expected because of the shorter T6ODMa amplicon used. Overall the estimated mutation frequency (1 per 288 kb) is similar to that estimated from the recovery of CODM mutations (1 per 395 kb).

Pool plate	Well	Screening		Rescreening		Plant Sd numbers				DNA Source	Sequencing Result	Seed batch number of mutant	Population
		Fragment 1 (bp)	Fragment 2 (bp)	Fragment 1 (bp)	Fragment 2 (bp)								
One	-	-	-	-	-	-	-	-	-	-	-	-	-
Two	-	-	-	-	-	-	-	-	-	-	-	-	-
Three	-	-	-	-	-	-	-	-	-	-	-	-	-
Four	-	-	-	-	-	-	-	-	-	-	-	-	-
Five	-	-	-	-	-	-	-	-	-	-	-	-	-
Six	-	-	-	-	-	-	-	-	-	-	-	-	-
Seven	B11	316	611	161	450	102690	102691	102692	102693	D-10025 E11-H11	Q141*	S-102693	B2
Eight	A12	58	612	103	509	103087	103088	103089	103090	D-10029 A12-D12	H162Y	S-103090	B2
	C9	43	609	147	468	103160	103161	103162	103163	D-10030 A9-D9	A146T	S-103161	B2
	E10	471	584	254	367	103317	103318	103319	103320	D-10031 A10-D10	F110F silent	-	C1
	F3	563	606	181	438	103257	103258	103259	103260	D-10031 E3-H3	G134G silent	-	C1
	G6	107	611	115	490	103397	103398	103399	103402	D-10032 A6-D6	L156L silent	-	C1
	G8	167	609	156	473	103419	103420	103421	103422	D-10032 A8-D8	A146T	S-103421	C1
Nine	-	-	-	-	-	-	-	-	-	-	-	-	-
Ten	D12	504	565	283	345	103991	103992	103997	103999	D-10036 E12-H12	Intron 1	-	B3
Eleven	A12	219	609	151	468	104549	104550	104551	104553	D-10039 A12-D12	W145*	S-104553	B3
Twelve	-	-	-	-	-	-	-	-	-	-	-	-	-

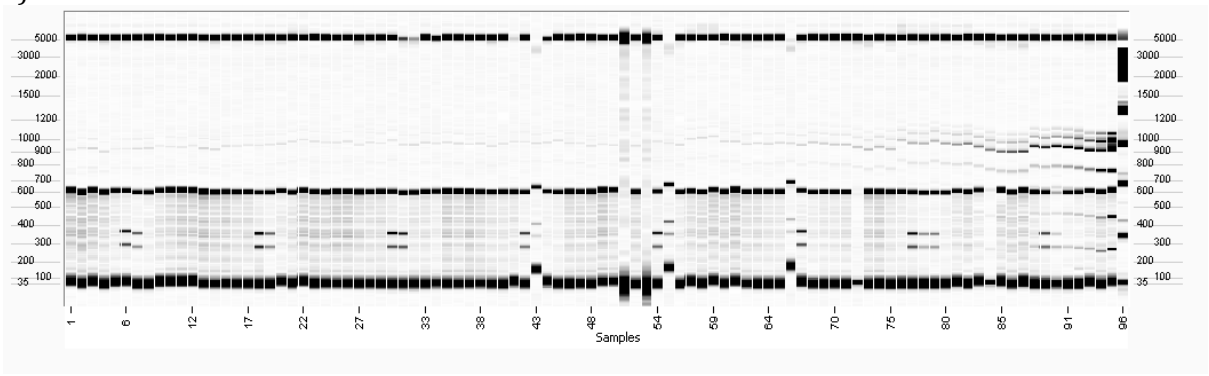
Table 16. Summary of identified T6ODMa mutations in the HM2 EMS mutagenised population. Individuals in the pool of four carrying the mutation are shaded. The shaded mutations were prioritised for further work.

3.6.2.2 A polymorphism in intron 1 of T6ODMa is present in a high thebaine forward screen line and commercial thebaine lines

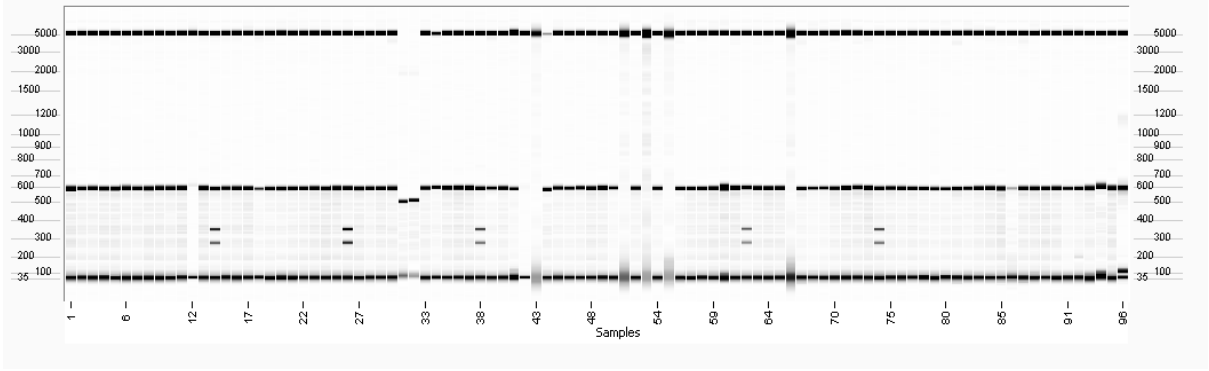
The same EcoTILLING strategy used previously (Figure 19) was employed to check whether there are any differences in the nucleotide sequence of T6ODM in high thebaine forward screen lines and WT HM2 and HN3 from which they were derived. These lines had undergone numerous rounds of selfing in the field in Tasmania. Capsules of these lines accumulate substantially more thebaine and oripavine than HM2 and HN3. The seed batches were sent to York for DNA extraction and analysis. DNA from the resulting seedlings was spiked with equal amounts of wild type DNA prior to PCR. Differences in sequence between wild type and mutant T6ODM would result in heteroduplexes and cleavage.

Figure 33a shows the discovery of a T6ODM polymorphism in all lanes where successful amplification using template DNA of plants derived from S-114249 (MAS 2150 B C) occurred. This is a high thebaine forward screen line derived from HM2 mutagenesis. The same polymorphism was identified in 5 of the 8 HT6 plants tested in Figure 33b. No polymorphisms were detected in the other two thebaine forward screen lines derived from HM2 mutagenesis or the four noscapine-free thebaine/oripavine only forward screen lines derived from HN3 mutagenesis. Just 626 bp of one copy of T6ODM was screened for mutations (Figure 32) so these lines and the other two other high thebaine lines from HM2 mutagenesis may have as yet undetected T6ODM polymorphisms (Appendix C).

a) D-10112 T6ODM EcoTILL



b) D-10836 T6ODM EcoTILL



c) Summary of lines screened and polymorphisms identified

CNAP seed batch id	GSK seed batch id	Original M2 seed batch (Population)	GSK capsule data Grown in field in Tasmania (%DW)				CNAP capsule data Grown in glasshouse in York (%DW)					Polymorphisms identified with Ecotilling	Gene affected
			M	O	C	T	M	O	C	T	CD dimer		
S-122070	MAS 2787 D	S-109141 (B6)	43.4	4.3	12.8	39.5	39.7	6.6	7.5	43.3	1.2	-	-
S-114249	MAS 2150 B C	S-107166 (B4)	15.5	19.4	6.5	58.5	19.1	16.9	2.4	57.7	2.9	**F271L Intron**	CODMb T6ODMa
S-114250	MAS 1956-4	S-106367 (B4)	16.1	1.7	4.0	78.2	47.9	3.8	5.0	39.3	2.9	-	-
S-125001	N/A	-	-	-	-	-	-	18.0	-	79.0	-	-	-
S-125002	N/A	-	-	-	-	-	-	24.0	-	72.0	-	-	-
S-125003	N/A	-	-	-	-	-	-	25.0	-	72.0	-	-	-
S-125004	N/A	-	-	-	-	-	-	27.0	-	70.0	-	-	-
HM2	WT	N/A	-	-	-	-	85.8	0.3	5.2	3.4	4.3	N/A	N/A

Figure 33. Identification of a polymorphism in intron 1 of T6ODMa in a) a high thebaine forward screen line and b) a GSK commercial thebaine variety, HT6 (5 of 8 plants tested-lanes 14, 26, 38, 62 and 74; absent in lanes 2, 50 and 86).

Sequencing T6ODMa of plants with the polymorphism revealed it is not in the coding sequence, but rather resides in intron 1 of T6ODM. It is unlikely that it was EMS induced as the polymorphism results in a change from T (HM2) to A (S-114249) at this position (Greene et al., 2003). Furthermore, the polymorphism (Tint) is present in all the commercial thebaine varieties from which DNA was available (HT1, HT4, HT5, HT6). It is probable that this allele was transferred to ancestors of S-114249 by a commercial T line in the field and may not

necessarily indicate a seed mix up at some point. At the time the polymorphism was discovered all HT1, HT4, HT5 and S-114249 individuals genotyped (3 of each) were homozygous for the polymorphism indicating that it is fixed. Subsequent work revealed that the polymorphism is not fixed both in HT5 and HT6 (Figure 34).

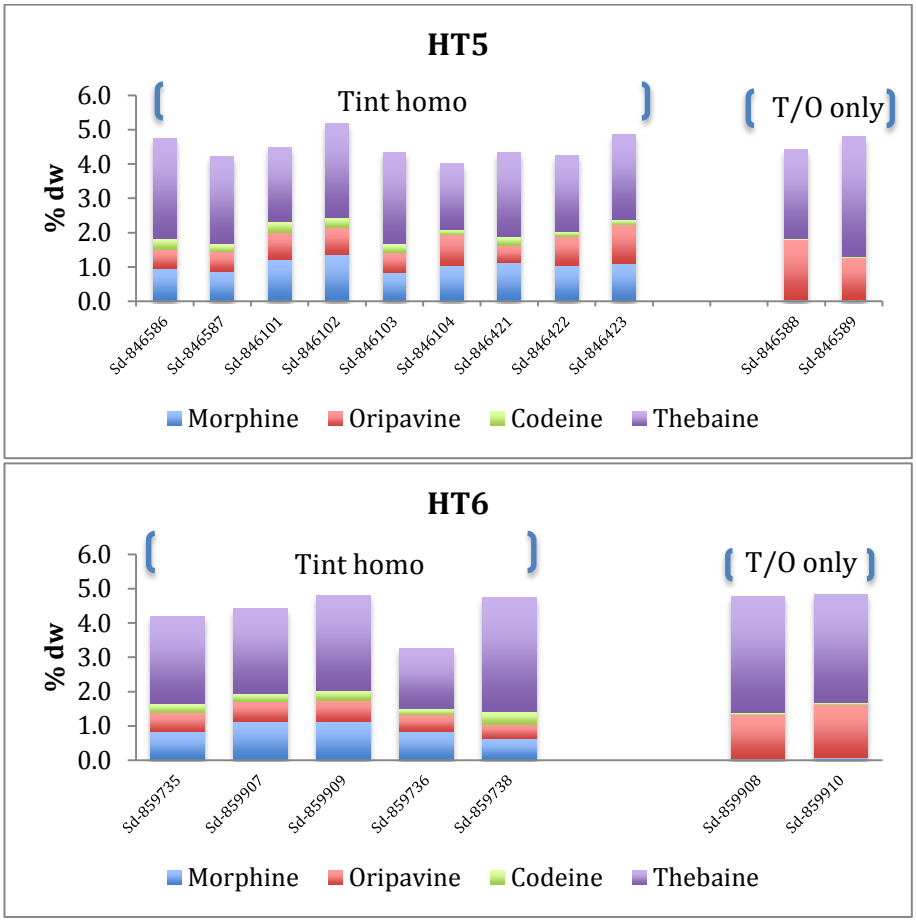


Figure 34. Phenotypes of HT5 and HT6 commercial thebaine lines grown as controls during Tasmanian field trials.

There are two distinct phenotypes in these lines-a morphine-thebaine-oripavine phenotype associated with the Tint polymorphism and a morphine-free thebaine-oripavine only phenotype with no detectable T6ODM polymorphisms. The latter has a complete block of metabolism at thebaine but the molecular basis of this phenotype is unknown. As knockouts of T6ODMa (W145* and Q141*) do not result in such complete blocks, presumably a regulator of T6ODM or a transporter of thebaine is affected. The Tint polymorphism could be used to assess future lines for homogeneity.

Within an intron, a donor site (5' end), a branch site (near the 3' end) and an acceptor site (3' end) are required for splicing. The donor and acceptor site typically have GU and AG

sequences, respectively. The branchpoint includes an adenine residue, which performs a nucleophilic attack on the donor site during splicing forming the lariat intermediate. In the second step the 3'OH of the released 5' exon then performs a nucleophilic attack at the acceptor site, thus joining the exons and releasing the intron lariat. Netgene2 was used to predict whether the polymorphism would have an effect on splicing (Hebsgaard et al., 1996).

```
Branch points, direct strand ----------------------------
acceptor     branch
pos 5'->3'  pos 5'->3'        strand    score     5'
3'
   402          374             +       -2.25       TTAGCTAGTGAATCTTACTT
```

Figure 35. Netgene2 analysis of intron 1 of T6ODMa. The T marked in bold (in commercial morphine and noscapine lines) is the base which is changed to A in commercial thebaine lines and the high T forward screen line S-114249.

The analysis predicted the branch point in HM2 T6ODM to be the A in italics. The T in bold, which is changed to A in the GSK commercial thebaine lines and S-114249, is situated five bases upstream. It is possible that this substitution disrupts splicing of intron 1 during pre mRNA processing. Perhaps the wrong A is used in the nucleophilic attack on the donor site. Alternatively the change in sequence may affect recruitment of elements of the spliceosome, the large RNA-protein complex that catalyses splicing. Splicing factor 1 binds to the intron branch point sequence prior to displacement by the U2 small nuclear ribonucleoprotein (snRNP).

A similar T to A mutation in the branch point consensus of the lecithin: cholesterol acyltransferase (LCAT) gene in humans is the cause of fish eye disease, a rare disorder characterised by low high-density lipoprotein (HDL) and dense corneal opacities (Li et al., 1998). LCAT is involved in the metabolism of HDL. The authors claimed that the mutation affected the interaction of U2 snRNA to the branch point. Binding of U2 snRNA enables bulging of the A for splicing to begin. The result of this mutation was a disruption of splicing and an RT-PCR demonstration of intron 4 retention. The LCAT protein was not detected. It may be that secretion was affected by the altered protein or the translation product was rapidly degraded.

cDNA of high thebaine lines were tested for retention of intron 1 of T6ODMa but no evidence could be found. The intron seems to be spliced out, unlike in the example above. It may be that the polymorphism affects mRNA stability in some way or that the polymorphism is merely associated with a high thebaine phenotype. For this reason, the polymorphism was tracked in crosses (Chapter 5). It was later proved to be an effective tool in predicting thebaine content among siblings in which the polymorphism was segregating (Appendix I-X051, S-186678). There are clear differences in thebaine yields of plants lacking the polymorphism (low thebaine), heterozygotes and homozygotes (highest thebaine).

3.6.2.3 Selection of best candidates from TILLING screen for producing heritable phenotypic changes

Q141* and W145* mutations should result in a truncated protein unable to catalyse O-demethylation of thebaine and oripavine. Further to the two nonsense mutations, two A146T mutants were also identified. These had the potential to be interesting as the substitution lies in an area of T6ODMa thought to be important for substrate recognition (Hagel and Facchini, 2010a; Runguphan et al., 2012). T6ODM and PODA have an A at this position, while CODM has a T. PSSM and SNAP analysis did not predict a deleterious effect for this missense mutation, but because of its potential role in substrate recognition it was brought forward together with the two nonsense mutations for further work.

3.6.3 Section summary

By exploiting a polymorphism between T6ODMa (the identified 2-ODD with demonstrated thebaine demethylation activity) and other copies, primers were designed to specifically amplify and use this gene for TILLING. Nonsense mutations were identified in the TILLING populations. Furthermore, a promising intron polymorphism was identified in lines associated with a high thebaine phenotype. These were prioritised for further work.

3.7 Recovery of mutations in M2 siblings and assessment for phenotypic changes

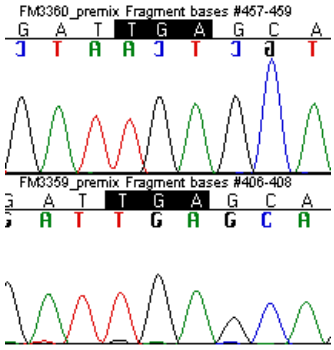
3.7.1 Recovery of mutations in M2 siblings

Recovery of T6ODMa mutations in M2 siblings was achieved by both TILLING and AS-PCR (Appendix C). Both approaches provided similar results in that the recovery rate for each mutation was low i.e. just two Q141* mutants were recovered among 38 siblings and just one A146T mutant was recovered among 15 siblings (Table 17).

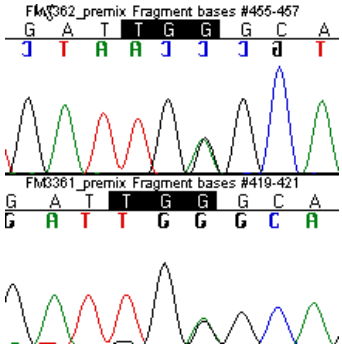
M2 seed batch id	T6ODMa polymorphism	Seedling_id	Zygosity	M3 seed batch id sown
S-104553	W145*	Sd-833305	**Homozygous**	S-123361
S-104553	W145*	Sd-833306	Heterozygous	-
S-104553	W145*	Sd-833308	Heterozygous	-
S-104553	W145*	Sd-833316	Heterozygous	-
S-102693	Q141*	Sd-833291	Heterozygous	S-123349
S-102693	Q141*	Sd-833298	Heterozygous	-
S-103421	A146T	Sd-833331	**Homozygous**	S-123374*
S-103161	A146T	Sd-833350	Heterozygous	S-124115

Table 17. Genotypes of M2 seedlings containing T6ODMa mutations. * indicates the M3 seed did not germinate.

Zygosity was determined by sequencing (Figure 36). M3 seed was collected and sown.



Sd-833305 example of W145* homozygote



Sd-833306 example of W145* heterozygote

Figure 36. Sequencing chromatograms showing the difference between a W145* homozygote and heterozygote. The third base in the codon for tryptophan is changed to A by the mutation, causing a premature stop codon (TGA). In the heterozygote there are signals for equal magnitude for the wild type (G) and mutant (A) allele. The homozygote has a complete change from G to A.

3.7.2 Analysis of M2 capsules-W145 mutant shows 5 fold increase in thebaine over HM2*

As mentioned above the W145* and Q141* T6ODMa mutations were segregating among their respective siblings. One W145* homozygote was recovered while two Q141* heterozygotes were identified and prioritised for selfing to produce M3 seed. M2 capsules were phenotyped (Figure 37). Although the numbers in each class are low, the results suggest that thebaine accumulation in the capsules is increased in T6ODMa W145* mutants. There is certainly a marked difference in thebaine content of the W145* homozygote (11 µg/mg dw (no SD as it was a single plant); 20% of total alkaloids) with that of siblings lacking the mutation (3.2±3.8 µg/mg dw; 7% of total alkaloids) and HM2 (1.4±1.5 µg/mg dw; 3% of total alkaloids). The three heterozygotes have an intermediate level of thebaine with a mean of 4.1±2.9 µg/mg dw (13% of total alkaloids). There is also a slight increase in the proportion of oripavine compared to HM2 (Appendix C).

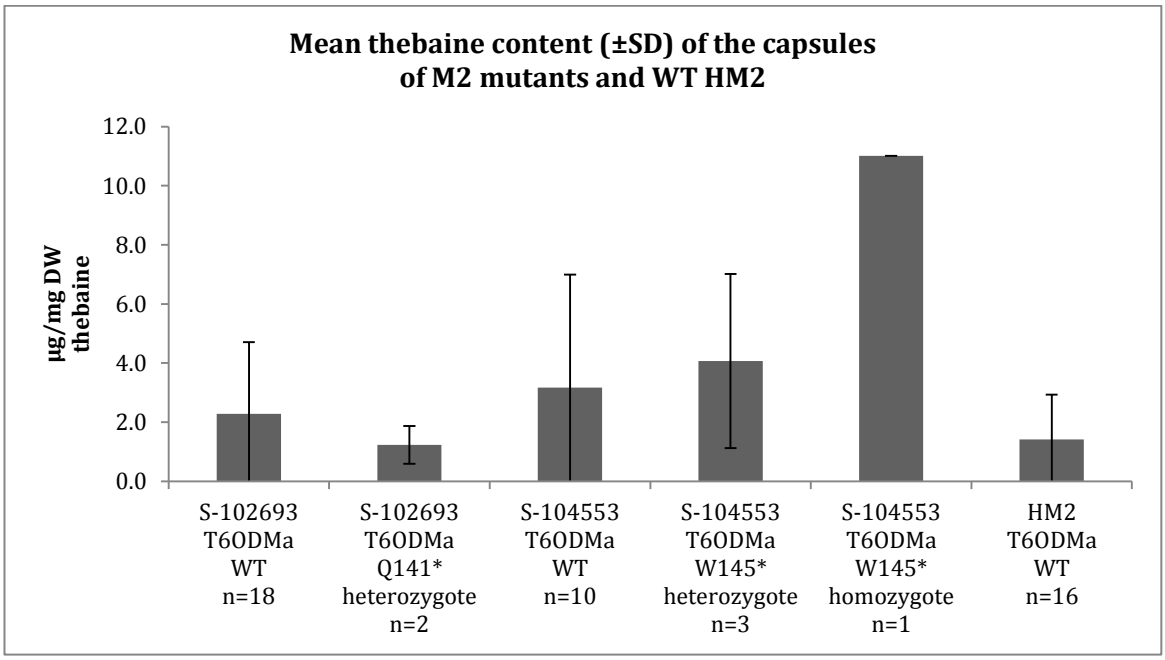


Figure 37. Thebaine content in the capsules of T6ODMa mutants and siblings with WT T6ODMa.

3.7.3 Analysis of M3 capsules-Increased thebaine in W145 homozygotes*

Seed collected from the homozygous W145* M2 was brought forward to M3. All the resulting plants were expected to be homozygous. In contrast, the Q141* and A146T

mutations were segregating in M3. Genotyping was completed using AS-PCR with visualisation of AS-PCR products on the Fragment Analyser.

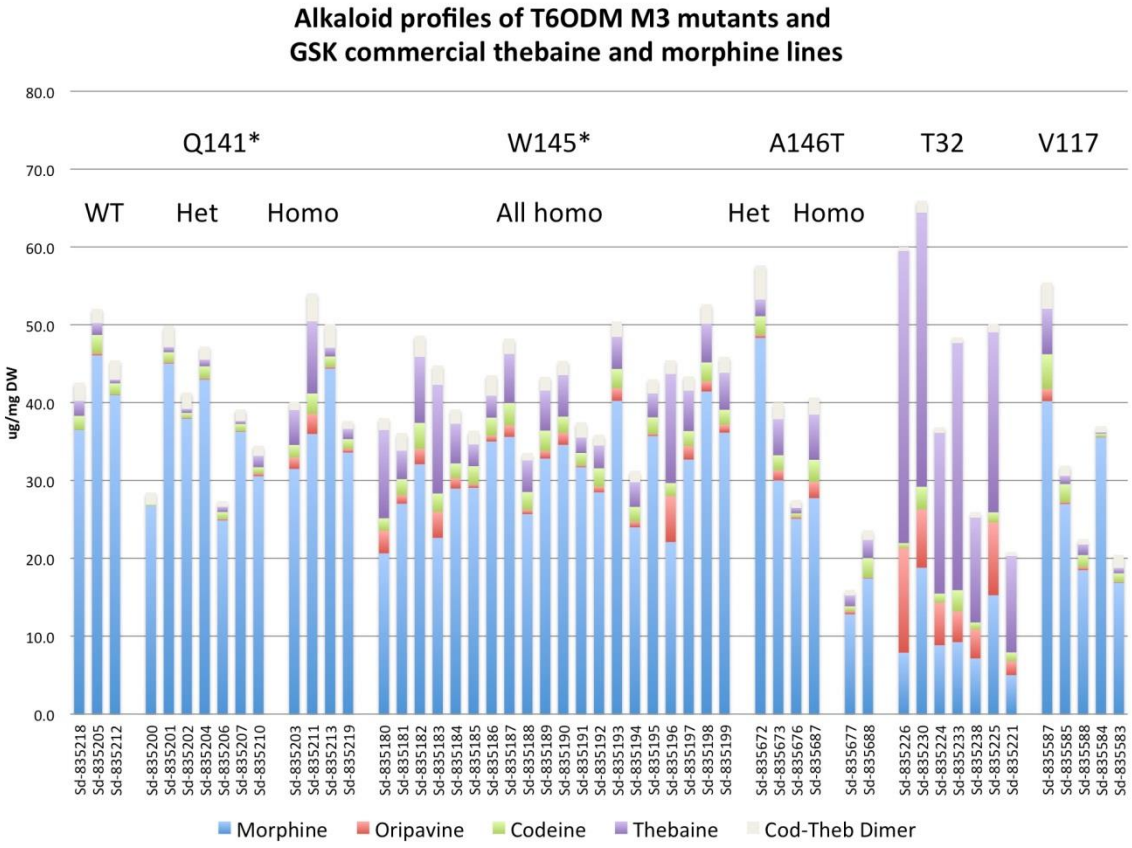


Figure 38. T6ODM M3 plants homozygous for T6ODM knockout mutations have higher thebaine and oripavine contents in their capsules than heterozygous and WT siblings and a commercial morphine line (HM6). However the relative amounts of these two alkaloids were lower than that in a current commercial thebaine line (HT5). The Q141* mutation was segregating among M3s. All W145* M3s were homozygous and displayed a partial block of metabolism at thebaine and oripavine. These are alkaloid profiles of single plants and not the mean for a line. Therefore error bars were not included.

The alkaloid content of the capsules of the various mutant classes is displayed in Figure 38. None of the mutants had thebaine and oripavine levels as high as the commercial thebaine line, HT5 (the molecular basis of which is unclear). However, there was substantially more thebaine in the capsules of two of the Q141* homozygotes (mean of all four was 4.0±3.8 µg/mg dw; 8% of total alkaloids) and all of the W145* homozygotes (5.7±3.6 µg/mg dw; 13% of total alkaloids) than commercial morphine lines HM2 (previously analysed in 3.8.2) and HM6 (1.8±2.3 µg/mg dw; 5% of total alkaloids). The two A146T homozygotes identified in M3 had low total alkaloid levels and a small proportion of thebaine (1.9±0.6 µg/mg dw; 9% of total alkaloids). It was decided to only continue with the knockout mutations as these displayed the intended changes in phenotype. Similar to CODM

nonsense mutants, knockouts of T6ODMa resulted in plants where the predominant alkaloid was still morphine. The mutations led to partial blocks at thebaine and oripavine. Additional copies of T6ODM may continue to O-demethylate thebaine and oripavine at levels similar to wild type HM2 or its successor HM6.

3.7.4 Section summary

The T6ODMa nonsense mutations resulted in partial blocks of metabolism at thebaine and oripavine among homozygous M3 individuals.

3.8 Discussion

EMS mutagenesis is a well established approach to inducing genetic variation and it causes a high density of irreversible point mutations (Henikoff and Comai, 2003). Mutagenesis of seed from the morphine line HM2 provided large mutant populations for analysis. As well as leading to lines with interesting metabolite profiles identified by forward screens, a proportion of the M2 lines were used in reverse genetic screens. Based on the recovery of T6ODM and CODM mutations, the estimated mutation frequencies were 1 per 288 kb and 1 per 395 kb, respectively. This compares favourably with estimated mutation frequencies based on results with other genes targeted previously (data not shown). With a genome size of 3724 Mbp, each M2 line could therefore contain in the region of 10,000 EMS induced mutations. Any line with a desired mutation would need to undergo numerous rounds of backcrosses to remove the undesired mutations. Mutant plants often had stunted growth, altered leaf shape, small capsules and flowered early. Some M2s even had albino sections, presumably resulting from mutations causing defects in chlorophyll production. Mutagenesis with EMS, then, provided a large population of plants with heritable changes in nucleotide sequence which could be detected in target genes by TILLING.

To our knowledge, a reverse genetics approach has yet to be reported with mutagenised populations of opium poppy. Mutation detection was carried out using Advanced Analytical's Fragment Analyser. Although this system has been sold by Advanced Analytical for a number of years, there is yet to be an example of its utilisation in TILLING in the literature.

CODMa and T6ODMa represent exact matches to the respective 2-ODD genes shown to be involved in O-demethylation steps in the final stages of morphine biosynthesis in 2010 (Hagel and Facchini, 2010a). Additional copies of both of these genes have been detected in gDNA of GSK commercial varieties. Copies of CODM have identical amino acid sequences and have just two polymorphic sites at the nucleotide level. In contrast there seems to be many more copies of T6ODM in the genome. Copies of T6ODM with minor amino acid differences to T6ODMa may have different activities or altered substrate preferences. Recombinant T6ODMa exhibited only trace to relatively minor activity with protopine alkaloids such as cryptopine, protopine and allocryptopine (Farrow and Facchini, 2013). However, these alkaloids accumulated in T6ODM VIGS silenced plants against expectations. VIGS would have silenced all copies of T6ODM including those with potential preferences for protopine alkaloids.

Sequencing CODM-containing BACs confirmed the presence of multiple copies of CODM and suggested that at least some of the copies are physically linked. Other genes involved in morphine biosynthesis were not detected on the sequence contigs analysed and so genes involved in morphine biosynthesis are not likely to be clustered in a manner similar to that with noscapine (Winzer et al., 2012). Nevertheless, evidence of duplication of genes involved in the metabolism of an important plant secondary metabolite was not unexpected. Gene duplications drive the recruitment of genes for secondary metabolism. In contrast to primary metabolism in which the most stable and functional enzymes are maintained, enzymes of secondary metabolism may have emerged through gene duplication and subsequent mutation accumulation altering substrate specificities or reaction activation barriers (Weng et al., 2012). Gene copies are modified over evolutionary time to create new genes in the family with different specificities and expression patterns adapted to the needs of the plant (Ober, 2005). At least one copy may be allowed to accumulate mutations leading to greater mechanistic elasticity and neofunctionalisation before inactivating mutations are introduced. Duplicated genes are often in tandem repeats, forming clusters within the plant genome. Tandem repeats of CODM were not observed but the results did show at least two copies of

CODMc are within 150kb of each other. The inability to combine mutations in CODMa/c and CODMb through crosses further suggests close physical linkage of the CODM copies. The precise arrangement and organisation is subject of future work. Chromosome walking efforts may close the gaps and confirm the architecture of the region of the genome with CODM. It would be interesting to see whether there are any other genes involved in the biosynthesis or regulation of morphine biosynthesis.

Various classes of transposable elements were present in the BACs sequenced. These complicate assembly of genomic sequence information, as sequence reads map with equal probability to multiple positions in the reference (Morrell et al., 2012). It is therefore understandable that most of the crop genomes sequenced to date have all been relatively small-pepper, tobacco and bread wheat being rare examples of genomes sequenced larger than opium poppy (3724 Mbp) and the average angiosperm genome (6000 Mbp). Transposable element induced null mutations are the most common cause of induced phenotypic changes and have been selected for many times in plant domestication. Insertions into repressors or enhancers can eliminate positive or negative regulatory functions e.g. insertion of *Mutator* elements into intron 1 of the *knotted1* gene leads to ectopic expression of this gene in maize leaves (Greene et al., 1994). Transposition into the 5' end of plant genes could also potentially introduce new regulatory information and alter gene expression. Transposable element activity is also thought to be involved in gene duplication and subfunctionalisation e.g. rearrangements at the *r1* locus in maize caused by aberrant transposition of a CACTA element caused changes in tissue specificity of duplicated gene copies (Walker et al., 1995). Transposable elements are highly sensitive to various abiotic and biotic stresses and so insertion upstream of host genes can confer stress responsiveness e.g. insertion of *mPing* elements conferred salt or cold inducibility in the EG4 cultivar of rice (Naito et al., 2009). Aberrant transposition of Class II elements can lead to deletions, inversions and translocations (Lisch, 2013). Movement of genes to a new chromosomal context can alter their regulation. Capture of gene fragments during movement can also impact on the evolution of new genes and so as well as impacting on genetic variation,

transposable elements have important roles in plant evolution. Of course at the moment there is limited evidence for gene duplication and repetitive elements surrounding these genes and it is not clear if the transposable elements are enriched in this area compared to other areas of the poppy genome. However, it is intriguing to hypothesise that transposable elements could possibly play a role in duplication events and in the regulation of the various copies of T6ODM and CODM.

Using RNA-seq over 50% of duplicated genes in soybean (which has undergone two separate polyploidy events 13 and 59mya) were shown to be differentially expressed and thus underwent expression subfunctionalisation (Roulin et al., 2013). Only a small fraction of duplicated genes were nonfunctionalised or neofunctionalised and overall the results suggested that polyploidy in soybean had impacts mainly at the regulatory level. At the beginning of this project only the publicly deposited sequences were available as targets for TILLING (GenBank accessions Q500141; GQ500139). During the validation of these targets a picture of multiple copies began to emerge, the completion of which goes beyond the scope of this thesis. Even though there are multiple copies of CODM and T6ODM in the genome, it is not clear which copies are most highly expressed and active in morphine biosynthesis. Also of the seven T6ODM copies sequenced there is only evidence of expression of two of these. Therefore, it was reasonable to proceed with targeting of these genes for reverse genetics with the focus being the characterised 2-ODDs, CODMa and T6ODMa (Hagel and Facchini, 2010a). Disruptions in CODM would increase yields of codeine, while disruption in T6ODM activities would increase yields of its substrates thebaine and oripavine. Even a partial block of metabolism to morphine resulting from a mutation in one of the target genes could be useful in efforts to breed new varieties of poppy with altered metabolite profiles.

Nonsense mutations were detected for both CODM and T6ODM in the reverse genetic screens which are highly likely to render the resulting proteins non-functional. These resulted in increases in codeine and thebaine/oripavine, respectively, as expected based on the substrate preferences of the enzymes (Figure 7). Despite identifying putative knockout mutations, complete blocks of the metabolism at the respective biosynthetic steps were not

observed and the increase in codeine and thebaine/oripavine, respectively, were relatively modest. This can be explained by functional redundancy conferred by the additional non-mutated copies of the gene. Virus induced gene silencing had resulted in much larger increases in the relative proportion of these compounds (Hagel and Facchini, 2010a). However, it is probable that the silencing construct used would have silenced more than one copy of the respective genes as it is now clear that there are a number of copies of each of these 2-ODD genes in the genome. More substantial blocks of metabolism at codeine or thebaine would require a plant with all copies of the relevant gene disrupted. This could be achieved without the use of genetic modification by crosses between individual mutants or by successive mutagenesis steps.

Two nonsense mutations were found in CODM among the TILLING population. The crossing strategy failed when trying to produce plants homozygous for both the CODMa/c W261* and CODMb Q254* mutations. This suggested close linkage of the copies in the genome and much larger F2 populations would be required to identify recombinants. Given that two copies are contained on the same BAC (within about 150 kb) this approach could prove infeasible even if mutations in all copies known so far could be identified. As an alternative to crosses, seed from the CODMb Q254* mutant line was bulked up for further EMS mutagenesis in an effort to introduce mutations into additionaly copies of CODM. M2 seedlings from this new mutant population would again be assessed in forward (Tasmanian field trials, section 5.6.3) and reverse genetic screens (to be completed).

EMS induced CODM E193K and R158K mutations were detected in high codeine forward screen lines taken forward from the initial mutagenesis. Genealogy records indicate that codeine content may have increased when these induced mutations were brought to the homozygous state by selection. Use of these lines in crosses and comparison between homozygotes, heterozygotes and wild types will clarify whether this is the case. It is not clear if either, or both, mutation(s) contribute to the phenotype or if they are merely linked to a causative mutation. The E193K mutation may reside in CODMa (Hagel and Facchini, 2010a) and as such may represent the most expressed or active copy of CODM. The W261* and

E301K mutations identified in the reverse genetic screen may be in less active copies of CODM. Alternatively, there may be no difference between the copies and both the E193K and R158K mutations (or an as yet unidentified linked mutation) have the effect of reducing or altering CODM activity in the high codeine forward screen lines.

Two nonsense mutations were also found in T6ODMa. For T6ODM so far only T6ODMa was screened due to its proven role in thebaine and oripavine O-demethylation and the inability to design primers to specifically amplify T6ODMb. Again these resulted in partial blocks of metabolism at thebaine indicating that there a multiple active copies of T6ODM in the genome that can demethylate thebaine and oripavine in plants homozygous for nonsense mutations in T6ODMa. Evidence for the presence of a second expressed copy of T6ODM was found in an in house EST library. A complete block at thebaine, like that observed with VIGS (Hagel and Facchini, 2010a; Wijekoon and Facchini, 2012), would require disruption of all active copies of T6ODM. An alternative strategy could be to use W145* material (homozygous seed has been bulked up) in another round of EMS mutagenesis in an effort to disrupt T6ODMb. It is also not clear what the molecular basis of the high thebaine phenotype in HT5 is. Introduction of the nonsense mutations into already high thebaine lines may further impact on flux to morphine and increase yields of thebaine and oripavine.

Because knockouts of T6ODMa provide only modest increases in thebaine, the Tint polymorphism detected in this gene in various high thebaine lines cannot in itself be the reason for the phenotype. Nevertheless the potential for the Tint polymorphism to affect splicing meant that this marker was included in those chosen to be followed in crosses set up between different chemotypes. Even if the polymorphism is not causative of the high thebaine phenotypes in commercial thebaine lines there is the potential that it is linked to a polymorphism that is.

It is worth re-emphasising that only a small portion of T6ODMa had been screened for polymorphisms. With more genomic sequence information i.e. upstream and downstream sequence that may in future be gleaned from sequencing T6ODM-containing BACs,

differences between copies of T6ODM could be exploited to design primers for TILLING different portions of the gene.

For future work, identifying and targeting of regulators of 2-ODD genes or transporters of metabolites may be better targets for reverse genetics. In particular, the mechanism by which thebaine is transported from the sieve elements to the laticifers should be studied further. If a transporter is involved then this would be an ideal target for reverse genetics as a disruption could cause a more extensive block at thebaine.

Chapter 4- Plant-wide impacts of selected alleles

4.1 Introduction

To assess the impact of CODM and T6ODM mutations on the accumulation of alkaloids in the plant an experiment was set up whereby plants from five mutant lines and wild type HM2 were sampled for alkaloids at different stages during their development. Also, a section of stem (approximately 1 inch in length) was taken immediately below the flower/capsule and used for RNA extraction. The latex that exuded was used for metabolite analysis. Likewise, the roots of the same plant were used for both RNA extraction and metabolite analysis. Because sampling involved sacrificing a plant, results for different growth stages are for siblings rather than a timecourse of a single plant.

Stem and root material was collected from early seedlings (es), plants at the hook stage (bud turned upside down, stem and bud take on a characteristic 'hook' shape; some examples in Figure 1) prior to flowering (hk), flowering plants (flw) and maturing plants 10-14 days post flowering after petal fall (mat). Latex was collected from hk, flw and mat plants only as it was not present in the young seedlings.

4.2 Phenotypes of the lines used in the experiment

Five mutant lines were included in the experiment along with wild type morphine variety, HM2, from which the mutants were originally derived (Table 18). Three lines (W261*, Q254* and W145*) were chosen as they contain knockout mutations for CODM or T6ODM. Previous experiments have indicated that the W261* and Q254* CODM mutations result in modest increases in codeine content of mature capsules (Figure 26). Likewise the W145* T6ODM mutation results in an increase in thebaine and oripavine content of mature capsules (Figure 38). The other two lines used in the experiment, E193K/R158K and HighT, came out of forward screens conducted by GSK in Tasmania (high codeine and high thebaine phenotypes, respectively). They have been repeatedly selfed over a number of years and their high codeine and high thebaine phenotypes are now stable. TILLING of DNA from these lines

identified CODM (Figure 20) and T6ODM (Figure 33) polymorphisms, which may cause or be linked to causes of their phenotypes.

4.2.1 Evaluation of the alkaloid content of capsules of siblings grown to maturity

Figure 39 shows the breakdown of the alkaloid content in capsules of the lines used in the study. Eight mature dried capsules were ground and analysed for each line. The results broadly agree with previous results obtained from these lines when plants were grown on soil, with the exception that codeine content of HM2 was higher than that seen before. Higher codeine content may possibly indicate a stressed plant. The reason the plants were grown on a 50:50 mix of silica sand and Terragreen as opposed to soil was to ensure that roots could be sampled without the damage that removing soil would entail. Washing the root system under a running tap was sufficient to remove the silica sand quickly and ensure the recovered root was flash frozen immediately before alkaloid and RNA extractions. The highest codeine was seen with the high codeine forward screen line with CODM mutant E193K/R158K polymorphisms. The W145* mutants displayed a partial block at T6ODM with levels of thebaine and oripavine are higher than in HM2. The HighT fwd line by contrast has a more substantial block at T6ODM whereby thebaine and oripavine are much more abundant at the expense of morphine.

Line (Population of origin)	M2	M3	M4	M5	M6	M7	M8	M9	Known mutations (homozygous in seed sown on 23DEC14)
HM2									Wild type CODM & T6ODM
E193K/ R158K (C1 pop)	S-103234 Selfed MUTS 437 G→ (Tas 07/08)	Selfed RMUUS 314 E→ (Tas 08/09)	Selfed MUVS 107 A→ (Tas 09/10)	Selfed UKMUVS 107 A→ (UK 2010)	Selfed MUWTS 1926 C→ (Tas 10/11)	Selfed MAS 2235 D→ (Tas 11/12)	S-114247 Sown 21FEB13 Selfed Sd-829638→	**S-122289**	CODMa/c E193K CODMb R158K
W261* (B2 pop)	S-102574 Sown 01AUG12 Selfed Sd-828053→	S-114037 Sown 13DEC12 Selfed Sd-829164→	**S-122152**						CODMa/c W261*
Q254* (B3 pop)	S-104530 Sown 01AUG12 Selfed Sd-828235→	**S-114052**							CODMb Q254*
W145* (B3 pop)	S-104553 Sown 25JUN13 Selfed Sd-833305→	S-123361 Sown 21NOV13 Selfed Sd-835199→	**S-124781**						T6ODMa W145*
HighT fwd (B4 pop)	S-107166 Selfed MUA 2642 A→ (Tas 09/10)	Selfed MUB 1013 B→ (Tas 10/11)	Selfed MAS 2150 B C→ (Tas 11/12)	S-114249 Sown 21FEB13 Selfed Sd-829680→	**S-114139**				CODMb F271L T6ODMa intron polymorphism

Table 18. Origins of seed batches (marked in bold) used to produce plants for the experiment to analyse the impact of CODM and T6ODM mutations in mutant lines. All five mutant lines came out of EMS mutagenesis of a GSK commercial morphine variety, HM2. The population ids refer to different EMS exposures and treatment dates of HM2 seed (Table 1)."S" numbers refer to York assigned seed batch ids with the resulting plants ("Sd" numbers) grown in York. Plants grown in Tasmania are indicated (Tas & growing season) along with the plot number and plant id. The W261*, Q254* and W145* mutations were identified in reverse genetic screens in York. These were selfed until homozygous plants were identified in M2 or M3. The high codeine (E193K/R158K) and high thebaine (HighT) lines came out of a forward screen in Tasmania. In 2013, seed was shipped to York where CODM and T6ODM mutations were identified (in plants derived from S-114247 and S-114249) which may contribute to their phenotypes.

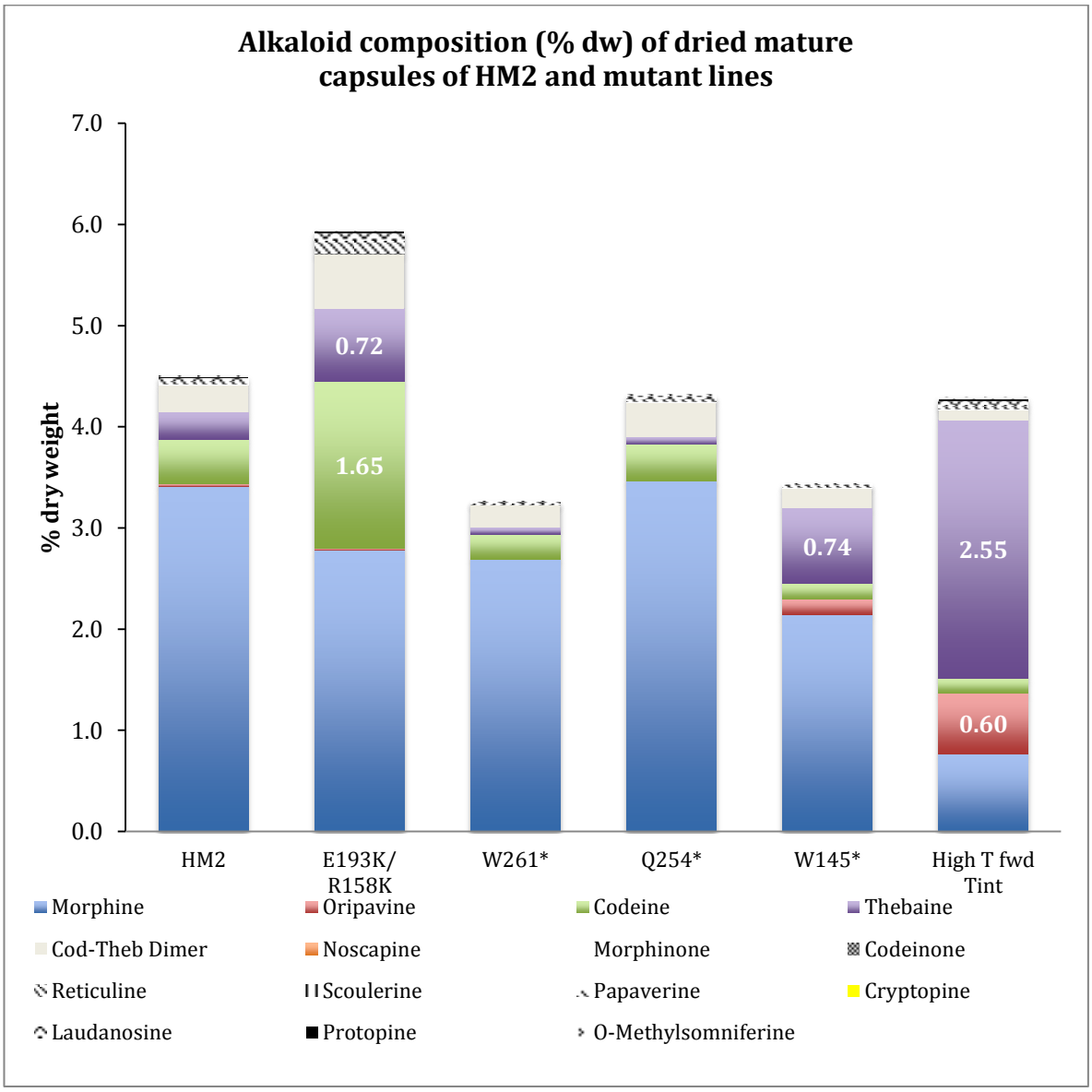


Figure 39. Mean amount (% dw) of each of the 15 most abundant alkaloids in the dried capsules. The figures are an average of 8 biological replicates with each having 3 technical replicates. These are siblings of the plants sacrificed and sampled for root and stem tissue and stem latex i.e. they were allowed to grow to maturity before dried capsules were harvested for analysis. HM2 is the wild type morphine (3.41% dw) line from which the mutants were derived. The E193K and R158K mutations are in CODMa/c and CODMb, respectively. The W261* and Q254* nonsense mutations are in CODMa/c and CODMb, respectively. The W145* nonsense mutation and the Tint polymorphism are in T6ODMa.

4.2.2 Evaluation of latex compositions of hk, flw and mat plants

Previous work in the lab has shown that latex analysis of plants post flowering (mat; 2-3 days after petal fall) is a useful indicator of what the final phenotype will be in the capsules.

shows the latex alkaloid composition of each of the six lines at different development stages-hook stage, flowering and maturing plants. For each line, approximately 90% of the alkaloids in the latex are accounted for by morphine, oripavine, codeine, thebaine and the codeine-thebaine dimer.

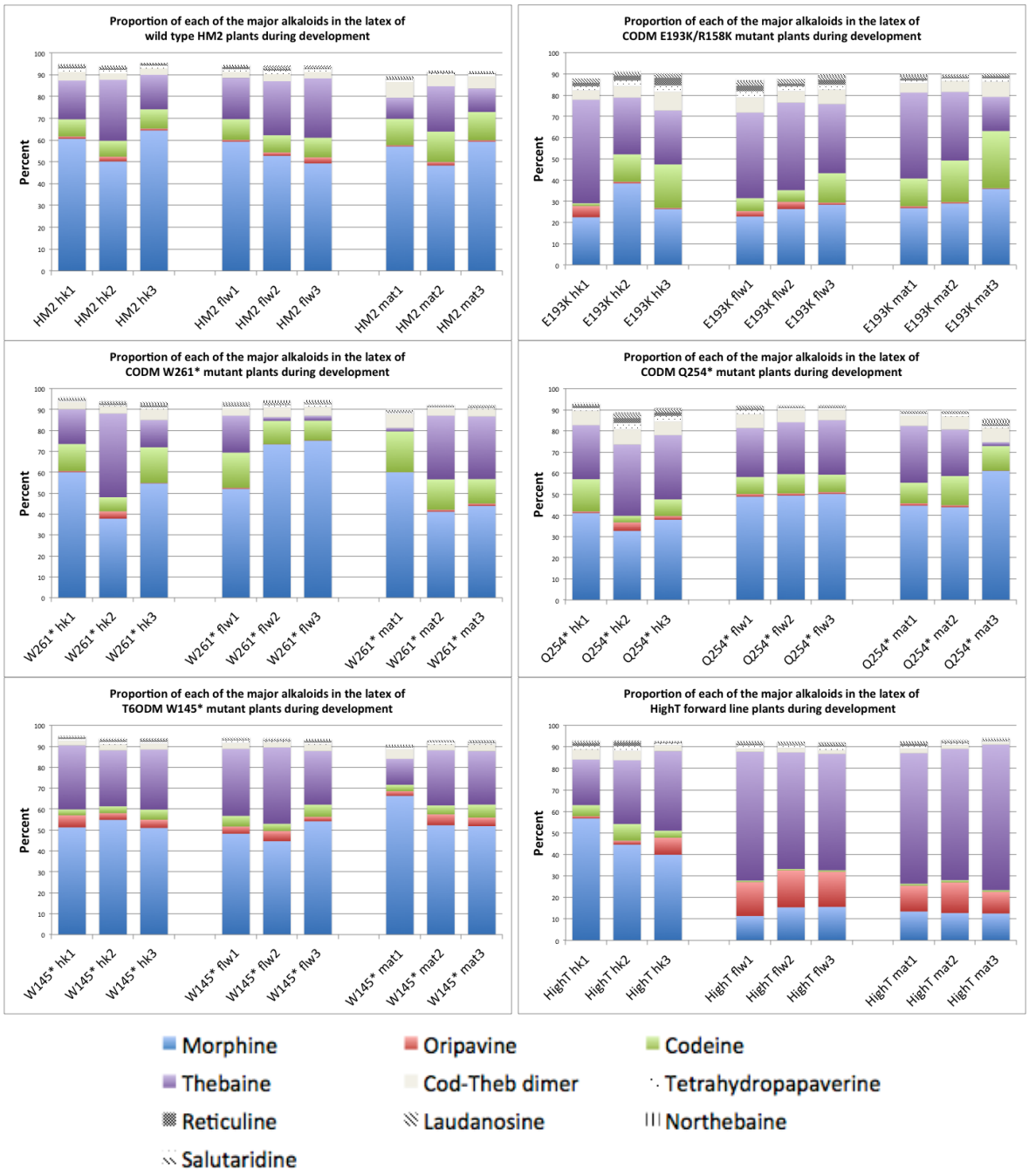


Figure 40. Proportion of each of the major alkaloids of the latex during development (hk= hook stage, flw=flowering and mat=maturing stages) in HM2 and the mutant lines. The top ten most abundant identifiable alkaloids are shown. Metabolite figures are an average of 3 technical replicates for each sample. Columns are not stacked to 100% because of the presence of unknown alkaloids and trace levels of known alkaloids.

In HM2, the proportion of codeine is higher after flowering. This appears to be at the expense of thebaine rather than morphine. This would indicate metabolism of thebaine to morphine through neopinone, codeinone and codeine. Typically, in mature capsules levels of each of codeine and thebaine are much reduced to around 3% of the total alkaloid content as the plant converts these alkaloids to the most abundant alkaloid morphine.

Surprisingly, in the high codeine forward screen line (E193K/R158K) the proportion of codeine was found to be higher than HM2 only in maturing plants. However, the proportion of thebaine is higher across all stages and this is at the expense of morphine. Therefore in at the hook and flowering stages the metabolic block may be at demethylation of thebaine to oripavine. In more mature plants, the block may be affecting demethylation of codeine to morphine, resulting in the final high codeine phenotypes. Alternatively, at the hook and flowering stage a partial block of the demethylation of thebaine to neopinone could explain the higher levels of thebaine. As the capsule matures this block may be released and the block affecting the demethylation of codeine to morphine becomes observable in elevated codeine levels.

Phenotypes reminiscent of HM2 are seen in the latex of W261* plants. Codeine levels are slightly increased at the expense of thebaine while morphine is still the predominant alkaloid in the latex. The latex profiles of the other CODM mutant line, Q254*, are broadly similar to that of HM2. However, this line has been shown previously (Figure 26) to have up to 3x more codeine in the mature capsules of plants grown on soil, which suggests that the CODMb may act at a later stage to convert any remaining codeine to morphine.

The T6ODM W145* mutant displays the predicted phenotype clearly. A block at T6ODM should result in a larger proportion of thebaine and oripavine and this is observed in the latex samples of plants at the hook, flowering and maturing stages. The effect is perhaps most evident with increases in the proportion of oripavine, which would indicate a disruption in the conversion of oripavine to morphinone and on to morphine. Additional functional copies of T6ODM would allow flux through the pathway to morphine when T6ODMa is knocked out. However, lack of functional T6ODMa protein is enough to create a metabolic bottleneck resulting in the higher proportions of thebaine and oripavine in the latex of T6ODMa W145* mutant compared to HM2 and the CODM mutant lines.

The most striking phenotypes were observed with the HighT forward screen line latex. In this line there is a clear change in phenotypes between plants at hook and flowering stages.

The time period between these stages is typically 3-4 days but in that time there is a pronounced change in alkaloid composition of the latex. At hook stage the latex is not too dissimilar to HM2. However, flowering appears to be a trigger for a large increase in the proportions of thebaine and oripavine at the expense of morphine. The fact that these plants produce morphine in the early stages suggests there is not a complete general block at T6ODM. However, this block appears to become stronger post flowering leading to increased levels of thebaine and oripavine. This line contains the T6ODMa intron polymorphism also present in GSK commercial thebaine varieties (Figure 33). This T6ODMa polymorphism cannot in itself be the cause of the block at T6ODM because a knockout in this gene (W145*) only provides a partial block at T6ODM. Unless the polymorphism affects splicing (no evidence found as yet) or mRNA stability and results in a general reduction in T6ODM protein synthesis, there must be another factor involved in determining the high thebaine phenotype. The fact that there is a marked change in phenotype between plants at the hook and flowering stages suggests that a regulator that becomes active at flowering may be involved.

4.3 Morphinan gene expression

4.3.1 CODM and T6ODM expression in HM2 stems is upregulated at flowering

Figure 41 shows the fold difference changes in CODM and T6ODM gene expression as wild type HM2 plants develop. Relative transcript levels were determined using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001) with beta actin serving as a normaliser and the mean expression levels of early seedling stems as calibrators.

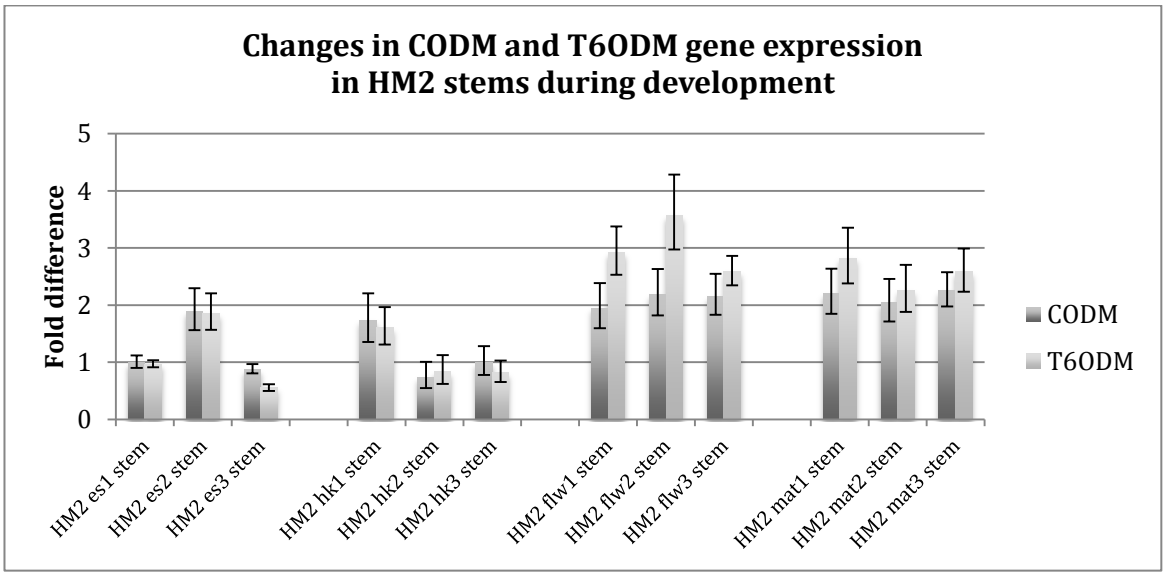


Figure 41. Changes in morphinan gene expression in HM2 stems during development from early seedlings to maturing plants. The mean expression levels of the three es stem samples were used as calibrator.

The results show that CODM and T6ODM gene expression in stems are similar in early seedling and hook stage plants. Expression is increased approximately twofold at flowering and this increase is maintained in maturing plants.

The relative change in abundance of T6ODM transcripts is higher than that for CODM in all flowering and maturing plants although overall CODM transcripts are more abundant (lower Ct values). This change is most pronounced in flowering plants and could relate to the point at which thebaine in the latex is beginning to be metabolised to morphine.

4.3.2 CODM and T6ODM expression in HM2 roots is also upregulated at flowering

Figure 42 shows the fold difference changes in CODM and T6ODM gene expression in the roots as the same (Figure 41) wild type HM2 plants develop. Relative transcript levels were determined using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001) with beta actin serving as a normaliser and the mean of early seedling root expression of each target gene as calibrators.

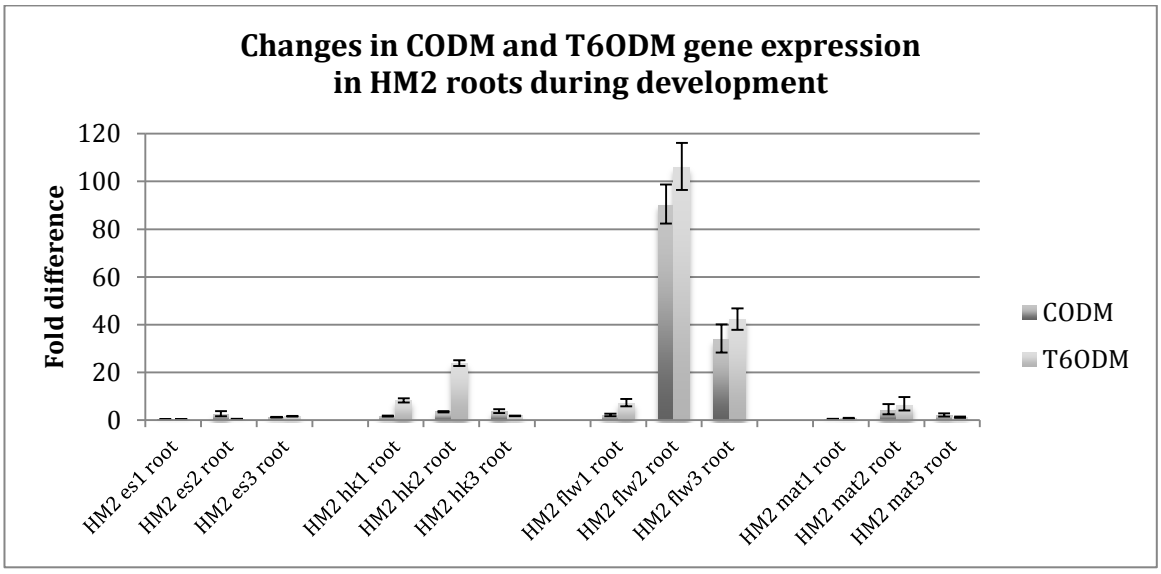


Figure 42. Changes in morphinan gene expression in HM2 roots during development from early seedlings to maturing plants. The mean expression levels of the three es root samples were used as calibrator.

CODM and T6ODM are expressed at lower levels in the roots compared to stems. The main site of BIA accumulation is the laticifers, which are associated with the phloem in the shoot. The latex analysed from these plants is simply the cytoplasm of laticifers. Latex does not exude from roots in the same way it does from shoots, although laticifers are present (Facchini and Bird, 1998). Laticifers are adjacent to other phloem cell types, known as sieve elements and companion cells, in aerial organs but are not necessarily proximal in roots (Liscombe and Facchini, 2008). Companion cells produce the enzymes required to synthesise alkaloids. These are translocated to sieve elements where alkaloids are produced. They are then transported to the laticifers for storage (Beaudoin and Facchini, 2014; Facchini et al., 2007; Hagel et al., 2008b).

However, CODM and T6ODM transcripts were detected in root material and alkaloids could be detected in dried and lyophilised roots. The transcript levels are comparatively low in early seedling and hook stage plants, but there appears to be a significant upregulation of expression at flowering time where expression levels are up to 100x higher than that in early seedlings.

4.3.3 Expression in High T fwd line stems is also upregulated at flowering

In the high thebaine forward screen line, CODM and T6ODM expression is increased at flowering and returns to similar levels to hook stage plants in maturing plants (Figure 43).

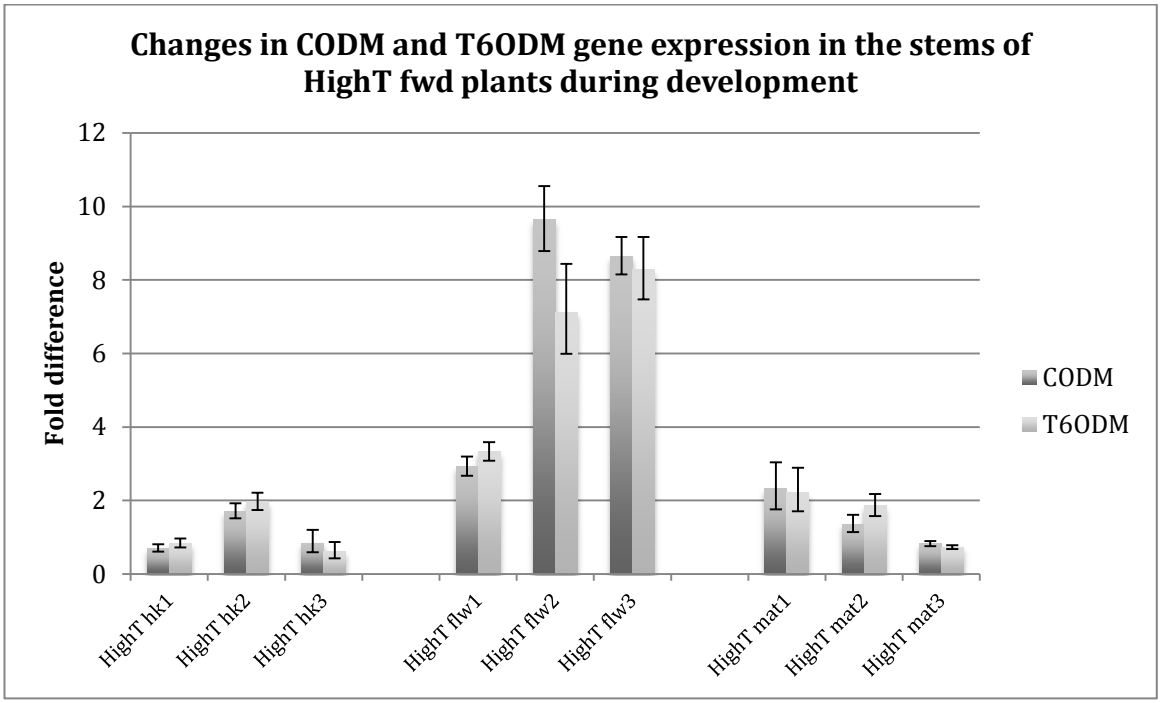


Figure 43. Changes in morphinan gene expression in HighT fwd line stems during development from hk stage plants to maturing plants. The mean expression levels of the three hook stem samples were used as calibrator as early seedling plants were not sampled.

Gene expression changes between hook and flowering stage plants correspond with a dramatic shift in the latex profiles with increases in the proportion of thebaine and oripavine and reductions in the proportions of codeine and morphine (Figure 40). Prior to flowering the plants are able to produce morphine but after flowering it appears that there is a block at thebaine that prevents the conversion of thebaine and oripavine to morphine that ordinarily occurs in HM2.

The fold changes in CODM/T6ODM transcripts at flowering are much greater than that observed in HM2, which could reflect a recognition by the plant of significantly high proportions of thebaine in the latex at flowering and an attempt to metabolise thebaine to morphine. The high levels of CODM/T6ODM expression at flowering suggests that factors other than gene expression levels are responsible for the metabolic block. Upregulation, unlike with HM2, appears to be transient as maturing plants have similar levels of CODM/T6ODM expression as hook stage plants.

4.4 Strategy for comparison of gene expression across mutant lines

The above results indicated that the most interesting development stage in terms of changes in gene expression and alkaloid profiles is flowering. For this reason it was decided to extract RNA from all the stems and roots of all flowering plants to compare morphinan gene expression. Flowering is the stage at which CODM and T6ODM are most highly expressed and so any differences in expression levels between the mutants were anticipated to be most evident at flowering.

4.4.1 2-ODD expression was highest in the stems of flowering High T fwd mutants

Figure 44 shows a comparison of morphinan gene expression in the stems of HM2 and the five mutant lines at the flowering stage. The mean HM2 flowering plants result for each target gene acted as calibrator and expression of each gene was calculated relative to beta actin.

CODM/T6ODM/COR expression levels are similar in the three HM2 flowering stems sampled. The levels of codeineone reductase (COR) are similar to HM2 in most of the mutants (the HighT line being the exception). The differences between the lines, however, are most clearly seen with CODM and T6ODM expression levels. The high codeine forward line flowering plants (E193K/R158K) have 5 fold lower levels of CODM/T6ODM than HM2. The lower T6ODM transcript levels reflect the high thebaine content of the latex at this stage. Similar lower expression levels were observed for the CODM nonsense mutants, W261* and Q254*. As one copy of CODM is knocked out in these lines, a lower CODM transcript level may be been expected. However, any reduction in CODM transcript levels seems to be associated with corresponding reduction in T6ODM transcript levels, suggesting a form of coordinate regulation.

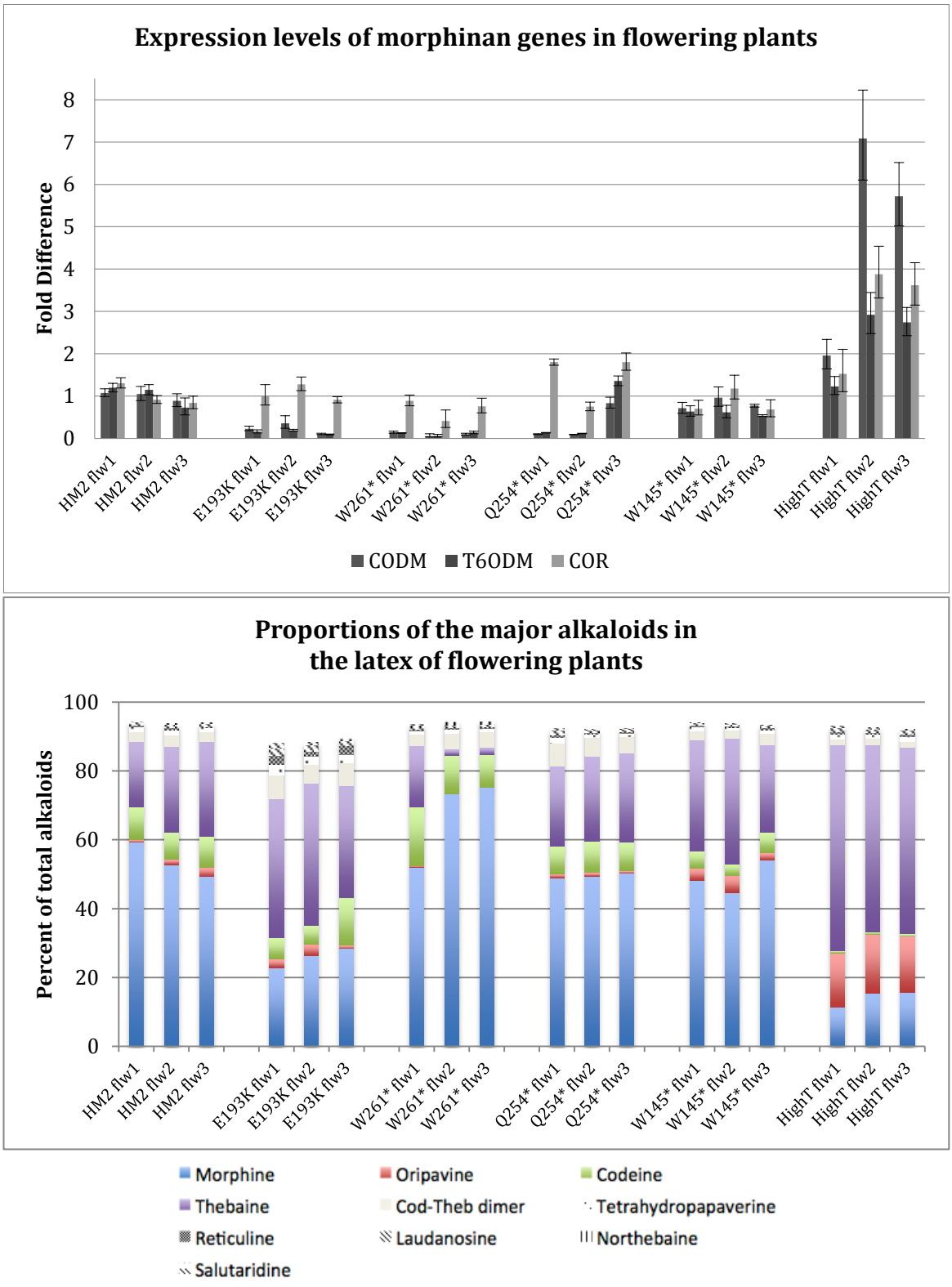


Figure 44. Morphinan gene expression and the corresponding latex alkaloid profiles of the stems of flowering plants. Metabolite figures are an average of 3 technical replicates for each sample. Columns are not stacked to 100% due the presence of unknown compounds. E193K is short for plants with both the CODMa/c E193K and CODMb R158K mutations.

T6ODM transcript levels are reduced in the W145* mutant to approximately 60% of the level in HM2. This is reflected in the slightly higher proportions of thebaine and oripavine in the latex of W145* flowering plants. Although T6ODMa is knocked out in this line, additional copies of T6ODM are being expressed and detected in the quantitative PCR.

There is a large increase in CODM/T6ODM/COR gene expression in the HighT fwd line in response to the block at thebaine and oripavine described above. Although levels of T6ODM transcripts are high, the plants appear not to metabolise thebaine like they did prior to flowering and consequently thebaine and oripavine accumulate in the latex. This suggests a cause other than expression levels of CODM/T6ODM/COR for the block in the conversion of thebaine and oripavine in the HighT line. A potential candidate may be an as yet unidentified transporter of thebaine that acts to shuttle thebaine from the sieve elements to the laticifers. Alkaloid transport in opium poppy is not yet well understood-it could occur via symplastic transport via plasmodesmata (Facchini and De Luca, 2008) or apoplastic via as yet unidentified transporters. Thebaine is thought to be the major translocated intermediate between sieve elements and laticifers (Onoyovwe et al., 2013).

4.4.2 Levels of protopine alkaloids are elevated in the roots of flowering mutant plants

Figure 44 shows a comparison of morphinan gene expression in the roots of HM2 and the five mutant lines. The mean HM2 flw results for each target gene acted as calibrator and expression of each gene was calculated relative to beta actin.

Gene expression levels in roots were highly variable, a phenomenon also observed previously (Facchini and Park, 2003). Even among the HM2 roots there were large differences in expression levels of CODM and T6ODM. Expression levels of these genes are down in the W145* mutants but this does not seem to be linked to any change in phenotype, i.e. root profiles for W145* are similar to HM2.

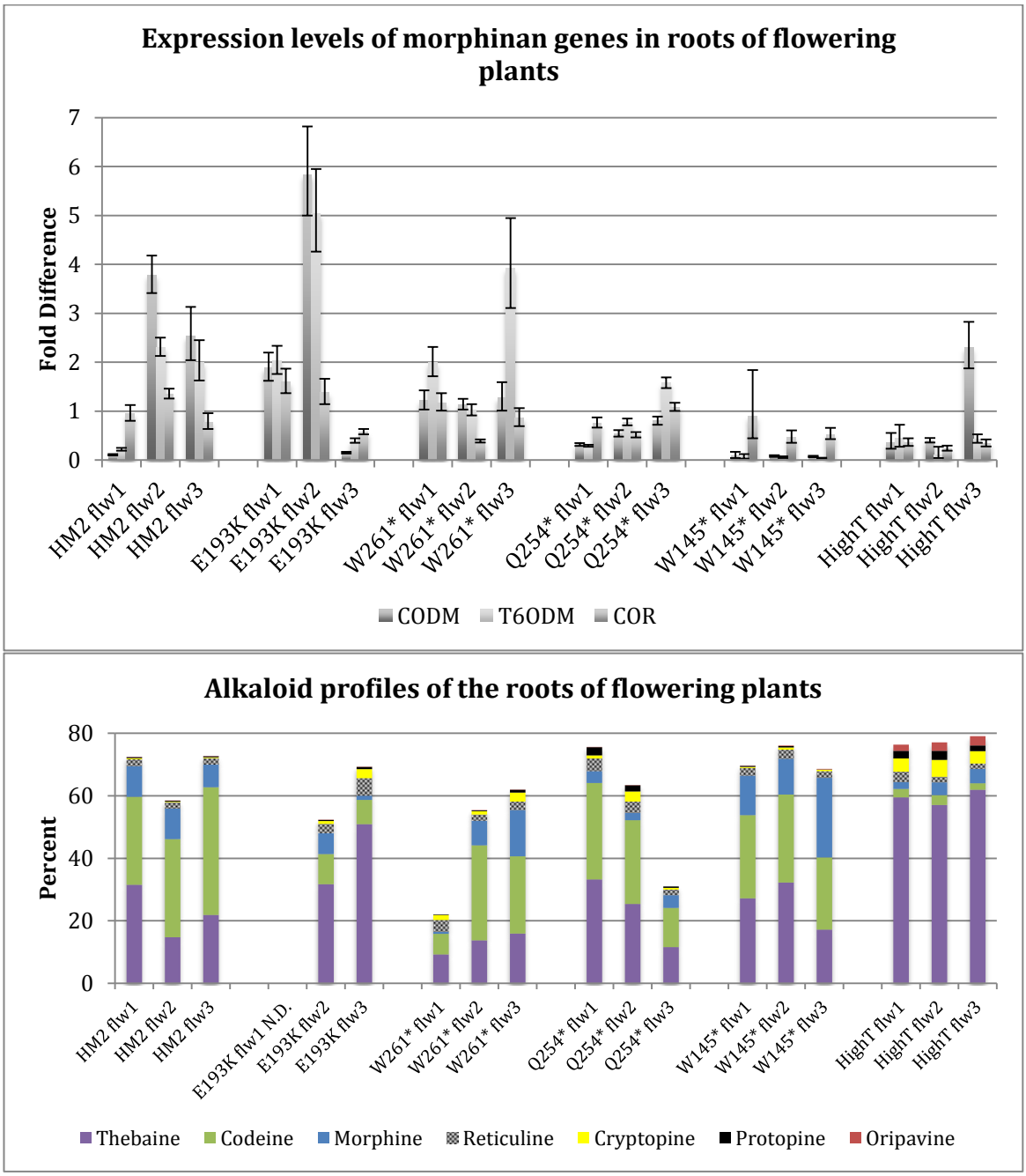


Figure 45. Morphinan gene expression and the corresponding alkaloid profiles of the roots of flowering plants. Metabolite figures are an average of 3 technical replicates for each sample. Columns are not stacked to 100% due the presence of unknown compounds, the proportion of which is much higher than that in stem latex. E193K is short for plants with both the CODMa/c E193K and CODMb R158K mutations.

There is generally a larger proportion of codeine and thebaine compared to stem latex of the same plants. In HM2 roots, codeine is the predominant alkaloid, while morphine is the most abundant alkaloid in the stem latex of the same plants. The proportions of unidentified alkaloids is much higher in the roots, in particular the roots of two plants W261* flw1 and Q254* flw3. Another interesting observation was the lack of sanguinarine in these roots. This was in contrast to a previous study that identified sanguinarine and dihydrosanguinarine as major alkaloids in an elite Tasmanian *Papaver somniferum* narcotic cultivar (Frick et al.,

2005) and the condiment "Marianne" cultivar (Facchini and De Luca, 1995; Frick et al., 2005). This may simply be due to differences between GSK's HM2 variety and previously studied opium poppy varieties where sanguinarine is detected (Facchini et al., 1996b; Facchini and Park, 2003; Huang and Kutchan, 2000). HM2 is the result of many years of breeding for morphine and as such it may divert all its available resources to that end. Opium poppy cell cultures are known to accumulate sanguinarine in response to fungal elicitor treatment (Alcantara et al., 2005; Facchini et al., 1996a; Zulak et al., 2006). The absence of fungi in the silica sand/Terragreen mixture used to grow the plants in for this experiment may be a reason why sanguinarine was not found in roots (and stem latex). Sanguinarine has been shown to be toxic to fungi at low concentrations (Cline and Coscia, 1988) and its presence roots may offer protection against pathogens in the soil.

Levels of cryptopine and protopine were higher among the mutants compared to HM2 (each approximately 0.5% of total alkaloids in roots). This is in agreement with a previous study which silenced CODM and T6ODM in opium poppy using Virus Induced Gene Silencing (VIGS; Farrow and Facchini, 2013). VIGS resulted in increases in cryptopine and protopine in the latex due to newly discovered roles for these enzymes in the metabolism of protopine, benzophenanthridine and rhoeadine alkaloids. The proportions of cryptopine and protopine are as high as 3% in some of the CODM mutants and remain unchanged in the T6ODMa W145* mutant (Figure 46). The High T fwd line displays the highest levels of protopines in its roots. Expression of the morphinan genes in the roots of this line is down compared to HM2 (Figure 45). This is in contrast to the situation in stems where expression levels are much higher than in HM2 (Figure 43).

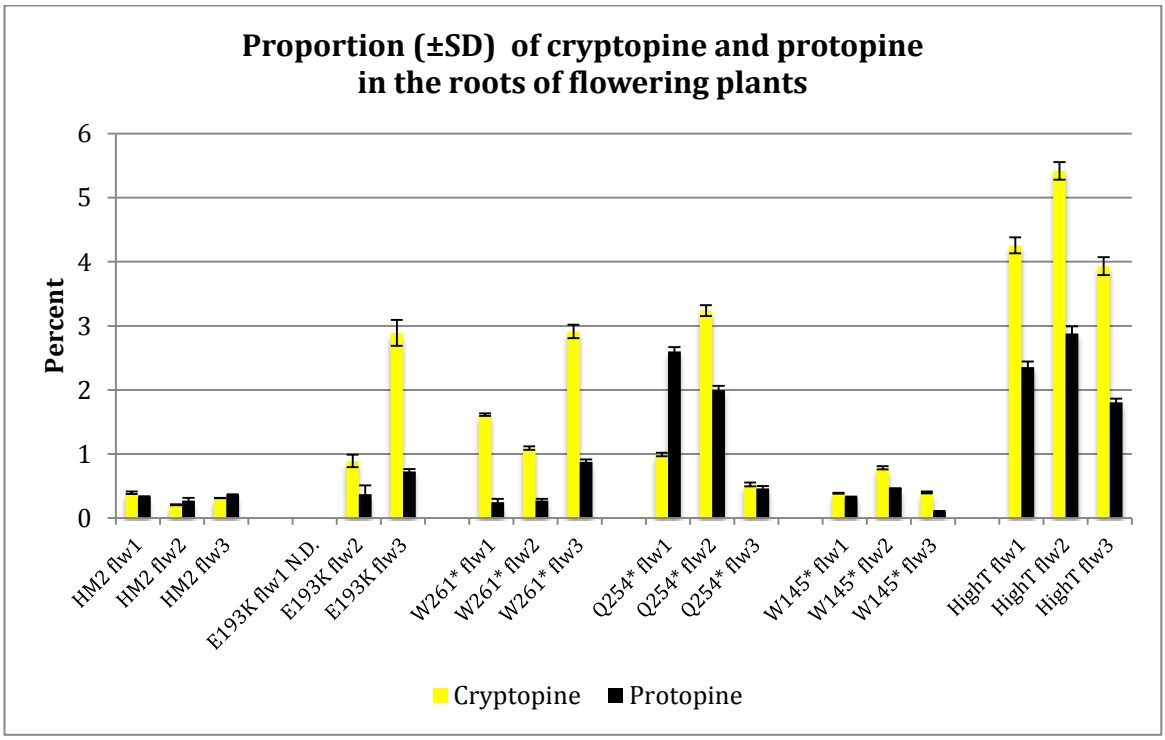


Figure 46. Proportion of cryptopine and protopine alkaloids in the roots of flowering plants. The figures are an average of 3 technical replicates for each sample. E193K is short for plants with both the CODMa/c E193K and CODMb R158K mutations.

4.5 Discussion

The results indicate that flowering seems to be trigger for a change in gene expression and alkaloid composition of the latex. Even though plants sampled at the hook stage are only 2-3 days younger than flowering plants, there is approximately a two fold increase in morphinan gene expression levels in flowering plants. Other examples where secondary metabolism has been shown to be coupled to a plant's developmental stage include *Hypericum*, where the amounts of several compounds change across five phenological stages (Cirak et al., 2013), and *Solanum lycopersicum* (tomato), where the levels of carotenoids, polyphenols, volatiles and alkaloids change during fruit ripening (Tohge et al., 2014).

Analysis of the latex of developing plants is a useful indicator of the composition. However, its limitation is that it is not quantitative. Therefore, even if absolute levels of a certain alkaloid, e.g. codeine or thebaine, increased in later stages of development, if levels of morphine also increased it may not evident from looking at percentages of each alkaloid present in latex. Nevertheless, a clear change in phenotype at flowering was observed with the high T forward screen line. Pre–flowering the morphine comprises approximately 50% of the alkaloids in the latex. At flowering this number reduces to 10% with corresponding

increases in the proportion of thebaine and oripavine. It may be the case that further morphine production is halted at flowering and the relatively small amounts of morphine in mature capsules were made prior to this stage. Post-flowering flux into the morphinan pathway was blocked at thebine with the result that the major alkaloid components of both the latex (Figure 40) and capsules (Figure 33) are thebaine and oripavine (Figure 47).

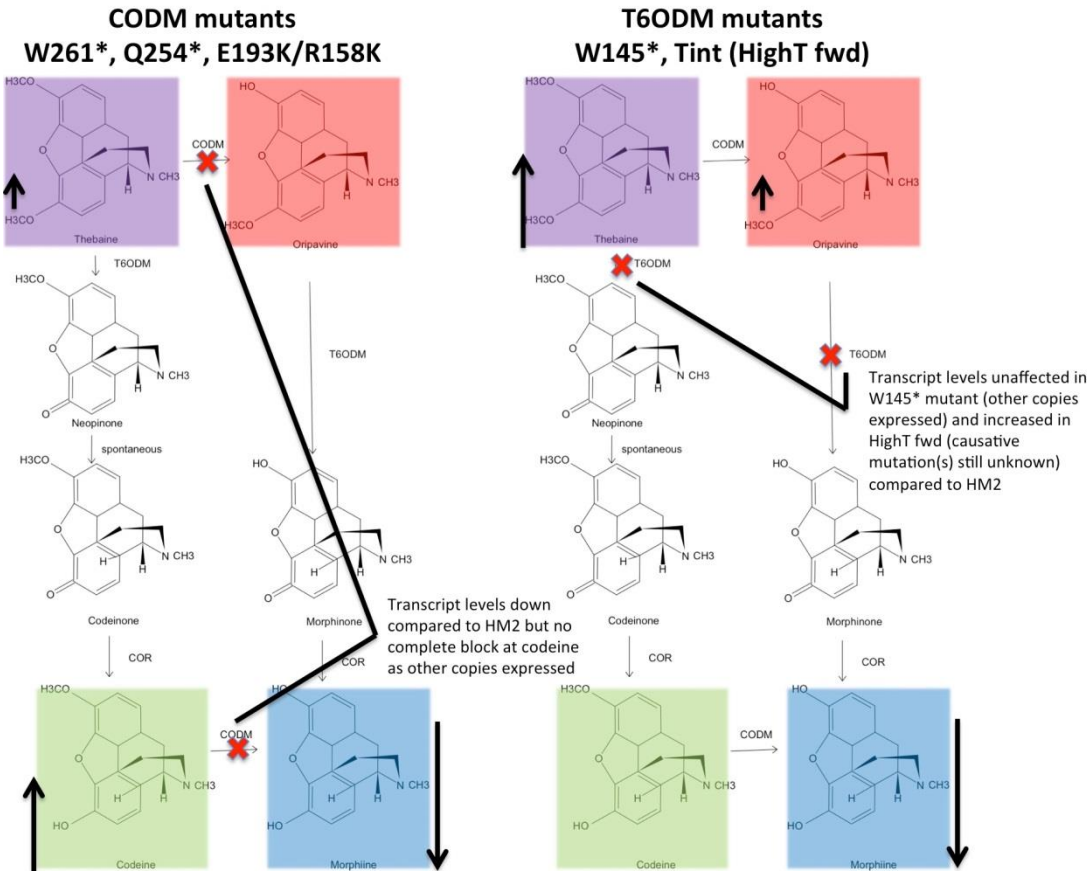


Figure 47. Summary showing the effects each of the known mutations have on flux through the morphinan pathway. In each case the red 'x' does not reflect a complete block because additional functional copies of CODM/T6ODM appear to operate in each of the mutants. The 'x' rather indicates a reduction in efficiency of the reactions indicated. After thebaine the pathway to morphine is bifurcated. The major route to morphine from thebaine is thought to occur via codeinone (Nielson et al., 1983). The minor route involves the intermediates oripavine and morphinone (Brochmann-Hanssen, 1984). Mutations in CODM therefore result in increases in the proportion of codeine (large black arrow) as a result of disrupting flux through the major pathway. Thebaine content (small black arrow) is also slightly higher than that found in HM2 as thebaine to oripavine reactions are also affected in these mutants. With the T6ODM W145* mutant thebaine and oripavine yields were increased at the expense of morphine due to disruption of thebaine and oripavine 6-O-demethylations. The HighT fwd has an intron polymorphism in T6ODMa or it may have other as yet undetected mutations in T6ODM, transporters of thebaine or regulators of morphinan gene activity.

The CODMa/c W261* and CODMb Q254* lines display modest increases in codeine content of mature capsules when their CODM mutations are homozygous (Figure 47). Analysis of the alkaloid composition of the roots of these lines, however, has revealed the increased levels of cryptopine and protopine alkaloids. Likewise, higher levels of protopines were observed in the roots of plants from the CODM E193K/R158K line. This confirms results

from a recent study showing that CODM has activities outside of the morphinan pathway (Farrow and Facchini, 2013) with activities in the metabolism of protopine, benzophenanthridine and rhoeadine alkaloids.

Levels of protopine alkaloids in T6ODMa W145* mutants were no different to that of wild type HM2. This was expected as recombinant T6ODMa showed no activity against protopines (Farrow and Facchini, 2013). T6ODM VIGS, however, did result in an increase in protopine and cryptopine content. It is likely therefore, that additional copies of T6ODM (b and/or c) with activities against protopine and cryptopine were co-silenced with T6ODMa in the VIGS experiments. Many of the amino acid differences between T6ODMa and T6ODMb (Table 9) also exist between CODM and T6ODMa i.e. T6ODMb is intermediate between CODM and T6ODMa. CODM has a clear activity against protopine and cryptopine and it may be that the T6ODMb is the copy of T6ODM which also metabolises these alkaloids. The highest levels of cryptopine and protopine were found in the High T forward line (2-5%) suggesting the same factor causing the metabolic block at thebaine and oripavine in stem latex is also leading to a reduction in protopine metabolism in roots.

Chapter 5- Use of detected CODM and T6ODM alleles in breeding for altered metabolite profiles

5.1 Establishment of AS-PCR for the detection of mutations/polymorphisms

Conventional plant breeding is based on phenotypic selection of superior genotypes within segregating progenies obtained from crosses. This approach is often complicated by reactions between genotypes and the environment. Marker assisted selection (MAS) avoids these problems by changing the selection criteria toward selection of genes rather than phenotypes (Francia et al., 2005). Molecular markers are unaffected by the environment and the conditions in which the plants are grown. Often quantitative trait loci (QTL) are used in MAS when the genes governing a desired trait are not known or mutations in the coding sequence which enhance a trait have not yet been identified. In this study, the markers are the single nucleotide polymorphisms (SNPs) detected in 2-ODD genes involved in morphine biosynthesis. These SNPs have been shown to alter flux to morphine in mutant lines. Crosses to combine CODM and T6ODM mutations may further increase yields of codeine, thebaine and oripavine and decisions on selection of seed for sowing could be made by marker analysis. This marker assisted gene (or QTL) pyramiding approach (MAGP) has been demonstrated successfully in a number of crops (Jiang et al., 2007; X. Li et al., 2010; Richardson et al., 2006; Wang et al., 2012). Alternatively, marker assisted recurrent selection (MARS) could be used to select genomic regions involved in the expression of complex traits and assemble the best performing genotype within a single, or across populations (Ribaut et al., 2010). Recurrent crosses could be initiated with HM6, HT5/HT6 and HN4, which are contemporary GSK morphine, thebaine and noscapine varieties, respectively. The mutations were derived from a previous morphine variety, HM2.

SNPs can be detected using allele-specific PCR primers designed such that the 3' nucleotide of a primer corresponds to the site of the SNP (Ugozzoli and Wallace, 1991). Thus, the allele-specific primer matches perfectly with one allele and has a 3' mismatch with the nonspecific allele. Because mismatched 3' termini are extended by DNA polymerases with

much lower efficiency than correctly matched termini, the allele-specific primer should preferentially amplify the specific allele (Cha et al., 1992).

Tracking of double mutants via AS-PCR would reduce the amount of expensive cleavage enzyme used if TILLING would be the sole genotyping strategy. Ensuring that different sized PCR products would be formed with the presence of various mutations would enable screening of each plant for multiple mutations simultaneously e.g. CODM and T6ODM mutations. This would save time and effort that would be spent TILLING the same DNA sample multiple times for different target genes.

Primers used to track CODM and T6ODM mutations and polymorphisms in AS-PCR are listed in Appendix D. Successful PCR with mutant primers would indicate the presence of a mutation. The limitation of this approach is that it cannot distinguish between homozygotes and heterozygotes. AS-PCR was used to confirm the presence of mutations in F1s (mutations assumed to be heterozygous as it was present in just one of the gametes). For F2s and subsequent generations, AS-PCR was used as first step in genotyping to detect the presence of segregating mutations. TILLING the same DNA samples would then clarify whether detected mutations were in homozygous or heterozygous states. In heterozygotes a heteroduplex would form between amplified wild type and mutant alleles, which would be cut and detected with the Fragment Analyser. In contrast PCR products generated from homozygotes would be identical, would not be cut and only full length PCR products would be visible on the TILLING gel. Those F2 candidates with desirable genotypes would then have CODM and/or T6ODM sequenced to confirm zygosity in a third genotyping step prior to advancement to F3.

5.2 Crossing strategy

TILLING and 'EcoTILLING' CODM and T6ODM identified a number of alleles with the potential to disrupt O-demethylation of codeine, thebaine and oripavine. Phenotypic analysis of M2s, M3s and forward screen lines with CODM/T6ODM polymorphisms demonstrated that the effects of the polymorphisms in isolation are limited. Efforts to combine CODM

polymorphisms were already detailed in Chapter 3 (3.4.5). It is likely that double mutants were not recovered due to the close proximity of copies of CODM in the genome. BAC sequence information that suggests that copies of CODM might be closely linked was not available at the time the crosses were being set up. Therefore, attempts were made at the time to combine CODM polymorphisms, in particular to set up crosses to replace the CODMb R158K polymorphism in the high codeine forward screen lines with the CODMb Q254* mutation found by reverse genetics. In addition, attempts were made to combine CODM and T6ODM mutations/polymorphisms to create plants with partial blocks in both sides of the pathway to morphine from thebaine. As EMS introduced random mutations across the genome and it is estimated that each M2 carries as much as 10,000 induced mutations, the process of backcrosses to GSK elite lines was initiated with crosses of mutants to current commercial morphine (HM6), thebaine (HT5 & HT6) and noscapine (HN4) lines. Finally, lines with interesting metabolite profiles, but for which no CODM or T6ODM polymorphisms have been found, were included in the crossing scheme. These include high thebaine forward screen lines that arose from HM2 mutagenesis and noscapine-free thebaine/oripavine only lines that arose from HN3 mutagenesis.

Figure 48 shows an overview of the 76 crosses carried out. More detail on each cross is provided in Appendix E. F1 seed was harvested from mature females and sown.

Figure 48. Overview of the parents used in each of the 76 crosses. Parents were drawn from forward screen, reverse screen lines and GSK commercial lines (blue). Those lines with known (dark) or predicted (light) codeine (green) and thebaine (purple) phenotypes are indicated alongside detected CODM and T6ODM mutations. Wk=week each seed batch was sown (if struckthrough the seed did not germinate). More information on each individual cross is provided in Appendix E.

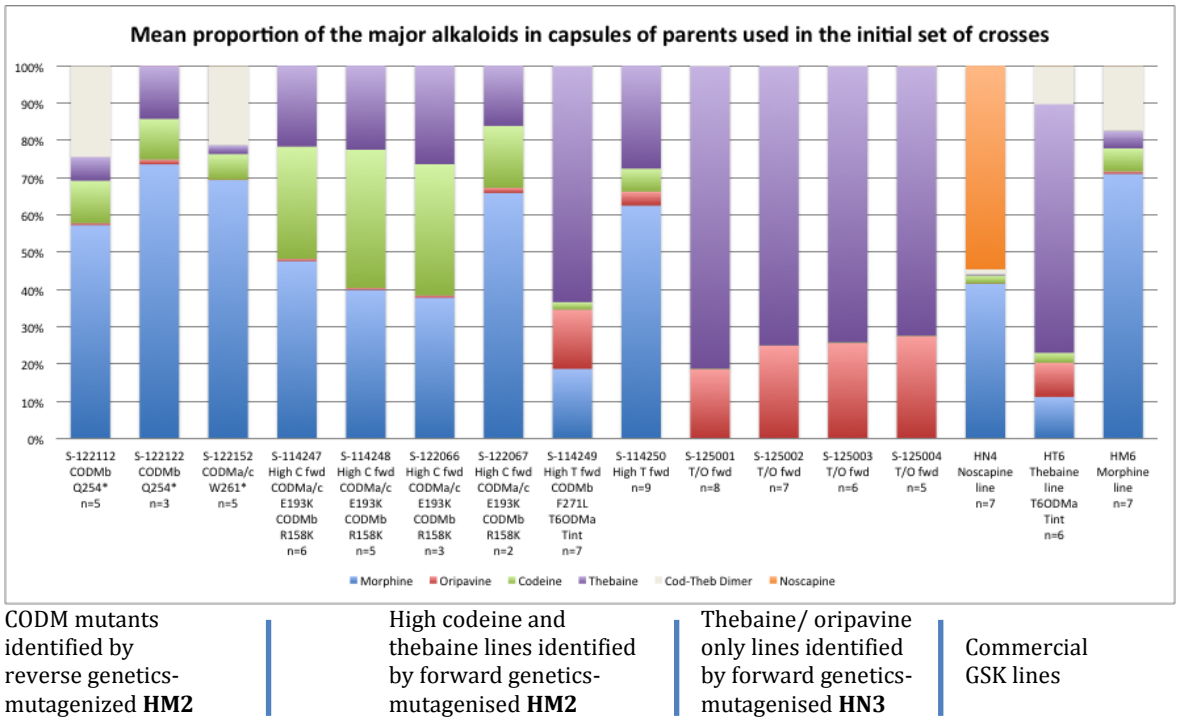


Figure 49. Phenotypes of the lines used in the initial set of crosses. The proportions of each of the major alkaloids in the capsules of the parents actually used in the crosses detailed in Figure 46 are shown. Known mutations and polymorphisms in each line are detailed.

Figure 49 provides an overview of the phenotypes of the parental lines used in the initial crosses (Appendix E, X001 to X056). Known mutations and polymorphisms in each line are indicated. CODMb Q254* TILLING mutants, in particular, had more codeine in their capsules than HM6. T6ODMa TILLING mutants are not included as they were discovered at a later date before being brought into the breeding programme. Their phenotypes are described in Figure 38. Four of the high codeine forward screen lines were included as parents. These lines contained CODM E193 and R158K polymorphisms (Figure 20). S-114249 and S-114250 are high thebaine forward screen lines that came out of HM2 mutagenesis. CODM and T6ODM polymorphisms were detected in the former using TILLING (Figure 33). The four thebaine-oripavine only lines (S-125001 to S-125004) came out of a forward screen of mutagenised HN3, a former noscapine line. The molecular basis of the phenotype remains unknown but they were included in the crosses because the high oripavine content was of particular interest.

5.3 Tracking of transferred mutant alleles in F1 individuals

DNA was extracted from ten F1 individuals for each seed batch sown. For the first batch of F1s, two (96 well) DNA plates were screened by AS-PCR for the Q254*, E193K, R158K and Tint SNPs. PCR products were visualised on agarose gels (Appendix F). Successful amplification indicated the presence of SNPs contributed by parents (Table 19). Additional F1s, including those resulting from crosses involving T6ODMa TILLING mutants (Appendix E, X057 to X075), were later screened using a similar approach but with visualisation of AS-PCR products on the Fragment Analyser (Appendix F).

In most cases alleles were successfully detected in the F1s. For example for X009 (Table 19a) all three CODM SNPs were detected (+ for each SNP in each of the ten F1s tested), two contributed by the high codeine forward screen line and the other from the TILLING mutant. The aim of this cross was to replace the R158K mutation in the high codeine forward screen line with the CODMb knockout mutation to further impact on the capacity of the plant to O-demethylate codeine. The F1s possess both R158K and Q254* mutations and perhaps segregation of these in F2 would result in the recovery of E193K-Q254* mutants.

X008, X001, X027, X030, X055 and X016 were carried out to begin the process of introducing detected CODM and T6ODM mutations to elite GSK germplasm i.e. morphine (HM6), thebaine (HT6) and noscapine (HN4) lines. The remaining crosses aimed to combine CODM, T6ODM and undetected mutations (in the high thebaine forward screen lines) to impact on both paths of the pathway to morphine from thebaine.

The T6ODMa intron polymorphism observed in GSK high thebaine lines was not transferred to F1s in X043 (Table 19b). The polymorphism is present in most but not all HT6 individuals (Table 19c and Figure 34). The HT6 individual used as a parent in X043 did not carry the polymorphism and so this cross was not progressed from this stage.

Mother Seed batch id Seedling id Known mutations	Father Seed batch id Seedling id Known mutations	F1 seed batch id Cross id	F1 Seedling id	T6ODMa intron SNP	CODMa/c E193K	CODMb R158K	CODMb Q254*
S-122112 Sd-833514 **CODMb Q254***	S-114247 Sd-833385 **CODMa/c E193K** **CODMb R158K**	S-124006 X009	Sd-835300	-	+	+	+
			Sd-835301	-	+	+	+
			Sd-835302	-	+	+	+
			Sd-835303	-	+	+	+
			Sd-835304	-	+	+	+
			Sd-835305	-	+	+	+
			Sd-835306	-	+	+	+
			Sd-835307	-	+	+	+
			Sd-835308	-	+	+	+
			Sd-835309	-	+	+	+
HM6 Sd-833541 M line	S-122112 Sd-833508 **CODMb Q254***	S-123958 X008	Sd-835310	-	N/A	N/A	+
			Sd-835311	-	N/A	N/A	+
			Sd-835312	-	N/A	N/A	+
			Sd-835313	-	N/A	N/A	+
			Sd-835314	-	N/A	N/A	+
			Sd-835315	-	N/A	N/A	+
			Sd-835316	-	N/A	N/A	+
			Sd-835317	-	N/A	N/A	+
			Sd-835318	-	N/A	N/A	+
			Sd-835319	-	N/A	N/A	+
S-122067 Sd-833459 **CODMa/c E193K** **CODMb R158K**	S-114139 Sd-833684 **CODMb F271L** **T6ODMa intron SNP**	S-124003 X050	Sd-835320	+	+	+	N/A
			Sd-835321	+	+	+	N/A
			Sd-835322	+	+	RF	N/A
			Sd-835323	+	+	+	N/A
			Sd-835324	+	+	+	N/A
			Sd-835325	+	+	+	N/A
			Sd-835326	+	+	+	N/A
			Sd-835327	+	+	+	N/A
			Sd-835328	+	+	+	N/A
			Sd-835329	+	+	+	N/A
S-114139 Sd-834053 **CODMb F271L** **T6ODMa intron SNP**	S-114156 Sd-834078 **CODMa/c E193K** **CODMb R158K**	S-124013 X051	Sd-835330	+	+	+	N/A
			Sd-835331	+	+	+	N/A
			Sd-835332	+	+	+	N/A
			Sd-835333	+	+	+	N/A
			Sd-835334	+	-	-	N/A
			Sd-835335	+	-	-	N/A
			Sd-835336	+	+	+	N/A
			Sd-835337	+	+	+	N/A
			Sd-835338	+	+	+	N/A
			Sd-835339	+	+	+	N/A
HT6 Sd-833560 T line **T6ODMa intron SNP**	S-114250 Sd-833486 High T Fwd	S-123968 X042	Sd-835340	+	N/A	N/A	N/A
			Sd-835341	+	N/A	N/A	N/A
			Sd-835342	+	N/A	N/A	N/A
			Sd-835343	+	N/A	N/A	N/A
			Sd-835344	+	N/A	N/A	N/A
			Sd-835345	+	N/A	N/A	N/A
			Sd-835346	+	N/A	N/A	N/A
			Sd-835347	+	N/A	N/A	N/A
			Sd-835348	+	N/A	N/A	N/A
			Sd-835349	+	N/A	N/A	N/A

Table 19a) Genotyping results for first batch of F1s. N/A indicates where DNA samples were not tested for this SNP because none of the parents had it. RF indicates reaction failure. Reactions were not repeated as there were sufficient siblings with the desired genotype (shaded) to take forward to F2.

Mother Seed batch id Seedling id Known mutations	Father Seed batch id Seedling id Known mutations	F1 seed batch id Cross id	F1 Seedling id	T6ODMa intron SNP	CODMa/c E193K	CODMb R158K	CODMb Q254*
HT6 Sd-833577 T line **T6ODMa intron SNP**	S-114250 Sd-833816 High T Fwd	S-123966 X043	Sd-835350	-	N/A	N/A	N/A
			Sd-835351	-	N/A	N/A	N/A
			Sd-835352	-	N/A	N/A	N/A
			Sd-835353	-	N/A	N/A	N/A
			Sd-835354	-	N/A	N/A	N/A
			Sd-835355	-	N/A	N/A	N/A
			Sd-835356	-	N/A	N/A	N/A
			Sd-835357	-	N/A	N/A	N/A
			Sd-835358	-	N/A	N/A	N/A
			Sd-835359	-	N/A	N/A	N/A
S-114250 Sd-833498 High T Fwd	S-114249 Sd-833464 **CODMb F271L** **T6ODMa intron SNP**	S-124001 X014	Sd-835360	+	N/A	N/A	N/A
			Sd-835361	+	N/A	N/A	N/A
			Sd-835362	+	N/A	N/A	N/A
			Sd-835363	+	N/A	N/A	N/A
			Sd-835364	+	N/A	N/A	N/A
			Sd-835365	+	N/A	N/A	N/A
			Sd-835366	+	N/A	N/A	N/A
			Sd-835367	+	N/A	N/A	N/A
			Sd-835368	+	N/A	N/A	N/A
			Sd-835369	+	N/A	N/A	N/A
S-114139 Sd-833690 **CODMb F271L** **T6ODMa intron SNP**	S-114250 Sd-833824 High T Fwd	S-124066 X044	Sd-835370	+	N/A	N/A	N/A
			Sd-835371	+	N/A	N/A	N/A
			Sd-835372	+	N/A	N/A	N/A
			Sd-835373	+	N/A	N/A	N/A
			Sd-835374	+	N/A	N/A	N/A
			Sd-835375	+	N/A	N/A	N/A
			Sd-835376	+	N/A	N/A	N/A
			Sd-835377	+	N/A	N/A	N/A
			Sd-835378	+	N/A	N/A	N/A
			Sd-835379	+	N/A	N/A	N/A
S-122112 Sd-833503 **CODMb Q254***	S-114249 Sd-833464 **CODMb F271L** **T6ODMa intron SNP**	S-124009 X015	Sd-835380	+	N/A	N/A	+
			Sd-835381	+	N/A	N/A	+
			Sd-835382	+	N/A	N/A	+
			Sd-835383	+	N/A	N/A	+
			Sd-835384	+	N/A	N/A	+
			Sd-835385	+	N/A	N/A	+
			Sd-835386	+	N/A	N/A	+
			Sd-835387	+	N/A	N/A	+
			Sd-835388	+	N/A	N/A	+
			Sd-835389	+	N/A	N/A	+
S-122122 Sd-833985 **CODMb Q254***	S-114139 Sd-834043 **CODMb F271L** **T6ODMa intron SNP**	S-124015 X054	Sd-835390	+	N/A	N/A	+
			Sd-835391	+	N/A	N/A	+
			Sd-835392	+	N/A	N/A	+
			Sd-835393	+	N/A	N/A	+
			Sd-835394	+	N/A	N/A	+
			Sd-835395	+	N/A	N/A	+
			Sd-835396	+	N/A	N/A	+
			Sd-835397	+	N/A	N/A	+
			Sd-835398	+	N/A	N/A	+
			Sd-835399	+	N/A	N/A	+
HM6 Sd-833543 M line	S-114248 Sd-833401 **CODMa/c E193K** **CODMb R158K**	S-123957 X001	Sd-835400	-	+	+	N/A
			Sd-835401	-	+	+	N/A
			Sd-835402	-	+	+	N/A
			Sd-835403	-	+	+	N/A
			Sd-835404	-	+	+	N/A
			Sd-835405	-	+	+	N/A
			Sd-835406	-	+	+	N/A
			Sd-835407	-	+	+	N/A
			Sd-835408	-	+	+	N/A
			Sd-835409	-	+	+	N/A
HM6 Sd-833559 M line	S-122066 Sd-833434 **CODMa/c E193K** **CODMb R158K**	S-123955 X027	Sd-835410	-	-	-	N/A
			Sd-835411	-	-	-	N/A
			Sd-835412	-	-	-	N/A
			Sd-835413	-	+	+	N/A
			Sd-835414	-	-	-	N/A
			Sd-835415	-	-	-	N/A
			Sd-835416	-	-	-	N/A
			Sd-835417	-	-	-	N/A
			Sd-835418	-	-	-	N/A
			Sd-835419	-	-	-	N/A

Table 19 b). Genotyping results for first batch of F1s. N/A indicates where DNA samples were not tested for this SNP because none of the parents had it. Those F1s with desirable genotypes are shaded.

Mother Seed batch id Seedling id Known mutations	Father Seed batch id Seedling id Known mutations	F1 seed batch id Cross id	F1 Seedling id	T6ODMa intron SNP	CODMa/c E193K	CODMb R158K	CODMb Q254*
HT6 Sd-833578 T line **T6ODMa intron SNP**	S-114247 Sd-833390 **CODMa/c E193K** **CODMb R158K**	S-123971 X030	Sd-835420	+	+	+	N/A
			Sd-835421	+	+	+	N/A
			Sd-835422	+	+	+	N/A
			Sd-835423	+	+	+	N/A
			Sd-835424	+	+	+	N/A
			Sd-835425	+	+	+	N/A
			Sd-835426	+	+	+	N/A
			Sd-835427	+	+	+	N/A
			Sd-835428	+	+	+	N/A
			Sd-835429	+	+	+	N/A
HT6 Sd-833942 T line **T6ODMa intron SNP**	S-114248 Sd-833628 **CODMa/c E193K** **CODMb R158K**	S-124068 X055	Sd-835430	+	+	+	N/A
			Sd-835431	+	+	+	N/A
			Sd-835432	+	+	+	N/A
			Sd-835433	+	+	+	N/A
			Sd-835434	+	+	+	N/A
			Sd-835435	+	+	+	N/A
			Sd-835436	+	+	+	N/A
			Sd-835437	+	+	+	N/A
			Sd-835438	+	+	+	N/A
			Sd-835439	+	+	+	N/A
HN4 Sd-833580 N line	S-114249 Sd-833464 **CODMb F271L** **T6ODMa intron SNP**	S-123972 X016	Sd-835440	+	N/A	N/A	N/A
			Sd-835441	+	N/A	N/A	N/A
			Sd-835442	+	N/A	N/A	N/A
			Sd-835443	+	N/A	N/A	N/A
			Sd-835444	+	N/A	N/A	N/A
			Sd-835445	+	N/A	N/A	N/A
			Sd-835446	+	N/A	N/A	N/A
			Sd-835447	+	N/A	N/A	N/A
			Sd-835448	+	N/A	N/A	N/A
			Sd-835449	+	N/A	N/A	N/A
S-114139 Sd-833680 **CODMb F271L** **T6ODMa intron SNP**	S-125001 Sd-833604 T/O line	S-124064 X041	Sd-835450	+	N/A	N/A	N/A
			Sd-835451	+	N/A	N/A	N/A
			Sd-835452	+	N/A	N/A	N/A
			Sd-835453	+	N/A	N/A	N/A
			Sd-835454	+	N/A	N/A	N/A
			Sd-835455	+	N/A	N/A	N/A
			Sd-835456	+	N/A	N/A	N/A
			Sd-835457	+	N/A	N/A	N/A
			Sd-835458	+	N/A	N/A	N/A
			Sd-835459	RF	N/A	N/A	N/A
HT6 T line **T6ODMa intron SNP**			Sd-835550	+	-	-	-
			Sd-835551	-	-	-	-
			Sd-835552	-	-	-	-
			Sd-835553	+	-	-	-
			Sd-835554	+	-	-	-
			Sd-835555	+	N/A	N/A	N/A
			Sd-835556	+	N/A	N/A	N/A
			Sd-835557	+	N/A	N/A	N/A
S-114139 **CODMb F271L** **T6ODMa intron SNP**		S-114139	Sd-835566	+	N/A	N/A	N/A
			Sd-835567	+	N/A	N/A	N/A
			Sd-835568	+	N/A	N/A	N/A
			Sd-835569	+	N/A	N/A	N/A
			Sd-835570	+	N/A	N/A	N/A
			Sd-835571	+	N/A	N/A	N/A
			Sd-835572	-	N/A	N/A	N/A
			Sd-835573	+	N/A	N/A	N/A
HN4 N line			Sd-835574	-	N/A	N/A	N/A
			Sd-835575	-	N/A	N/A	N/A
			Sd-835576	-	N/A	N/A	N/A
			Sd-835577	-	N/A	N/A	N/A
			Sd-835578	-	N/A	N/A	N/A
			Sd-835579	+	N/A	N/A	N/A
			Sd-835580	-	N/A	N/A	N/A
			Sd-835581	-	N/A	N/A	N/A
HM6 M line			Sd-835582	-	-	-	-
			Sd-835583	-	-	-	-
			Sd-835584	-	-	-	-
			Sd-835585	-	-	-	-
			Sd-835586	-	-	-	-
			Sd-835587	-	N/A	N/A	N/A
			Sd-835588	-	N/A	N/A	N/A
			Sd-835589	-	N/A	N/A	N/A

Table 19 c). Genotyping results for first batch of F1s, the high T forward screen line S-114139 and GSK commercial lines HT6, HN4 and HM6. N/A indicates where DNA samples were not tested for this SNP because neither of the parents had it. RF indicates reaction failure. Reactions were not repeated as there were sufficient siblings with the desired genotype (shaded) to take forward to F2.

For X027, just one of the ten F1s screened possessed the E193K and R158K markers inherited from the high codeine parent (likely heterozygous for these mutations; Table 19b). Nevertheless, this plant produced viable seed to progress this cross to F2.

All mutations were assumed to be in the heterozygous state and segregation in the F2 would enable recovery of homozygotes. Figure 27 details the possible scenarios involved in the recovery of genotypes of interest in F2. If the genes involved are unlinked 1 in every 16 F2s would be double mutants. If the genes are linked a greater number of F2s would be required to recover double mutants. It was decided to sow approximately 25 F2s for each cross that was taken forward due to limitations with greenhouse space. Figure 27 also explains how 6 in every 16 F2s would end up homozygous for one mutation and heterozygous for the other. In this case, one more selfing would lead to recovery of double mutants in F3s (1 in 4).

5.4 F2s with desirable genotypes were identified using AS-PCR and sequencing

Seed collected from F1s with confirmed genotypes was used to generate the F2 individuals. DNA was extracted from F2s after four weeks and screened for the relevant CODM and T6ODM polymorphisms as before (visualisation with the Fragment Analyser). In the second step, the zygosity of each mutant was determined by TILLING the relevant gene in each plant. Finally, the genotypes of those individuals estimated to be homozygotes were confirmed by sequencing the portion of the gene with the mutation. As well as F2s, back-cross 1 individuals (BC1) were genotyped. In these cases, the F1 was again crossed to a recurrent parent. Genotyping of BC1s would determine which individuals still possessed the CODM and T6ODM SNPs introduced by the original mutant parent (only half of the gametes transferred by the F1 would contain the SNPs). Appendix G details the genotyping of F2 and BC1 individuals. Genotypes confirmed by sequencing are marked in bold. If appropriate, these individual plants were allowed to self to produce F3 seed for further analysis.

5.5 Selection of material for Tasmanian field trials

Approximately 600 F2s and BC1s were genotyped in the first instance. At this time the crosses involving T6ODMa TILLING mutants were at the F1 stage and it was not possible to produce F3 seed (producing plants with stable genotypes and phenotypes) in time for the

2014/2015 growing season in Tasmania. However, F3 seed arising from the crosses detailed in Table 19 was shipped to Tasmania for field trials. The seed was collected from F2 individuals with desirable genotypes (Appendix G; D-10858 to D-10865).

Seed batches were sown alongside controls for morphine (HM6 and HM7), thebaine (HT5 and HT6) and noscapine (HN1). Controls were randomly sown in duplicate within the same plot (Appendix H). HM2, the morphine line from which the mutant parents were originally derived, was sown in plots on an adjacent block in the same field. Each plot contained a row of approximately 20 plants. In a number of cases the F2s that produced the F3 seed had one mutation in the heterozygous state and so the resulting F3s would be segregating e.g. Q254* x HT1, Q254* x HN1, X008, X009, X042 and X051 (Table 20). For these lines DNA was requested from all individuals for genotyping. For the other lines, DNA would be extracted from a smaller number of individuals for QC genotyping to confirm expected genotypes (Table 20). Appendix H details the genotyping of field grown material carried out once DNA was shipped back to York.

As the CODM E193K and R158K polymorphism were always inherited together in previous generations, only the R158K mutation was screened for. If a plant was homozygous for R158K it was assumed to be homozygous for E193K also.

Original cross Generation	Original Mother	Original Father	Seed batch id	Expected genotypes	Expected phenotypes
Q254* x HN1 F3	CODMb Q254*	HN1	S-186637	Q254* homozygous, N cluster potentially segregating	C+N
Q254* x HN1 F3	CODMb Q254*	HN1	S-186640	Q254* heterozygous, N cluster potentially segregating	C+N
Q254*x HT1 F3	CODMb Q254*	Tint	S-186628	Q254* segregating in high T background (Tint homozygous)	C+T
Q254*x HT1 F3	CODMb Q254*	Tint	S-186632	Q254* homozygous, T6ODM Tint segregating	C+T
X001 F3	HM6	CODM E193K/R158K	S-186617	CODM mutations E193K/R158K homozygous in high M background	C
X001 F3	HM6	CODM E193K/R158K	S-186729	CODM mutations E193K/R158K homozygous in high M background	C
X001 F3	HM6	CODM E193K/R158K	S-186728	CODM mutations E193K/R158K homozygous in high Mbackground	C
X008 F3	HM6	CODMb Q254*	S-186601	CODMb Q254* homozygous in HM6 background	C
X009 F3	CODMb Q254*	CODM E193K/R158K	S-186609	segregating CODM mutations	C
X009 F3	CODMb Q254*	CODM E193K/R158K	S-186671	segregating CODM mutations	C
X014 F3	High T fwd (no markers)	CODMb F271L Tint	S-186607	homozygous for T6ODM Tint polymorphism	T+O
X016 F3	HN4	CODMb F271L Tint	S-186619	homozygous for T6ODM Tint in N background	T+N
X016 F3	HN4	CODMb F271L Tint	S-186620	homozygous for T6ODM Tint in N background	T+N
X016 F3	HN4	CODMb F271L Tint	S-186621	homozygous for T6ODM Tint in N background	T+N
X016 F3	HN4	CODMb F271L Tint	S-186622	homozygous for T6ODM Tint in N background	T+N
X016 F3	HN4	CODMb F271L Tint	S-186624	homozygous for T6ODM Tint in N background	T+N
X023 F2	HN4	T/O only from HN3 mutagenesis	S-186682	No markers	T+O+N
X027 F3	HM6	CODM E193K/R158K	S-186658	CODM mutations E193K/R158K homozygous in high morphine background	C
X027 F3	HM6	CODM E193K/R158K	S-186659	CODM mutations E193K/R158K homozygous in high morphine background	C
X027 F3	HM6	CODM E193K/R158K	S-186660	CODM mutations E193K/R158K homozygous in high morphine background	C
X029 F2	HN4	T/O only from HN3 mutagenesis	S-186685	No markers	T+O+N
X041 F3	CODMb F271L Tint	T/O only from HN3 mutagenesis	S-186667	F2 M/T/O only and homozygous for T6ODM Tint	T+O
X041 F3	CODMb F271L Tint	T/O only from HN3 mutagenesis	S-186662	F2 T/O only and homozygous for T6ODM Tint	T+O
X041 F3	CODMb F271L Tint	T/O only from HN3 mutagenesis	S-186665	F2 T/O only and homozygous for T6ODM Tint	T+O
X042 F3	Tint	High T fwd (no markers)	S-186610	T6ODM Tint homozygous in high T forward screen line background	T
X042 F3	Tint	High T fwd (no markers)	S-186611	T6ODM Tint homozygous in high T forward screen line background	T
X042 F3	Tint	High T fwd (no markers)	S-186613	T6ODM Tint homozygous in high T forward screen line background	T
X042 F3	Tint	High T fwd (no markers)	S-186716	T6ODM Tint homozygous in high T forward screen line background	T
X042 F3	Tint	High T fwd (no markers)	S-186719	T6ODM Tint segregating in high T forward screen line background	T
X042 F3	Tint	High T fwd (no markers)	S-186720	T6ODM Tint segregating in high T forward screen line background	T
X042 F3	Tint	High T fwd (no markers)	S-186717	T6ODM Tint segregating in high T forward screen line background	T
X051 F3	CODMb F271L Tint	CODM E193K/R158K	S-186608	homozygous for T6ODM Tint and CODM E193K/R158K	C+T
X051 F3	CODMb F271L Tint	CODM E193K/R158K	S-186678	T6ODM Tint segregating in a background homozygous for CODM E193K/R158K	C+T
X051 BC	CODMb F271L Tint	CODM E193K/R158K	S-186605	CODM E193K/R158K segregating in a background homozygous for the T6ODM Tint	C+T

Table 20. Material sent to Tasmania for field trials in 2014/2015 growing season. M=morphine, C=codeine, O=oripavine, T=thebaine

5.6 Towards a high codeine line

Worldwide demand for codeine far exceeds the amount extracted from plants. As discussed in section 1.7.1 codeine is obtained from morphine through the process of O-methylation, a time consuming and expensive process that produces toxic byproducts and waste that needs to disposed of safely. A plant with the morphinan pathway arrested at CODM would allow the direct recovery of codeine from the plant. It is not clear whether total alkaloid yield would be affected in such a plant but any reduction in yield and therefore commercial viability would be balanced by the current costs faced by the industry with the production of codeine from morphine crops.

EMS-induced CODM mutations were introduced to the current commercial morphine variety HM6 to begin the process of removing other potentially deleterious mutations. As this is process is likely to take a longer time than permitted for this studentship, homozygous mutants were recovered in F2 to demonstrate the effects of the CODM mutations on codeine content in the capsules.

5.6.1 X008-introducing the CODMb Q254* mutation to HM6

Mother Seed batch id Seedling id Known mutations	Father Seed batch id Seedling id Known mutations	F1 seed batch id Cross id	F1 seedling id	F2 seed batch id	F2 seedling id Genotype	F3 seed batch id
HM6 Sd-833541 M line	S-122112 Sd-833508 **CODMb Q254***	S-123958 X008	Sd-835319	S-122988	**Sd-837535** Q254* homo	**S-186601**

Table 21. X008 description and followup

11 of the 34 F2s were genotyped as Q254* homozygotes (Appendix G, D-10858). Two were confirmed to have this genotype by sequencing, including Sd-837535 which, when selfed, produced S-186601 seed. This F3 seed was sown in Tasmania and it produced plants with on average 0.46% dw codeine (10.3% of total alkaloids) in the capsules (Figure 50). This is over double the amount of codeine observed in the commercial morphine varieties HM2 (0.18% dw; 3.6% of total alkaloids) and HM6 (0.19% dw; 4.0% of total alkaloids). An F4 seed batch sown in the glasshouse in York (S-188301) also gave rise to Q254* homozygote plants with on average 0.39% dw codeine (11.1% of total alkaloids; Appendix I, X008). The increase

in codeine content suggests a partial block of metabolism at codeine, caused by the plants having one copy of CODM knocked out.

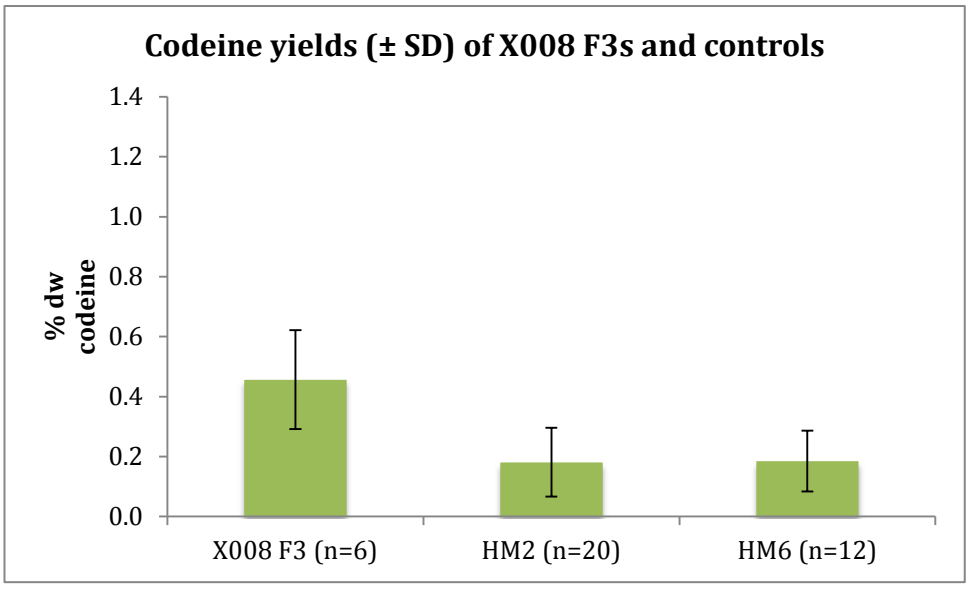


Figure 50. Increased codeine in CODMb Q254* homozygotes. F3s homozygous for the mutation have twice as much codeine in the capsules as HM2 (the line in which the mutation was induced) and HM6 (a more recent commercial morphine line into which the mutation was introduced in X008).

5.6.2 X001 & X027-introducing the CODMa/c E193K and CODMb R158K mutations to HM6

Mother Seed batch id Seedling id Known mutations	Father Seed batch id Seedling id Known mutations	F1 seed batch id Cross id	F1 seedling id	F2 seed batch id	F2 seedling id Genotype	F3 seed batch id
HM6 Sd-833543 M line	S-114248 Sd-833401 **CODMa/c E193K** **CODMb R158K**	S-123957 X001	Sd-835406	S-122996	**Sd-837876** **Sd-832178** **Sd-832179** E193K/R158K homo	**S-186617** **S-186729** **S-186728**
HM6 Sd-833559 M line	S-122066 Sd-833434 **CODMa/c E193K** **CODMb R158K**	S-123955 X027	Sd-835413	S-122982	**Sd-837766** **Sd-837779** **Sd-837784** E193K/R158K homo	**S-186658** **S-186659** **S-186660**

Table 22. X001 & X027 description and followup

4 of the 41 F2s derived from the X001 F2 S-122996 were genotyped as homozygous for the CODM E193K and R158K mutations (Appendix G, D-10861 & D-10865). Three of these produced F3 seed for Tasmanian field trials. Likewise, 3 of the 25 F2s derived from X027 F2 S-122982 were homozygous for both mutations (Appendix G, D-10861).

An important point to note is that the two polymorphisms were always inherited together i.e. if a plant was heterozygous for one it was also heterozygous for the other. Furthermore, X009 (Table 19), which attempted to replace the CODMb R158K mutations with the CODMb nonsense mutation (Q254*), failed to achieve the desired result in both F2 and F3 generations. These observations lend further support to the suggestion that copies of

CODM are physically linked. Nevertheless, the greater contribution of the E193K/R158K mutations over the Q254* mutation to a high codeine phenotype was demonstrated among the F3s in which these mutations were segregating (Figure 51). The Q254* homozygotes had on average 0.61% dw codeine (11.1% of total alkaloids) while the E193K/R158K homozygotes had on average 1.01% dw codeine (17.0% of total alkaloids).

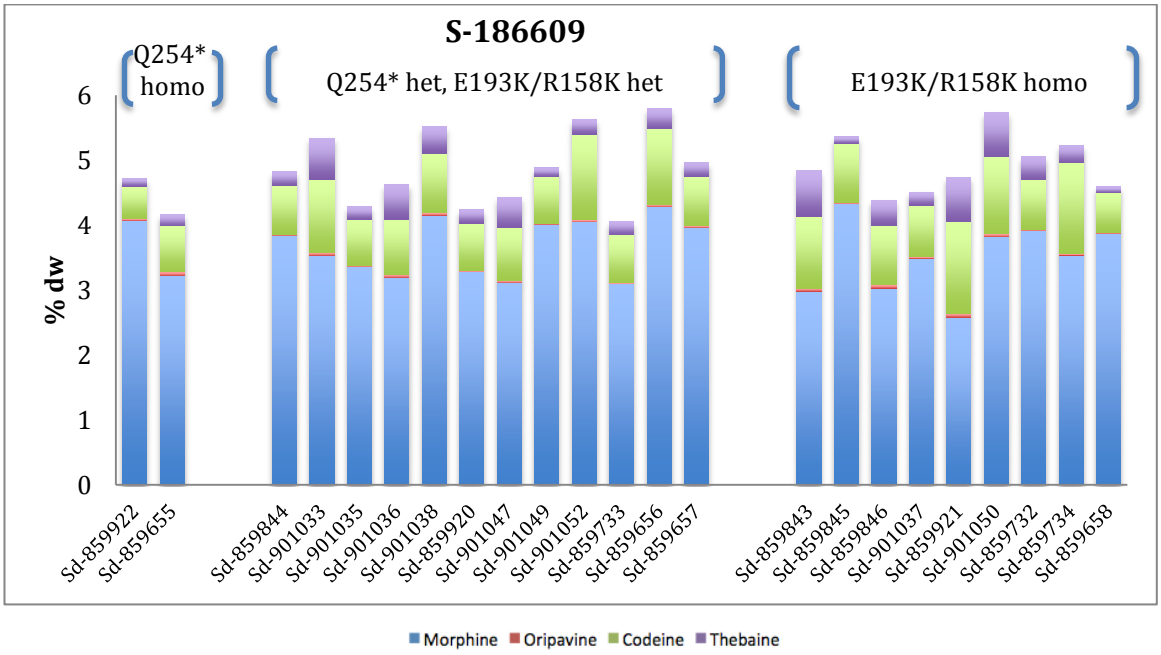


Figure 51. X009 F3s demonstrating the contribution of the E193K/R158K polymorphisms to a codeine phenotype.

Similar yields of codeine were observed in X001 and X027 F3s genotyped as E193K/R158K homozygotes (Figure 52 and Appendix I, X001 & X027). The largest number of E193K/R158K homozygote F3s genotyped and phenotyped came from S-186658. These had consistently high codeine, a 5-6 fold improvement over the yield of codeine from HM2 (the line in which the mutations were induced) and HM6 (the line into which the mutations were transferred). Again, a complete block at codeine would require disruptions to all actively expressed copies of CODM. However, plants which possess these two CODM mutations clearly are associated with high codeine phenotypes. It may be useful to continue backcrossing these lines to HM6 to reduce the number of background mutations and ensure the resulting codeine line will have the same agronomic traits as HM6 but with a higher yield of codeine at the expense of morphine.

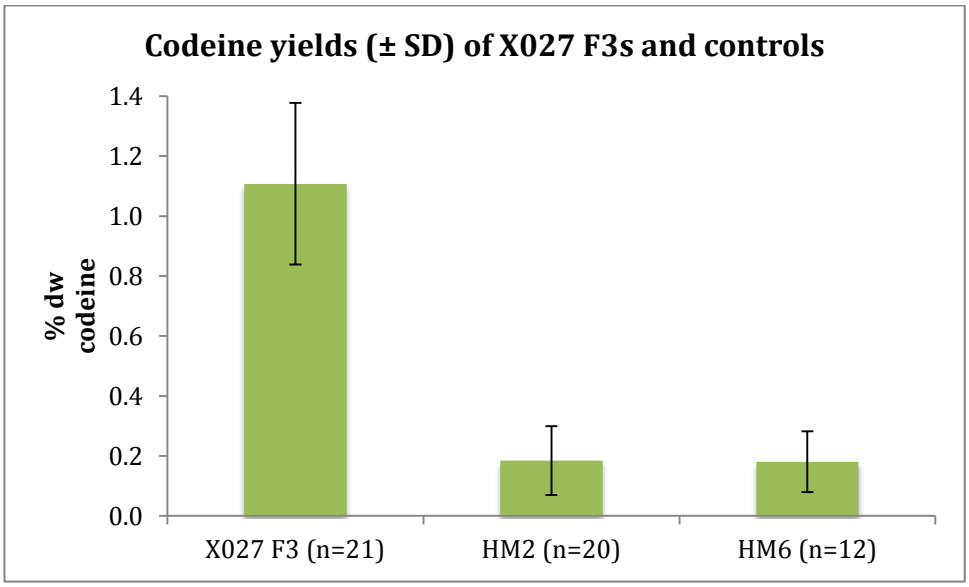


Figure 52. Increased yields of codeine in plants from X027 F3 seed batch S-186658. All F3s were homozygous for the CODM E193K/R158K mutations and showed significantly higher codeine (mean 1.1%) than commercial morphine lines, HM2 (0.19%) and HM6 (0.18%).

5.6.2.1 HM7 GSK line

One of the controls grown alongside the F3 material in Tasmania was GSK variety, HM7, into which the CODMa/c E193K and CODMb R158K mutations were introduced (Appendix H; D-11114 B2-E2, D-11115 G12-H12, D-11116 A1-B1). Codeine content among the seven HM7 plants phenotyped averaged at 0.81% dw (15% of total alkaloids).

A larger trial of this line last year also demonstrated a higher proportion of codeine in this line compared to e.g. HM6. However, the total alkaloid content is approximately 60% of current commercial morphine lines and so it would not be commercially viable at present. The high codeine forward screen lines came from a heavily mutagenized M2 individual which underwent numerous rounds of selfing (Appendix B). The high codeine forward screen seed had a relatively poor germination rate and produced short plants which flowered early. They were clearly stressed plants with a high mutation burden. The creation of HM7, X001 and X027 represent the first backcrosses into elite material. Further backcrosses aided by tracking of the polymorphisms for high codeine will reduce the number of background mutations while maintaining the desirable trait. In the end the CODM mutations can be brought to homozygosity and the result should be a line with a high total alkaloid content, a large proportion of which would be codeine. The results from X027 demonstrate what can be achieved with just one backcross to HM6. There are also QTLs for numerous other agronomic

traits (developed in another arm of the project), which could be incorporated into the marker assisted breeding strategy. The end result would be a high codeine line with all the agronomic traits of the recurrent parent.

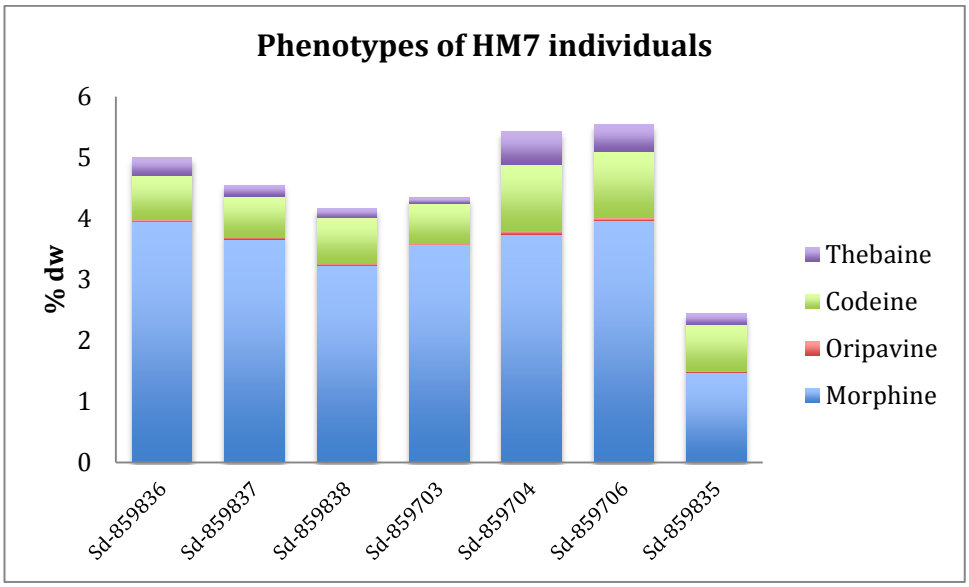


Figure 53. Phenotypes of HM7 individuals grown in the field in Tasmania alongside F3s. HM7 is a candidate commercial codeine line.

5.6.3 Further mutagenesis of bulked up Q254 material*

When crosses between CODM Q254* and W261* mutants failed to result in a double mutant with a more substantial block of metabolism at codeine and it became increasingly likely that copies of CODM may be linked, an alternative strategy for obtaining a CODM double mutant was required. Seed of the CODM mutant with the clearest change in phenotype, Q254*, was bulked up in the M4 generation. M5 seed was then subjected to another round of EMS mutagenesis (Table 2 & Figure 54) in an attempt to introduce mutations in other copies of CODM (or any other genes which when disrupted would result in increased codeine).

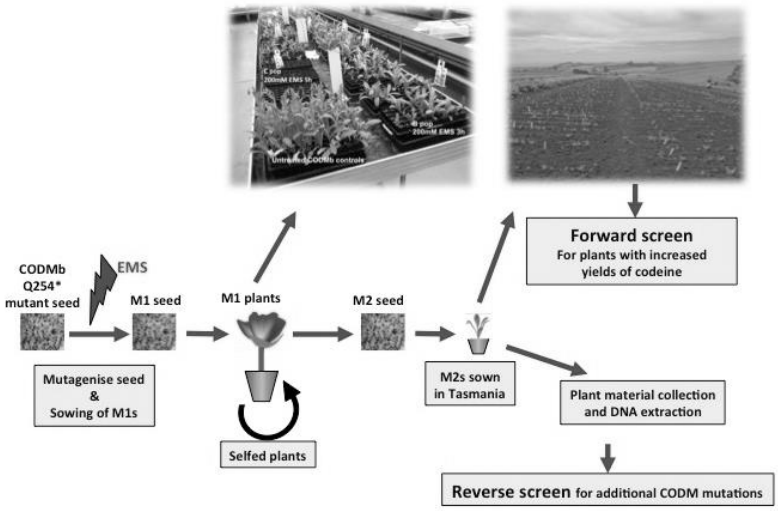


Figure 54. Strategy for introducing new mutations in CODMb Q254* material. M5 seed was mutagenised in April 2014. M1s were grown in the greenhouse in York over the summer before M2 seed was shipped to Tasmania for the 2014/2015 growing season. DNA has been collected from M2 individuals which could be used for a reverse genetic screen.

M1 seedlings were grown in the glasshouse in York and selfed manually to ensure M2 seed production. M2 seed was sown in Tasmania in a number of different blocks within the field trial site. A row (plot) of plants was generated with each seed batch and mature capsules from four of these siblings were assessed for their alkaloid content at GSK's Port Fairy site. As with the original TILLING population, it was anticipated that any induced mutations would likely be in the heterozygous state at this stage. Therefore, some sibling variation was expected and the aim was to identify lines with increased codeine content that have the potential to be replicated in the M3 generation.

As well as M2 seed from the B and C subpopulations, various controls (Table 23) were grown in each block (HM2, HM7, HT7) and among the blocks (non-EMS treated CODMb Q254* M5s). HM2 was the morphine line originally subjected to EMS mutagenesis resulting in CODM Q254*, E193K and R158K mutations. HM7 is a candidate codeine line in which the CODM E193K/R158K mutations were detected (section 5.6.2.1). HT7 has been bred by GSK for high thebaine and oripavine and little morphine and codeine are present in mature capsules. The three remaining controls are the Q254* M5 seed used for mutagenesis. The strategy was to identify M2 lines with codeine content approaching 1% or higher which could be the result of newly introduced mutations and the Q254* mutation already present.

Control/ Seed batch id	Block	GSK plot id Mean codeine yield (% dw)	Description
HM2	15	MD 1524 & MD 1574 0.27 (n=7)	Morphine line subject of original EMS mutagenesis.
	16	MD 1634 & MD 1687 0.17 (n=6)	
HM7	15	MD 1578 & MD 1596 0.61 (n=7)	Candidate codeine line with EMS induced CODM E193K/R158K polymorphisms
	16	MD 1623 & MD 1657 0.60 (n=7)	
HT7	15	MD 1553 & MD 1585 0.05 (n=7)	Thebaine/Oripavine line
	16	MD 1611 & MD 1669 0.05 (n=8)	
S-123173	16	MD 1640 0.25 (n=4)	CODMb Q254* M5 seed used to make the new B population
	20	MD 2024 0.46 (n=4)	
S-124701	19	MD 1937 0.55 (n=3)	CODMb Q254* M5 seed used to make the new C population
S-122971	24	MD 2493 0.40 (n=4)	

Table 23. Controls grown alongside the new mutagenised population in field trials in Tasmania. HM2 was the morphine line originally subjected to EMS mutagenesis. The Q254*, E193K and R158K mutations were introduced in this way, polymorphisms which resulted in minor blocks of metabolism at codeine. HM7 is a candidate codeine line in which the CODM E193K/R158K polymorphisms are segregating. Individuals homozygous for the polymorphisms have high codeine levels. HT7 is a contemporary thebaine/oripavine line with little codeine. The three remaining controls represent Q254* M5 seed which was not subjected to EMS mutagenesis earlier in the year. Any M2 line with codeine content greater than 0.6% would be an improvement on the Q254* M5 material and HM7. All of the yields quoted were determined in field trials that were also the source of the material in Table 24.

Table 24 shows the top 25 M2s with regard to yields of codeine (that successfully produced M3 seed). Many are approaching or exceeding 1% codeine, which if replicated in the M3 generation, would be an improvement over the CODMb Q254* M5s before treatment with EMS and HM7, the candidate codeine line.

Mutation detection efforts should focus on these 25 M2 lines in the first instance. Sowing the M3s would indicate whether the phenotypes are stable and if any *de novo* mutations are found then markers for the new mutation and Q254* could be used to begin backcrosses to elite material. These lines are the result of two rounds of EMS mutagenesis and the extremely low germination rates (Table 2) are testament to the burden of background mutations that they carry.

M2 seed batch id	Block Pop	GSK plot id Mean codeine yield (% dw)	Individual capsules with highest codeine yield* (% dw)	Resulting M3 seed batch id
S-186932	15 B	MD 1581 0.67	MD 1581-1 0.91	S-130267
S-186894	15 B	MD 1540 0.64	MDS 1540-3 0.87	S-130243
S-186878	15 B	MD 1523 0.67	MDS 1523-4 0.83	S-141377
S-186883	15 B	MD 1529 0.58	MDS 1529-2 0.74	S-141396
S-187016	16 B	MD 1676 0.80	**MDS 1676-1 2.02**	TBD
S-186990	16 B	MD 1646 0.88	**MDS 1646-1 1.23**	TBD
S-187011	16 B	MD 1671 0.57	**MDS 1671-2 1.1**	TBD
S-186981	16 B	MD 1636 0.86	MDS 1636-2 0.99	TBD
S-187012	16 B	MD 1672 0.75	MDS 1672-2 0.92	TBD
S-186993	16 B	MD 1650 0.64	MDS 1650-2 0.91	TBD
S-186971	16 B	MD 1624 0.7	MDS 1624-4 0.89	TBD
S-187063	19 B	MD 1928 0.89	MDS 1928-1 1.04	S-130702
S-187141	20 B	MD 2006 0.58	MDS 2006-2 1.12	S-130904
S-187206	20 B	MD 2072 0.81	MDS 2072-3 0.94	S-131119
S-187133	20 B	MD 2001 0.79	MDS 2001-4 0.91	S-132069
S-187315	23 B	MD 2370 0.84	MDS 2370-1 0.93	S-131782
S-187317	23 C	MD 2375 0.44	MDS 2375-1 0.96 (0.79M+1.06T)	S-131797
S-187320	23 C	MD 2378 0.36	MDS 2378-2 0.89	S-131810
S-187310	23 B	MD 2366 0.71	MDS 2366-2 0.88	S-131777
S-187330	23 C	MD 2389 0.84	MDS 2389-1 0.84	S-131500
S-187372	24 C	MD 2435 0.89	MDS 2435-2 0.99	S-131502
S-187457	28 C	MD 2822 0.7	MDS 2822-3 0.92	S-131724
S-187463	28 C	MD 2830 0.67	MDS 2830-2 0.91	S-131749
S-187483	28 C	MD 2855 1.03	MDS 2855-4 0.79	S-132234
S-187491	28 C	MD 2864 0.89	MDS 2864-1 0.71	S-132264

Table 24. Top 25 codeine yields (% dw) of M2s which successfully produced M3 seed. M2 seed was sown in plots (MD numbers) and four capsules were assessed for alkaloid content. The highest yields of codeine (marked in bold) were present in block 16.

5.7 Towards a high thebaine line

5.7.1 X051-cross between high codeine and high thebaine forward screen lines

Mother Seed batch id Seedling id Known mutations	Father Seed batch id Seedling id Known mutations	F1 seed batch id Cross id	F1 Seedling id	F2 seed batch id	F2 Seedling id Genotype	F3 seed batch id
S-114139 Sd-834053 **CODMb F271L** **T6ODMa intron SNP**	S-114156 Sd-834078 **CODMa/c E193K** **CODMb R158K**	S-124013 X051	Sd-835330	S-122983	**Sd-837576** E193K/R158K homo Tint homo **Sd-832031** E193K/R158K homo Tint het **Sd-832040** E193K/R158K het Tint homo	**S-186608** **S-186678** 1in4 F3s should be homo for all 3 SNPs **S-186605** 1in4 F3s should be homo for all 3 SNPs

Table 25. X051 description and followup.

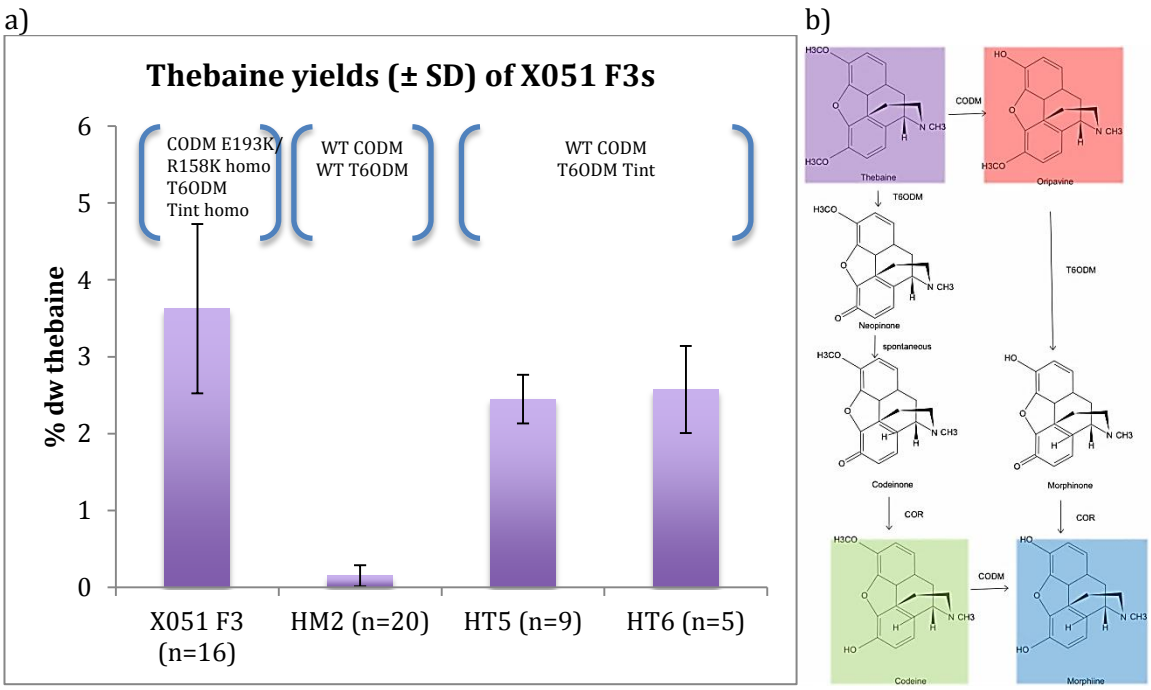


Figure 55. a)Thebaine yields of X051 F3s (from seed batch S-186608) homozygous for both the CODM E193K and R158K mutations and the Tint polymorphism. b) Pathways to morphine from thebaine. The minor route is thought to proceed via oripavine and morphinone.

Just 1 of the 37 F2s genotyped were homozygous for all three SNPs-CODMa/c E193K, CODMb R158K and T6ODMa Tint. Seed was collected from this individual and resulting F3s had significant increases in the yields of thebaine (mean 3.62% dw; 54.4% of total alkaloids). This is a not just an improvement over HM2, into which the mutations were induced, but also current commercial thebaine lines grown alongside the F3s in Tasmania (Figure 55a). Increased yields of thebaine may be attributed disruption of thebaine to oripavine reactions in the minor route from thebaine to morphine (Figure 55b). This may complement a block at

T6ODM activity on thebaine, linked to the Tint polymorphism present in commercial thebaine lines and the high thebaine forward screen line used as a parent in this cross.

The impact of segregating CODM E193K/R158K and T6ODMa Tint on F3 phenotypes is demonstrated in Appendix I (X051). The results show the usefulness of tracking these SNPs in crosses and the predictable impact on phenotypes. The CODM mutations, always inherited together, clearly increase codeine content- at the expense of morphine in X001 and X027 but at the expense of oripavine in X051. Likewise the impact of the Tint polymorphism (or the unknown T6ODM mutation it is linked to) has a clear effect on thebaine accumulation in capsules. Those without the polymorphism had on average 0.23% dw thebaine (4.7% of total alkaloids). Thebaine is increased in both heterozygotes (mean 1.29% dw; 25.0% of total alkaloids) and homozygotes (2.71% dw; 50.1% of total alkaloids).

5.8 Towards a high oripavine line

Oripavine is a much sought after alkaloid due to its use in the manufacture of semi-synthetic drugs such as for naloxone and naltrexone, for which the market continues to grow. A number of crosses were attempted to boost yields of oripavine. Current commercial thebaine lines HT5 and HT6 had oripavine yields of 0.75% dw (14.8% of total alkaloids) and 0.54% dw (11.0% of total alkaloids) when grown alongside the F3s in the 2014/2015 growing season. The thebaine/oripavine only lines derived from HN3 mutagenesis (Figure 49) may have mutations in a regulator of T6ODM or a transporter responsible for transport of thebaine between sieve elements and laticifers.

5.8.1 X016-Attempt to introduce noscapine into a high thebaine forward line resulted in a high oripavine line

Mother Seed batch id Seedling id Known mutations	Father Seed batch id Seedling id Known mutations	F1 seed batch id Cross id	F1 Seedling id	F2 seed batch id	F2 Seedling id Genotype	F3 seed batch id
HN4 Sd-833580 N line	S-114249 Sd-833464 **CODMb F271L** **T6ODMa intron SNP**	S-123972 X016	Sd-835442	S-122995	8 of 36 F2s were Tint homo including **Sd-837722** (2.6%N), **837723** (3.3%N), **837726** (4.1%N), **837728** (2.9%N), **837734** (2.3%N)	**S-186619, 186620, 186621, 186622, 186624**

Table 26. X016 description and followup.

X016 was set up initially in an attempt to introduce noscapine into the high thebaine forward screen line derived from HM2 mutagenesis. Seed from five of the F2s identified as being homozygous for the Tint polymorphism was shipped to Tasmania for field trials (Table 20). All of these F2s had relatively small amount of noscapine in their capsules indicating that the noscapine gene cluster was likely in a hemizygous state. Therefore segregation of the noscapine gene cluster in the F3s in Tasmania was expected to result in plants with circa 50% noscapine/50% morphinans, similar phenotypes to F2s, or plants with no noscapine in their capsules. For four of the five F3 seed batches this was the case (Appendix I) but the capsules of plants derived from S-186622 had no noscapine or morphine and had on average 1.52% dw oripavine (Figure 56; 35.6% of total alkaloids) which is a 2-3 fold improvement on HT5 and HT6 controls-oripavine yields of 0.75% dw (14.8% of total alkaloids) and 0.54% dw (11.0% of total alkaloids), respectively. These plants also lacked the Tint polymorphism so the block at thebaine and oripavine must be caused by additional mutation contributed by the high thebaine forward screen line that then segregates independently in the F2. The alkaloid profiles are similar to those of the thebaine/oripavine only lines derived from HN3 mutagenesis (Figure 49).

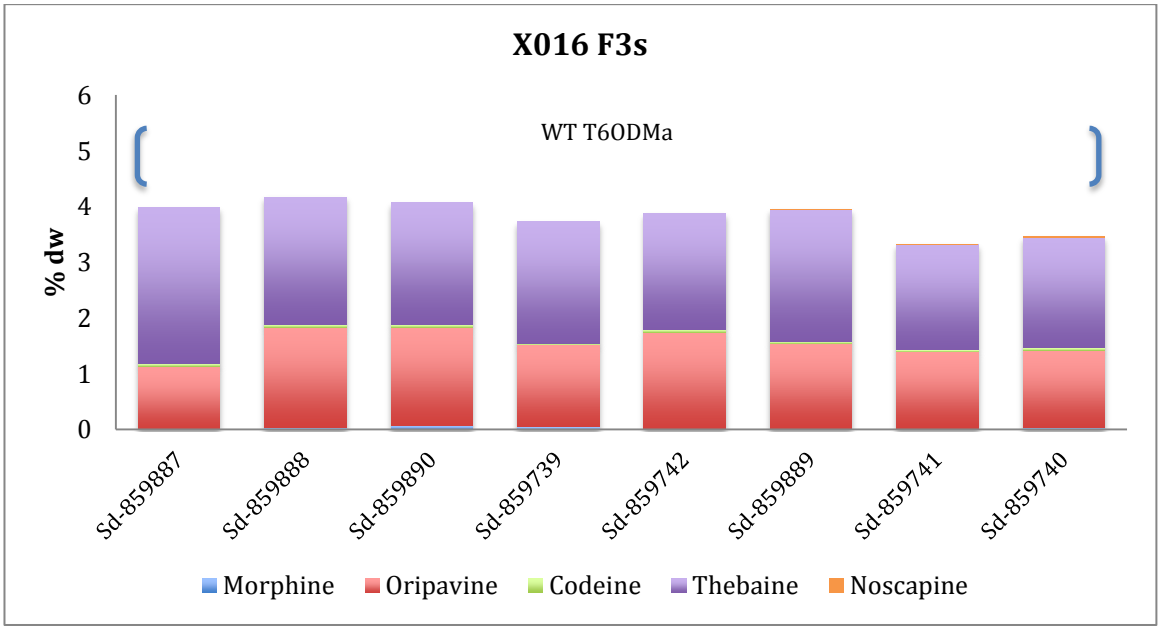


Figure 56. Alkaloid profiles of X016 F3s-noscapine and morphine free plants with high oripavine content.

5.8.2 X041-cross between two high thebaine forward screen lines

Mother Seed batch id Seedling id Known mutations	Father Seed batch id Seedling id Known mutations	F1 seed batch id Cross id	F1 Seedling id	F2 seed batch id	F2 Seedling id Genotype	F3 seed batch id
S-114139 Sd-833680 **CODMb F271L** **T6ODMa intron SNP**	S-125001 Sd-833604 T/O line	S-124064 X041	Sd-835453	S-122994	11 of 12 F2s genotyped as Tint homo 2 had gentoypes confirmed including **Sd-837703**	**S-186667** **Others sown** **S-186661, S-186664, S-186665,** **S-186740, S-186742**

Table 27. X041 description and followup.

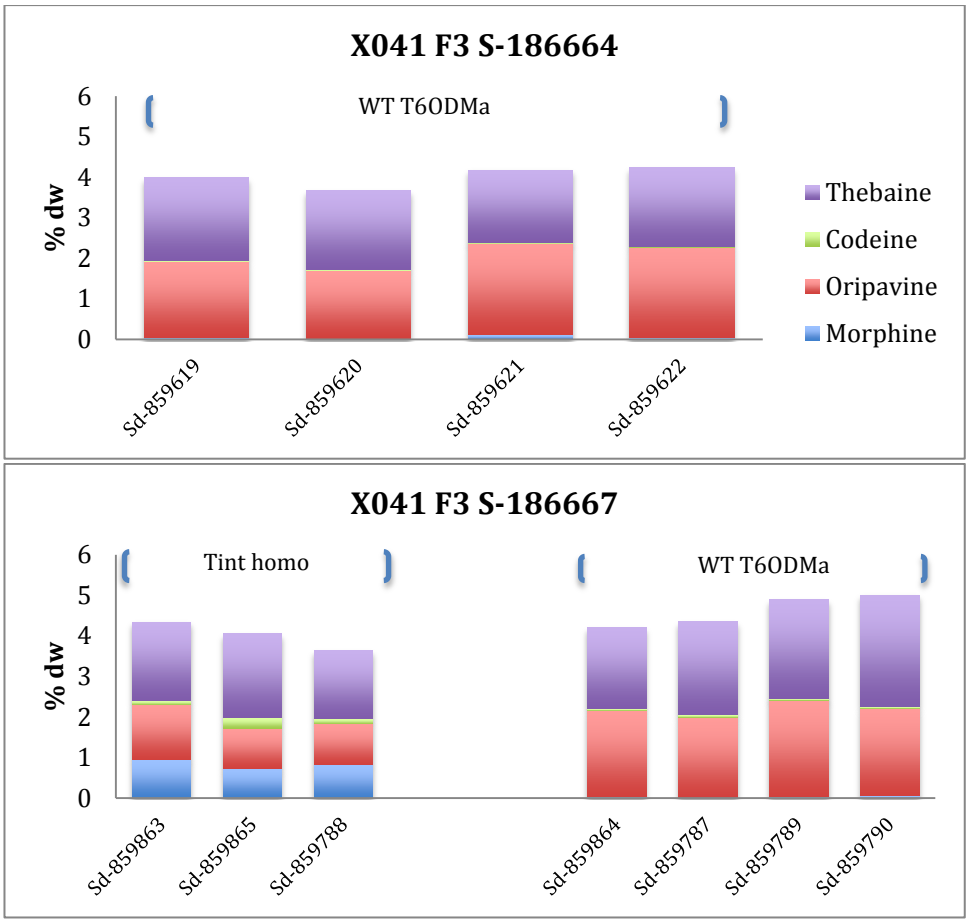


Figure 57. High oripavine yields among X041 F3s. The Tint polymorphism was observed to segregated independently of the thebaine/oripavine phenotype. Absence of the Tint polymorphism was an indicator of a metabolic block towards morphine and a high yield of thebaine and oripavine. Plants homozygous for Tint were able to produce morphine suggesting only a partial block of metabolism at thebaine.

The high thebaine forward screen lines derived from both HM2 and HN3 mutagenesis were crossed together in an effort to combine whatever mutations cause their phenotypes and further increase yields of thebaine and oripavine. Among the F3s there were a number of plants with more oripavine than thebaine in their capsules. The parental line derived from HN3 mutagenesis has approximately a 60-40% split in favour of thebaine. Some of the F3s, however, flipped this ratio in favour of oripavine with yields of over 2% dw in some cases (Figure 57). It is postulated that the T6ODM Tint polymorphism alone is not responsible for high thebaine/oripavine in the High T forward screen line used as parent in X041. There

could be an as yet undetected independently segregating mutation that contributes to the phenotype and when this was combined with the mutation causing the thebaine/oripavine only phenotype in the forward line derived from HN3 mutagenesis the result was high oripavine plants.

5.9 Towards a morphine free noscapine line

Mother Seed batch id Seedling id Known mutations	Father Seed batch id Seedling id Known mutations	F1 seed batch id Cross id	F1 Seedling id	F2 seed batch id
HN4 Sd-833583 N line	S-125002 Sd-833842 T/O line	S-123977 X029	Sd-831918 Sd-831913	S-186685 (Tas) S-186686 (York)

Table 28. X029 description and followup.

An important final objective was to attempt to create a noscapine/thebaine variety. GSK and its partners have found it possible to separate these alkaloids in the industrial extraction process. In the field trials control noscapine line HN4 had on average 1.64% morphine and 2.16% noscapine in mature capsules with trace amounts of thebaine, codeine and oripavine. Currently one way to increase thebaine content of noscapine lines is to use a chemical spray containing the acylcyclohexanedione trinexapac-ethyl (Moddus). Introduction of mutations preventing flux from thebaine to morphine into noscapine lines should increase thebaine/oripavine levels and may possibly avoid the need for spraying with Moddus.

A step towards this was taken with one of the F3s from X016 described previously (Appendix I, S-186620) with a phenotype of 0.25% oripavine, 1.06% thebaine, 1.55% morphine and 1.93% noscapine. The noscapine cluster in this plant is restored to the homozygous state and it also contains the Tint polymorphism which may cause or be linked to the cause of the partial block of metabolism at thebaine.

However, X029 provided an even better phenotype in field trials of F2 individuals. The parents in this case were HN4 and the thebaine/oripavine only line that came out of HN3 mutagenesis. Among the X029 F2s grown in York (Figure 58) were individuals with the noscapine cluster restored to display a phenotype like HN4, some heterozygotes with a small

amount of noscapine and a third group comprising of only one morphine-free individual with 0.99% oripavine, 1.85% thebaine and 1.93% noscapine. Seed from this plant should be sown to ensure that the phenotype is reproducible.

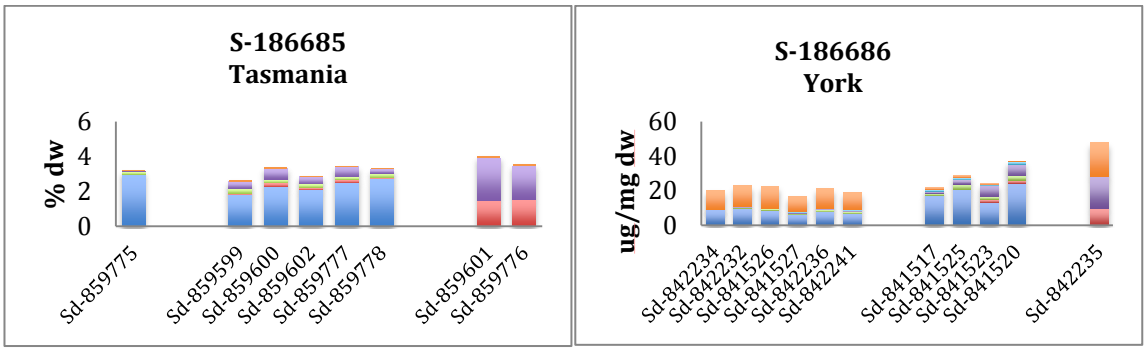


Figure 58. Alkaloid profiles of X029 F2s sown in both the field in Tasmania and the glasshouse in York. No markers were available for this cross. However plants with circa half of their alkaloid content being noscapine are interpreted as being homozygous for the noscapine gene cluster.

Two of the F2s in Tasmania had a thebaine/oripavine phenotype similar to one of the original parents. The noscapine cluster must be in a heterozygous state in these individuals and so approximately 25% of the F3s should have the desired noscapine/thebaine/oripavine profiles.

5.10 Discussion

Tracking mutations using AS-PCR and visualisation of products with the Fragment Analyser machine was an efficient and cost effective way of tracking mutations through the generations. Because different sized PCR products were generated for each SNP each gel could test 96 plants for the presence of 4-5 mutations. Once a mutation was found the be present, a TILLING assay was used to determine zygosity-PCR products would remain uncut if amplified from homozygotes, while PCR products would be cut into two expected fragments if amplified from heterozygotes. This strategy avoided the use of expensive labelled primers and overcame the challenges presented by multiple copies of the 2-ODD genes, which may have frustrated efforts to assign zygosity in e.g. KASP assays where an additional copy of CODM to the one with the mutation would always provide a wild type signal.

Although a large number of crosses were carried out only a portion were brought through to field trials in Tasmania. The remaining crosses are still at the F1 or F2 stages and the seed could be a resource in future efforts to combine mutant alleles of CODM and/or T6ODM. In particular, crosses involving T6ODM TILLING mutants should be considered for future trials. If possible they should be crossed with high thebaine lines for which no molecular markers are available. Introduction of a T6ODMa knockout mutation may contribute to an even more comprehensive block at thebaine/oripavine.

Steps towards the creation of new high codeine varieties of opium poppy were taken. Introducing the CODMb Q254* TILLING mutation to the commercial morphine line, HM6, resulted in a doubling of codeine content in capsules of F3s (Figure 50). Likewise, introducing the EMS-induced CODM E193K/R158K mutations into HM6 resulted in a more substantial block at codeine with a five fold increase in F3s compared to HM6 (Figure 52). Due to the time limitations of this project these lines were brought to homozygosity to assess the impacts of these alleles in the F3. In the medium to long term, these lines should continue to be backcrossed to HM6 or another suitable recurrent parent and the alleles tracked by AS-PCR. Each step would remove potentially debilitating background mutations and produce a crop with all the agronomic traits exhibited by HM6 but with a higher proportion of codeine. Furthermore, if other alleles of CODM are discovered these could be introduced through crosses and further impact on the disruption of codeine demethylation in the plant. Although some of the M2s in the new TILLING population (Figure 54, Table 24) displayed high yields of codeine, this population has yet to be screened for CODM mutations. This resource would likely be a source of new CODM (and T6ODM) alleles. In the meantime, the M3 seed of the highest codeine yielders should be sown to assess whether the high yields are stable and heritable. DNA from these M3s could be the starting point for a new reverse genetic screen.

The usefulness of tracking the Tint polymorphism in crosses for high thebaine was demonstrated in X051 (Appendix I). The Tint polymorphism was segregating among plants resulting from seed batch S-186678. Three clear classes are visible in the chart-those with low thebaine (wild type T6ODM), moderate thebaine (heterozygotes) and high thebaine

(homozygotes). This cross also resulted in plants with the highest thebaine content with a 1.5 fold increase in thebaine over current commercial thebaine lines HT5 and HT6 (Figure 55). This is presumably the result of the additional block at thebaine (in thebaine to oripavine conversions) contributed by the CODM E193K/R158K mutations. X051 led to lines with much potential. One possible way forward would be to set up crosses to elite commercial thebaine lines and to follow the segregation of the CODM E193K/R158K and T6ODM Tint polymorphisms by genotyping. As these mutations came from mutant lines they are likely to be associated with a high number of background mutations. Recurrent backcrosses will remove potentially deleterious mutations while retaining the CODM and T6ODM mutations which when brought to homozygosity at the end of the programme should result in the desired high thebaine phenotype.

Increased yields of oripavine were also achieved in this project. F3s from X016 and X041 displayed a 2-3 fold increase in oripavine content of capsules compared to current commercial thebaine lines HT5 and HT6. Although molecular markers for the high oripavine trait have not yet been identified these lines should be progressed to establish stable high oripavine lines. It is possible that as yet undiscovered transporters of thebaine/oripavine or regulators of T6ODM expression may be affected by EMS mutagenesis in these lines i.e. the T/O only parent derived from HN3 mutagenesis has a complete block of metabolism at T6ODM. Finally, seed from the noscapine-thebaine-oripavine only F2 grown in York should be sown to assess whether this trait is stable and heritable because it would be a useful addition to GSK's portfolio, providing three sought after alkaloids without producing any morphine and codeine which could be produced from other crops.

Chapter 6- General Discussion

EMS mutagenesis of opium poppy seed has proved a successful approach for the generation of alleles of key genes involved in the biosynthesis of important opiates. New lines of opium poppy with desirable alkaloid profiles have been created by means of forward and reverse genetics. The promising CODM alleles produce E193K, R158K and E301K subtitutions and truncations (nonsense mutations). Two nonsense mutations were also identified in T6ODM and an interesting polymorphism in intron 1 of T6ODM was found to be linked to a thebaine phenotype. However, the presence of multiple copies of CODM and T6ODM in the genome of opium poppy has meant that, although knockout mutations in both genes were recovered, complete metabolic blocks at specific alkaloids was not achievable. Despite this the mutant alleles could be used to fine tune metabolic phenotypes. The lack of genome sequence for opium poppy meant that the presence of multiple copies of 2-ODD genes was not known at the beginning of this project. It has now been demonstrated that the plant does indeed have multiple functional copies of CODM and T6ODM. Previous VIGS experiments, demonstrating the effect of silencing CODM and T6ODM, would have silenced all copies simultaneously (Hagel and Facchini, 2010a; Wijekoon and Facchini, 2012). A genome sequence or more sequence information around the site of these 2-ODD genes would have aided primer design for screening additional copies. A single polymorphism between T6ODMa and other copies of T6ODM was exploited to design a primer for the specific amplification of T6ODMa for TILLING (Table 8, Figure 32). It was not possible to do the same for T6ODMb. Sequence information upstream or downstream of the coding sequence would have revealed differences between the copies and enabled primer design for TILLING T6ODMb. To date, no genome sequence for opium poppy is available. This may be due to the fact that it remains quite a niche medicinal crop that has not attracted the interest of large consortia that typically are involved in genome sequencing. Also, opium poppy has a relatively large genome and although the cost of sequencing has come down in recent years the task of assembling the short reads into a genomic sequence is not trivial especially when

you consider the amount of repetitive DNA. The high frequency of repetitive sequences was demonstrated when BACs containing the noscapine gene cluster were sequenced and characterised (Winzer et al., 2012). Different sequencing methods needed to be deployed to achieve a good assembly of even a relatively small amount of genomic DNA (400kb). Of course it is technically feasible to sequence the entire genome but the challenge would more likely be accomplished with adoption of new sequencing technologies which generate read lengths of several thousand base pairs, such as MinION nanopore sequencing (Oxford Nanopore) and PacBio single molecule real time sequencing (Pacific Biosciences). Assembly of larger fragments into a functional genome sequence would then be a much less complex task and therefore the prospect of obtaining a genome sequence for opium poppy is much more feasible. Access to a genome sequence in conjunction with high density genetic maps will aid in the identification of genes involved in BIA biosynthesis and other important agronomic traits. A genome sequence will also allow for studies of the evolutionary relationships between other alkaloid producing plant species and give some insight into how the pathways evolved.

As well as continuing to identify alleles of CODM and T6ODM for incorporation into the breeding programme, rational engineering of BIA metabolism in opium poppy will require a much deeper knowledge of how BIA synthesis is regulated in the plant at the transcriptional, cellular and biochemical levels. Obvious future targets would be transcription factors involved in initiating synthesis of key enzymes and metabolic components. Furthermore, there is still much to be learned about intra-and inter-cellular transport of pathway intermediates. If transporters are involved then regulation of their synthesis or activities would be one way to impact on BIA metabolism. Comparative transcriptomic profiling of morphine accumulating lines and, for example, the thebaine/oripavine only mutants or high thebaine mutants for which no T6ODM mutations have been found but transcript levels are reduced, may be a starting point in identifying putative regulators or transporters which may be affected in high thebaine lines.

Another way to employ marker assisted selection for the improvement of opium poppy crops is to identify resistance markers for downy mildew (*Peronaspora*). An outbreak of downy mildew was first reported in a commercial crop in Tasmania in 1996 (Scott et al., 2004) and the 2014-2015 growing season in Tasmania featured many outbreaks throughout the state with many commonly used fungicides proving ineffective. A genetic study of downy mildew resistance indicated polygenic control (Dubey et al., 2008). Identification of these alleles and their incorporation into commercial crops should be prioritised.

Overall TILLING has proved to be an effective strategy for detecting induced genetic variation. It is recognised as a useful tool when working with neglected or underutilised crops, so-called orphan crops, where access the genetic diversity is a problem for researchers (Jain and Gupta, 2013). EcoTILLING can be deployed to mine for alleles in other varieties but in the case of opium poppy access to genetic diversity in difficult to come by. Nevertheless, this study has demonstrated that utilisation of chemically induced variation is possible in poppy. The TILLING platform used allows for screening large sections of a target gene in a cost effective manner. TILLING using next generation sequencing platforms requires complex barcoding and is limited to read length of the machine used while high resolution melting platforms are limited to 300-400 bp TILLING 'windows'.

An alternative to TILLING would have been to use genetic modification to knockout the desired genes and achieve the required phenotypes. As discussed in section 1.8.1 genetically modified opium poppy crops are not currently used and in Tasmania, where half of the world's opiate based medicines originate, there is a moratorium on the use of GM crops. During the lifetime of this project, the CRISPR/Cas9system, adapted from a bacterial immune system, has come to force as a simple, inexpensive and versatile tool for genome editing (Bortesi and Fischer, 2015). Two years ago five reports were published at the same time detailing the first applications of the CRISPR/Cas9-based genome editing in plants (Feng et al., 2013; Li et al., 2013; Nekrasov et al., 2013; Shan et al., 2013; Xie and Yang, 2013). Genome editing promises to accelerate plant breeding by helping to overcome problems with conventional breeding such as finding and exploiting natural genetic variation. Furthermore,

random mutagenesis programmes such as this one require large-scale screens to identify mutants with desirable properties and an extensive series of backcrosses to remove undesired mutations. With genome editing it is possible to introduce specific and predictable modifications directly into elite material. With the sequence similarity observed between copies of CODM, a guide RNA for CODM might efficiently disrupt all copies. It remains to be seen whether new crops developed with this technology will be classified as GM and whether they will carry the same regulatory burdens (Hartung and Schiemann, 2014). Use of site-specific nucleases means that the number of off-site mutations is much less than using chemical and physical mutants and means there is less of the perceived concern with GM technology that transgenes will integrate randomly in the genome and produce unintended consequences such as disrupting host metabolism or synthesis of toxic or allergenic compounds (Podevin et al., 2013). Although there have been several cases where the US Department of Agriculture ruled products of this technology do not have to be treated as GM organisms, the EU commission has not yet adopted a stance and this is frustrating efforts to adopt gene-editing throughout the 28 member states (Abbott, 2015). The technology has progressed rapidly in recent years, and if starting the project again now, genome editing would have to be considered as a realistic means to achieve the end goals. Likewise the cost of next generation sequencing has continued to come down and perhaps with a genome sequence of opium poppy, and prior knowledge of the number of copies of DIOX genes and their locations, different decisions could have been made and different points in the project i.e. more sequence information may have aided primer design for TILLING, the screens would have been limited to only those copies likely to be functional, crosses between plants with mutations in closely linked genes may not have been attempted (i.e. the W261* x Q254* cross) and the regulatory regions of DIOX genes could have been assessed for induced changes in nucleic acid sequence in those lines which have high thebaine/oripavine but for which no change in the coding sequence has yet been detected.

The work has also demonstrated the expression levels of CODM and T6ODM at various stages of plant development. Flowering was shown to be an important time for changes in

gene expression and changes in alkaloid composition of the latex. The activities of these genes in pathways other than morphine biosynthesis were also confirmed. CODM mutant plants had increased levels of cryptopine and protopine alkaloids in their roots while the levels in the roots of T6ODMa knockout was similar to wild type HM2. The highest levels were observed in the HighT fwd screen mutant, presumably due to reduced activities of T6ODMb or other copies of T6ODM. The exact roles of each the different copies of T6ODM needs to be investigated. It is probable that gene duplication has resulted in a family of genes with roles in a variety of biosynthetic pathways-the morphinan but also in the pathways to protopines (Farrow and Facchini, 2013) and pacodine (Farrow and Facchini, 2015). T6ODMb, for example, looks to be intermediary between T6ODMa (98% sequence identity at the amino acid level) and CODM i.e. in those positions where amino acid differences exist between T6ODMa and T6ODMb, the amino acid of T6ODMb corresponds to the amino acid at the equivalent position in CODM. It remains to be seen what its substrate specificities are and how efficient it is in comparison to CODM and T6ODMa.

Combining detected mutations in different copies of CODM by crossing failed due to close linkage. Evidence to support this conclusion was gained when the BAC sequences also suggested physical proximity of some copies. Therefore, alternative strategies for obtaining complete blocks a certain metabolic steps would be required. By embarking on another round of EMS mutagenesis of bulked up CODMb Q254* mutant seed an alternative strategy was initiated. Physical rather than chemical mutagens could be deployed if further BAC or genomic sequencing reveals close linkage of copies of CODM. Gamma radiation can produce small indels as well as point mutations (mostly transversions as opposed to transitions) as demonstrated by exposure of seeds of *japonica* rice to 500 Gy of gamma rays (Sato et al., 2006). Fast neutrons are uncharged particles of high kinetic energy and are generated in nuclear reactors or in accelerators. Fast neutrons impart some of their high kinetic energy via collisions, largely with protons within the material. Fast neutron mutagenesis generally causes genomic deletions rather than point mutations. “Delete-a-gene” technology can be used to screen for deletions using primers flanking the target gene. Discovery can be aided by

adjusting the PCR extension time to promote the preferential amplification of shorter amplicons (Li et al., 2002; Li and Zhang, 2002). This method was improved upon in deletion-based (de)-TILLING whereby much smaller deletions could be detected in a very large *Medicago truncatula* fast neutron generated mutant population (Rogers et al., 2009). This is a variation on TILLING where mutant populations are screened for deletions using two rounds of nested PCRs in which a third 'poison' primer (in the deleted region) is included. The poison primer amplifies a 'suppressor' product which, being smaller, is amplified from the wild type more efficiently. The amplicon cannot act as a template in the second round of PCR, however, and the mutant product is amplified efficiently. Fast neutron mutagenesis of opium poppy has been described before but the molecular basis of the observed phenotypes has not been elucidated (Fist et al., 2013). In order to carry out effective TILLING or (de)-TILLING studies additional sequence information is required to understand the genomic context of these 2-ODD genes. With the advent of ever more affordable next generation sequencing platforms, the possibility of sequencing the genome of opium poppy must be a realistic prospect. In the meantime, the material generated in this project provides a good starting point for further marker assisted breeding to create bespoke varieties of opium poppy. For example, introducing the CODMa/c E193K and CODMb R158K mutations into a commercial morphine line increased codeine yields 6 fold. Introducing these CODM mutations to a high thebaine line had the effect of increasing thebaine yields still further. Finally, further work is required to identify the causative mutations involved in the high oripavine phenotypes observed in this project.

A growing area of research is in the development of microbial based systems for the production of plant natural products. Common reasons cited for this approach include supply inefficiencies and variability in yields associated with the impact of climate and disease. The general objective of such work is to engineer microbes to complement or even replace plant based production systems. One such example is the supply of the antimalarial drug artemisinin. It can be isolated from the medicinal plant *Artemisia annua* (Sweet wormwood) which is cultivated in China, Vietnam and East Africa. Recently, synthetic biology researchers

have generated a strain of yeast that contains the required plant genes to reconstitute the pathway to a precursor of artemisinin, artemisinic acid. Since 2013, French company Sanofi have been using this technology for the commercial production of artemisinin at a facility in Italy. The production costs are in the region of $350-400 per kg which is comparable to the cost of plant biomass based extraction (Schaefer, 2015). This approach is seen as a steady and less volatile supply source of the drug and could supplement plant based production systems which still have major role with the advent of new high yielding varieties developed with molecular breeding techniques (Graham et al., 2010; Townsend et al., 2013). As mentioned in section 1.6.4 there has been a push recently to move BIA biosynthesis from opium poppy into yeast based production systems. Yields of, for example, thebaine from yeast based systems are still relatively small and inefficient-a yield of 6.4 μg/l thebaine was reported from strains of yeast with an impressive construction of 24 heterologous expression cassettes, 21 new enzyme activities and overexpression/inactivation of native enzymes (Galanie et al., 2015). The authors recognise that several orders of magnitude greater yields will be required if commercial production of morphinans in yeast based systems is to be feasible. It was claimed that a target of 5 g/l would need to be reached. A 5 ml culture would therefore produce 25 mg thebaine and potentially reduce the need for the cultivation of 0.2m^2 of opium poppy. However thebaine yields of 70 kg per hectare (10000 m^2) are achievable in Tasmania which equates to 1400 mg per 0.2 m^2. Therefore the 5 g/l target may still be an underestimation (50x) of the yield that would be required to become competitive with the existing poppy industry. However, the potential of the approach may well be in the formation of derivatives, both known and new to nature. Experimenting with various tailoring enzymes on a plant-derived scaffold may in future produce novel alkaloids with medicinal uses. It may also be useful for the production of compounds that are extracted from other plants with less established agricultural practices, those more susceptible to changes in climate and disease or with long generation times.

Appendices

Appendix A- Creation of DNA pool plates for TILLING

Numbers refer to the Sd numbers of plants from which DNA was placed in each pool i.e. well A1 of pool plate 1 contained DNA from Sd-100001, Sd-100002, Sd-100003 and Sd-100004.

Pool Plate 1

	row A	1	2	3	4	5	6	7	8	9	10	11	12
	A	100001	100009	100017	100025	100033	100041	100049	100057	100065	100073	100177	100185
	B	100002	100010	101203	100026	100034	100042	100050	100058	100066	100074	100178	100186
	C	100003	100011	100019	100027	100035	100043	100051	100059	100067	100075	100179	100187
Source Plate D-10001	D	100004	101202	100020	100028	100036	100044	100052	100060	100068	100076	100180	100188
	row B	1	2	3	4	5	6	7	8	9	10	11	12
	E	100005	100013	100021	100029	100037	100045	100053	100061	100069	100077	100181	100189
	F	100006	100014	100022	100030	100038	100046	100054	100062	101213	100078	100182	100190
	G	100007	100015	100023	100031	101204	100047	100055	100063	100071	100079	100183	100191
	H	100008	100016	100024	100032	100040	100048	100056	100064	100072	100080	100184	100192
	row C	1	2	3	4	5	6	7	8	9	10	11	12
	A	100081	100089	100097	100105	100113	100121	101211	100137	100145	100153	100161	100169
	B	100082	101206	101209	100106	100114	100122	100130	100138	100146	100154	101214	100170
	C	100083	100091	100099	100107	100115	100123	100131	100139	100147	100155	100163	100171
Source Plate D-10002	D	101205	100092	100100	100108	100116	100124	100132	100140	101212	100156	100164	100172
	row D	1	2	3	4	5	6	7	8	9	10	11	12
	E	100085	101207	1000101	100109	100117	100125	100133	100141	100149	100157	100165	100173
	F	100086	100094	100102	100110	100118	100126	100134	100142	100150	100158	100166	100174
	G	100087	100095	100103	101210	100119	100127	100135	100143	100151	100159	100167	100175
	H	100088	100096	100104	100112	100120	100128	100136	100144	100152	100160	100168	100176
	row E	1	2	3	4	5	6	7	8	9	10	11	12
	A	100193	100281	100289	100297	100305	100313	100321	101230	100337	100345	100353	100361
	B	100194	100282	101227	100298	100306	100314	100322	101330	100338	100346	100354	100362
	C	100195	100283	100291	100299	101228	100315	100323	101331	100339	100347	100355	101231
Source Plate D-10003	D	100196	100284	100292	100300	100308	100316	100324	101332	100340	100348	100356	100364
	row F	1	2	3	4	5	6	7	8	9	10	11	12
	E	100197	100285	100293	100301	101229	100317	100325	100333	100341	100349	100357	100365
	F	101215	100286	100294	100302	100310	100318	100326	100334	100342	100350	100358	100366
	G	100199	100287	100295	100303	100311	100319	100327	100335	100343	100351	100359	100367
	H	100200	100288	100296	100304	100312	100320	100328	100336	101344	100352	100360	100368
	row G	1	2	3	4	5	6	7	8	9	10	11	12
	A	100201	100209	100217	100225	100233	100241	100249	100257	100265	100273	100369	100377
	B	100202	101216	100218	100226	101220	100242	100250	100258	100266	100274	100370	100378
	C	100203	100211	100219	101217	100235	100243	100251	100259	100267	100275	100371	100379
Source Plate D-10004	D	100204	100212	100220	100228	100236	100244	100252	100260	100268	100276	100372	100380
	row H	1	2	3	4	5	6	7	8	9	10	11	12
	E	100205	100213	100221	100229	101221	**100245**	100253	100261	100269	101232	100373	100381
	F	100206	100214	100222	101218	100238	100246	101223	100262	100270	100278	100374	100382
	G	100207	100215	100223	100231	100239	100027	100255	100263	100271	10079	100375	100383
	H	100208	100216	100224	101219	101222	100248	100256	100264	100272	100280	100376	100384

Pool Plate 2

	row A	1	2	3	4	5	6	7	8	9	10	11	12
	A	100401	100409	100417	100425	100433	100441	100449	100457	100465	100473	100481	100489
	B	100402	100410	100418	100426	100434	100442	100450	101236	100466	100474	100482	100490
	C	100403	101224	100419	100427	100435	100443	100451	100459	100467	100475	100483	100491
Source Plate D-10005	D	100404	100412	101233	100428	100431	100444	100452	100460	100468	100476	100484	100492
	row B	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	100405	100413	100421	100429	100437	100445	100453	101237	100469	100477	100485	100493
	F	100406	100414	100423	100430	100438	100446	100454	100462	100470	100478	100486	100494
	G	100407	100415	100423	100431	100434	100447	100455	100463	101238	100479	100487	100495
	H	100408	101226	100424	101234	100440	100448	100456	100464	100472	100480	100488	100496
	row C	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	A	100497	100505	100513	100521	100529	100537	100545	100553	100561	100569	100577	100585
	B	100498	100506	100514	100522	100530	100538	101164	100554	100562	100570	100578	100586
	C	100494	100507	100515	100523	100531	100539	100547	100555	100563	100571	100579	100587
Source Plate D-10006	D	100500	100508	100516	101162	100532	100540	100548	100556	100564	100572	100580	100588
	row D	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	100501	100509	101161	100525	100533	100541	100549	100557	100565	100573	100582	100589
	F	100502	100510	100518	100526	100534	100542	100550	100558	100566	100574	100581	100590
	G	100503	100511	100519	100527	100535	100543	100551	100559	100567	100575	100583	100591
	H	101240	100512	100520	100528	100536	100544	100552	100560	100568	100576	100584	100592
	row E	**1**	**2**	**3**	**4**	**5**	**3**	**7**	**8**	**9**	**10**	**11**	**12**
	A	100593	100601	100609	100617	100625	100633	100641	100649	100657	100664	101177	100680
	B	100594	100602	100610	100618	100626	100634	100642	100650	100658	100665	100673	101181
	C	100595	100603	100611	100619	100627	100635	100643	100651	100659	100666	101178	101182
Source Plate D-10007	D	100596	100604	100612	100620	100628	100636	101169	100652	100660	100667	100675	101185
	row F	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	100597	100605	100613	100621	100629	100637	100645	100653	100661	100668	101179	101186
	F	100598	100606	100614	100622	100630	100638	100646	101172	101174	100669	100677	101187
	G	100599	100607	100615	100623	100631	100639	100647	100655	100663	100670	100678	101188
	H	100600	100608	100616	100624	100632	100640	101171	100656	100662	100671	101180	101189
	row G	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**3**	**11**	**12**
	A	100385	100393	100761	100769	100777	100785	100793	100681	100689	100697	100705	100713
	B	100386	100394	100762	101165	100778	100786	100794	100682	100690	100698	101190	100714
	C	100387	100395	100763	100771	100779	100787	100795	100683	100691	100699	100707	100715
Source Plate D-10008	D	100388	100396	100764	100772	100780	100788	100796	100684	100692	100700	100708	100716
	row H	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	100389	100397	100765	100773	100781	100789	100797	100685	100693	100701	100709	100717
	F	100390	100398	100766	100774	100782	100790	100798	100686	100694	100702	100710	100718
	G	100391	100399	100767	100775	100783	100791	100799	100687	100695	100703	100711	100719
	H	100392	100400	100768	100776	100784	100792	100800	101175	100696	100704	100712	100720

Pool Plate 3

	row A	1	2	3	4	5	6	7	8	9	10	11	12
	A	100721	100729	100737	100745	100753	100841	100849	100857	100865	100873	100801	100809
	B	100722	100730	100738	100746	100754	100842	100850	100858	100866	100874	100802	100810
	C	100723	100731	100739	100747	100755	100843	100851	100859	100867	100875	100803	100811
Source Plate D-10009	D	100724	100732	100740	100748	100756	100844	100852	100860	100868	101184	100804	100812
	row B	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	100725	100733	100741	100749	100757	100845	100853	100861	100869	100877	100805	100813
	F	100726	100734	100742	100750	100758	100846	100854	100862	100870	100878	100806	100814
	G	100727	100735	100743	100751	100759	100847	100855	100863	100871	100879	100807	100815
	H	100728	100736	101166	101168	101170	101173	100856	100864	100872	100880	10088	101191
	row C	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	A	100881	100889	100897	100905	101154	100921	100929	100937	100945	100953	100961	100969
	B	100882	100890	100898	100906	100914	100922	100930	100938	100946	100954	100962	101197
	C	100883	100891	100899	100907	100915	100923	100931	100939	100947	100955	100963	100971
Source Plate D-10010	D	100884	100892	100900	100908	100916	100924	100932	100940	100948	100956	100964	100972
	row D	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	100885	100893	100901	100909	100917	100925	100933	101195	100949	100957	100965	100973
	F	100886	100894	100902	100910	100918	100926	100934	100942	100950	100958	100966	100974
	G	100887	100895	100903	100911	100919	100927	100935	100943	100951	100959	100967	100975
	H	100888	100896	101193	100912	100920	100928	100936	100944	100952	100960	101196	100976
	row E	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	A	100817	100825	100833	101121	101009	101017	101025	101033	101131	101089	101135	101138
	B	100818	100826	100834	101002	101010	101018	101026	101130	101082	101133	101098	101106
	C	100819	100827	100835	101003	101011	101019	101027	101035	101083	101091	101099	101107
Source Plate D-10011	D	100820	100828	100836	101004	101012	101020	101028	101036	101084	101092	101100	101108
	row F	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	100821	100829	100837	101005	101013	101021	101029	101037	101085	101093	101101	101109
	F	100822	100830	100838	101006	101014	101022	101030	101038	101086	101094	101136	101110
	G	101198	100831	100839	101007	101015	101023	101031	101039	101087	101134	101103	101111
	H	100824	100832	101200	101008	101016	101024	101128	101040	101132	101096	101137	101112
	row G	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	A	100977	100985	100993	101041	101049	101057	101065	101073	101140	101149	101158	
	B	100978	100986	100994	101042	101050	101058	101066	101074	101141	101150	101159	
	C	100979	100987	100995	101043	101051	101059	101067	101075	101142	101151	101114	
Source Plate D-10012	D	100980	100988	101122	101044	101052	101060	101068	101076	101143	101152	101116	
	row H	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	100981	100989	100997	101045	101053	101061	101069	101077	101145	101153	101118	
	F	100982	101199	100998	101046	101054	101062	101070	101078	101146	101154		
	G	100983	100991	100999	101123	101055	101063	101071	101079	101147	101155		
	H	100984	100992	101000	101124	101056	101064	101072	101080	101148	101157		

Pool Plate 4

	row A	1	2	3	4	5	6	7	8	9	10	11	12
	A	101297	101265	101273	101262	101289	101297	101306	101314	101322	101330	101338	101348
	B	101249	101266	101274	101281	101290	101298	101307	101315	101323	101331	101339	101349
	C	101241	101267	101275	101282	101291	101299	101308	101316	101324	101332	101341	101350
Source Plate D-100013	D	101256	101268	101276	101283	101292	101300	101309	101317	101325	101333	101342	101351
	row B	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	101258	101269	101277	101284	101293	101301	101310	101318	101326	101334	101343	101352
	F	101259	101270	101278	101285	101294	101302	101311	101319	101327	101335	101344	101353
	G	101261	101271	101279	101286	101295	101303	101312	101320	101328	101336	101345	101357
	H	101263	101272	101280	101287	101296	101304	101313	101321	101329	101337	101346	101358
	row C	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	A	101359	101367	101375	101383	101394	101405	101419	101428	101439	101448	101456	101464
	B	101360	101368	101376	101384	101396	101410	101420	101429	101440	101449	101457	101465
	C	101361	101369	101377	101386	101397	101411	101421	101430	101441	101450	101458	101466
Source Plate D-10014	D	101362	101370	101378	101387	101398	101413	101422	101431	101442	101451	101459	101467
	row D	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	101363	101371	101379	101388	101399	101414	101424	101432	101443	101452	11460	101468
	F	101364	101372	101380	101389	101400	101415	101425	101434	101444	101453	101461	101469
	G	101365	101373	101381	101390	101401	101417	101426	101437	101446	101454	101462	101470
	H	101366	101374	101382	101392	101403	101418	101427	101438	101447	101455	101463	101471
	row E	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	A	101472	101480	101488	101496	101508	101522	101542	101560	101694	101712	101729	101738
	B	101473	101481	101489	101498	101509	101526	101543	101683	101695	101714	101730	101739
	C	101474	101482	101490	101499	101510	101528	101549	101684	101696	101720	101731	101740
Source Plate D-10015	D	101475	101483	101491	101500	101512	101531	101550	101685	101699	101721	101732	101741
	row F	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	101476	101484	101493	101501	101514	101534	101551	101686	101701	101722	101733	101742
	F	101477	101485	101494	101502	101515	101535	101552	101687	101704	101723	101734	101743
	G	101478	101486	101495	101503	101519	101536	101554	101688	101706	101726	101735	101744
	H	101479	101487	101497	101506	101520	101541	101557	101691	101711	101728	101736	101745
	row G	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	A	101564	101576	101586	101594	101604	101613	101621	101630	101638	101646	101656	101664
	B	101566	101577	101587	101596	101605	101614	101623	101631	101639	101649	101657	101665
	C	101567	101578	101588	101597	101606	101615	101624	101632	101640	101650	101658	101666
Source Plate D-10016	D	101569	101580	101589	101598	101607	101616	101625	101633	101641	101651	101659	101667
	row H	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	101570	101582	101590	101599	101608	101617	101626	101634	101642	101652	101660	101668
	F	101573	101583	101591	101600	101610	101618	101627	101635	101643	101653	101661	101669
	G	101574	101584	101592	101602	101611	101619	101628	101636	101644	101654	101662	101670
	H	101575	101585	101593	101603	101612	101620	101629	101637	101645	101655	101663	101671

Pool Plate 5

	row A	1	2	3	4	5	6	7	8	9	10	11	12
	A	101747	101755	101763	101772	101780	101788	101796	101804	101815	101822	101831	101840
	B	101748	101756	101764	101773	101781	101789	101797	101805	101814	101823	101832	101961
	C	101749	101757	101765	101774	101782	101790	101798	101806	101816	101825	101833	101962
Source Plate D-10017	D	101750	101758	101766	10775	101783	101791	101799	101807	101817	101826	101834	101964
	row B	1	2	3	4	5	6	7	8	9	10	11	12
	E	101751	101759	101767	101776	101784	101792	101800	101808	101808	101827	101835	101963
	F	101752	101760	101768	101777	101785	101793	101801	101810	101810	101828	101837	101965
	G	101753	101761	101769	101778	101786	101794	101802	101812	101812	101829	101838	101966
	H	101754	101762	101771	101779	101787	101795	101803	101813	101813	101830	101839	101967
	row C	1	2	3	4	5	6	7	8	9	10	11	12
	A	101672	101841	101852	101861	101871	101885	101894	101902	101910	101918	101926	101934
	B	101673	101842	101853	101862	101872	101886	101895	101903	101911	101919	101927	101935
	C	101675	101843	101854	101864	101874	101888	101896	101904	101912	101920	101928	101936
Source Plate D-10018	D	101676	101845	101855	101865	101875	101889	101897	101905	101913	101921	101929	101937
	row D	1	2	3	4	5	6	7	8	9	10	11	12
	E	101677	101847	101856	101866	101878	101890	101898	101906	101914	101922	101930	101938
	F	101678	101848	101857	101868	101880	101891	101899	101907	101915	101923	101931	101939
	G	101679	101849	101858	101870	101881	101892	101900	101908	101916	101924	101932	101940
	H	101680	101851	101860	101869	101884	101893	101901	101909	101917	101925	101933	101941
	row E	1	2	3	4	5	3	7	8	9	10	11	12
	A	101968	101977	101985	101985	101993	102001	102009	102018	102029	102041	102059	102067
	B	101969	101978	101986	101986	101994	102002	102010	102019	102031	102042	102060	102068
	C	101970	101979	101987	101987	101995	102003	102011	102020	102032	102043	102061	102069
Source Plate D-10019	D	101972	101980	101988	101988	101996	102004	102012	102021	102033	102044	10206	102070
	row F	1	2	3	4	5	6	7	8	9	10	11	12
	E	101973	101981	101989	101989	101997	102005	102013	102022	102034	102046	102063	102071
	F	101974	101982	101990	101990	101998	102006	102014	102023	102038	102047	102064	102072
	G	101975	101983	101991	101991	101999	102007	102015	102024	102939	102049	102065	102073
	H	101976	101984	101992	101992	102000	102008	102016	102028	102040	102048	102066	102074
	row G	1	2	3	4	5	6	7	8	9	3	11	12
	A	101942	101950	101959	102087	102095	102104	102112	102120	102288	102296	102305	102313
	B	101943	101951	101960	102088	102096	102105	102113	102281	102289	102297	102306	10314
	C	101944	101952	102081	102089	102097	102106	102114	102282	102290	102298	102307	102315
Source Plate D-10020	D	101945	101954	102082	102090	102098	102107	102115	102282	102291	102300	102308	102316
	row H	1	2	3	4	5	6	7	8	9	10	11	12
	E	101946	101955	102084	102091	102100	102108	102116	102284	102292	102301	102309	102317
	F	101947	101956	102083	102092	102101	102109	102117	102285	102293	102302	102310	102318
	G	101948	101957	102085	102093	102102	102110	102118	102286	102294	102303	102311	102319
	H	101949	101958	102086	102094	102103	102111	102119	102287	102295	102304	102312	102320

Pool Plate 6

	row A	1	2	3	4	5	6	7	8	9	10	11	12
	A	102075	102163	102171	102179	102188	102198	102209	102218	102227	102236	102245	102253
	B	102076	102164	102172	102180	102189	1021990	102210	102219	102228	102237	102246	102254
	C	102077	102165	102173	102181	102190	102200	102211	102220	102229	102238	102247	102255
Source Plate D-10021	D	102078	102166	102174	102182	102191	102203	102212	102221	102230	102239	102248	102256
	row B	1	2	3	4	5	6	7	8	9	10	11	12
	E	102079	102167	102175	102183	102192	102204	102213	102223	102231	102241	102249	102257
	F	102080	102168	102176	102184	102195	102205	102214	102224	102232	102242	102250	102258
	G	102161	102169	102177	102186	102196	102206	102216	102225	102234	102243	102251	102259
	H	102162	102170	102178	102187	102197	102207	102217	102226	102235	102244	102252	102260
	row C	1	2	3	4	5	6	7	8	9	10	11	12
	A	102321	102329	102337	102346	102354	102364	102372	102383	102393	102404	102413	102423
	B	102322	102330	102338	102347	102355	102365	102373	102385	102394	102405	102415	102425
	C	102323	102331	102340	102348	102356	102366	102374	102386	102396	102406	102417	102426
Source Plate D-10022	D	102324	102332	102341	102349	102357	102367	102377	102387	102397	102407	102418	102427
	row D	1	2	3	4	5	6	7	8	9	10	11	12
	E	102325	102333	102342	102350	102359	102368	102378	102388	103398	102409	102419	102428
	F	102326	102334	102343	102351	102360	102369	102379	102389	102399	102410	102420	102429
	G	102327	102335	102344	102352	102361	102370	102380	102390	102401	102411	102421	102430
	H	102328	102336	102345	102353	102362	102371	102381	102391	102402	102412	102422	102431
	row E	1	2	3	4	5	6	7	8	9	10	11	12
	A	102261	102270	102432	102440	102448	102456	102467					
	B	102262	102273	102433	102441	102449	102457	102468					
	C	102263	102274	102434	102442	102450	102458	102471					
Source Plate D-10023	D	102264	102276	102435	102443	102451	102460	102473					
	row F	1	2	3	4	5	6	7	8	9	10	11	12
	E	102265	102277	102436	102444	102452	102462	102474					
	F	102266	102278	102437	102445	102453	102463	102475					
	G	102268	102279	102438	102446	102454	102464	102476					
	H	102269	102280	102439	102447	102455	102466	102477					
	row G	1	2	3	4	5	6	7	8	9	10	11	12
	A	102482	102490	102499	102507	102515	102563	102571	102579	102587	102596	102644	102652
	B	102483	102491	102500	102508	102508	102564	102572	102580	102588	102597	102645	102653
	C	102484	102492	102501	102509	102509	102565	102573	102581	102589	102598	102646	102654
Source Plate D-10024	D	102485	102493	102502	102510	102510	102566	102574	102582	102590	102599	102647	102655
	row H	1	2	3	4	5	6	7	8	9	10	11	12
	E	102486	102494	102503	102511	102511	102567	102575	102583	102591	102600	102648	192656
	F	102487	102495	102504	102512	102512	102568	102576	102584	102592	102641	102649	102657
	G	102488	102496	102505	102513	102513	102569	102577	102585	102593	102642	102650	102658
	H	102489	102498	102506	102514	102514	102570	102578	102586	102595	102643	102651	102659

Pool Plate 7

	row A	1	2	3	4	5	6	7	8	9	10	11	12
	A	102521	102531	102539	102548	102556	102605	102613	102621	102629	102637	102686	102694
	B	102522	102532	102540	102549	102557	102606	102614	102622	102630	102638	102687	102695
	C	102524	102533	102541	102550	102558	102607	102615	102623	102631	102639	102688	102696
Source Plate D-10025	D	102525	102534	102542	102551	102559	102608	102616	102624	102632	102640	102689	102697
	row B	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	102526	102535	102543	102552	102560	102609	102617	102625	102633	102681	102690	102698
	F	102527	102536	102544	102553	102601	102610	102618	102626	102634	102682	102691	102699
	G	102529	102537	102545	102554	102603	102611	102619	102627	102635	102683	102692	102700
	H	102530	102538	102547	102555	102604	102612	102620	102628	102636	102684	102693	102701
	row C	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	A	102660	102668	102676	102726	102735	102744	102754	102832	102843	102852	102860	102869
	B	102661	102669	102677	102727	102736	102746	102755	102833	102844	102853	102861	102870
	C	102662	102670	102678	102728	102738	102747	102756	102835	102845	102854	102852	102912
Source Plate D-10026	D	102663	102671	102679	102729	102739	102748	102757	102836	102844	102863	102864	102914
	row D	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	102664	102672	102680	102730	102740	102749	102758	102837	102847	102856	102865	102915
	F	102665	102673	102721	102731	102741	102750	102759	102840	102848	102857	102866	102916
	G	102666	102674	102723	102733	102742	102751	102760	102841	102849	102858	102867	102917
	H	102667	102675	102724	102734	102743	102753	102831	102842	102851	102859	102868	102918
	row E	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	A	102702	102710	102718	102786	102777	102786	102799	102877	102885	102893	102901	102910
	B	102703	102711	102719	102769	102778	102788	102800	102878	102886	102894	102902	102951
	C	102704	102712	102720	102771	102779	102789	102871	102879	102887	102895	102903	102952
Source Plate D-10027	D	102705	102713	102761	102772	102780	102790	102872	102880	102888	102896	102904	102953
	row F	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	102706	102714	102762	102773	102781	102791	102873	102881	102889	102897	102905	102954
	F	102707	102715	102763	102774	102782	102792	102874	102882	102890	102898	102906	102955
	G	102708	102716	102765	102775	102783	102797	102875	102883	102891	102899	102908	102956
	H	102709	102717	102767	102776	102785	102798	102876	102884	102892	102900	102909	102957
	row G	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	A	102919	102927	102938	102992	103002	103011	103020	103193	103201	103209	103217	103225
	B	102920	102928	102941	102994	103003	103012	103021	103194	103202	103210	103218	103226
	C	102921	102929	102943	102996	103004	103014	103022	103195	103203	103211	103219	103227
Source Plate D-10028	D	102922	102930	102944	102997	103005	103015	103024	103196	103204	103212	103220	103228
	row H	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	102923	102933	102945	102998	103006	103016	103027	103197	103205	103213	103221	103229
	F	102924	102935	102946	102999	103007	103017	103028	103198	103206	103214	103222	103230
	G	102925	102936	102950	103000	103009	103018	103030	103199	103207	103215	103323	103232
	H	102926	102937	102991	103001	103010	103019	103192	103200	103208	103216	103324	103233

Pool Plate 8

	row A	1	2	3	4	5	6	7	8	9	10	11	12
	A	102958	102966	102974	102982	102990	103038	103047	103055	103063	103071	103079	103087
	B	102959	102967	102975	102983	103031	103039	103048	103056	103064	103072	103080	103088
	C	102960	102968	102976	102984	103032	103041	103049	103057	103065	103073	103081	103089
Source Plate D-10029	D	102961	102969	102977	102985	103033	103042	103050	103058	103066	103074	103082	103090
	row B	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	102962	102970	102978	102986	103034	103043	103051	103059	103067	103075	103083	103091
	F	102963	102971	102979	102987	103035	103044	103052	103060	103068	103076	103084	103092
	G	102964	102972	102980	102988	103036	103045	103053	103061	103069	103077	103085	103093
	H	102965	102973	102981	102989	103037	103046	103054	103062	103070	103078	103086	103094
	row C	**1**	**2**	**3**	**4**	**5**	**3**	**7**	**8**	**9**	**10**	**11**	**12**
	A	103095	103103	103111	103119	103128	103136	103144	103152	103160	103168	103176	103184
	B	103096	103104	103112	103120	103129	103137	103145	103153	103161	103169	103177	103185
	C	103097	103105	103113	103121	103130	103138	103146	103154	103162	103170	103178	103186
Source Plate D-10030	D	103098	103106	103114	103122	103131	103139	103147	103155	103163	103171	103179	103187
	row D	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	103099	103107	103115	103123	103132	103140	103148	103156	103164	103172	103180	103188
	F	103100	103108	103116	103124	103133	103141	103149	103157	103165	103173	103181	103189
	G	103101	103109	103117	103125	103134	103142	103150	103158	103166	103174	103182	103190
	H	103102	103110	103118	103127	103135	103143	103151	103159	103167	103175	103183	103191
	row E	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**3**	**11**	**12**
	A	103234	103244	103252	103261	103270	103279	103289	103298	103307	103317	103328	103337
	B	103235	103245	103253	103262	103271	103282	103290	103299	103309	103318	103330	103338
	C	103236	103246	103255	103264	103272	103283	103291	103300	103310	103319	103331	103339
Source Plate D-10031	D	103239	103247	103256	103265	103273	103284	103292	103301	103311	103320	103332	103340
	row F	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	103240	103248	103257	103266	103274	103285	103293	103302	103312	103321	103333	103341
	F	103241	103249	103258	103267	103276	103286	103295	103303	103314	103322	103334	103342
	G	103242	103250	103259	103268	103277	103287	103296	103304	103315	103324	103335	103343
	H	103243	103251	103260	103269	103278	103788	103297	103305	103316	103325	103336	103346
	row G	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	A	103348	103357	103366	103378	103389	103397	103409	103419	103428	103438	103447	103456
	B	103349	103358	103367	103379	103390	103398	103411	103420	103429	103439	103448	103458
	C	103350	103359	103369	103381	103391	103399	103412	103421	103432	103440	103449	103459
Source Plate D-10032	D	103351	103360	103370	103382	103392	103402	103413	103422	103433	103442	103451	103460
	row H	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	103352	103361	103371	103383	103393	103405	103415	103423	103434	103443	103452	103461
	F	103353	103362	103373	103385	103394	103406	103416	103424	103435	103444	103453	103462
	G	103354	103364	103375	103386	103395	103407	103417	103426	103436	103445	103454	103464
	H	103356	103365	103376	103387	103396	103408	103418	103427	103437	103446	103455	103465

Pool Plate 9

	row A	1	2	3	4	5	6	7	8	9	10	11	12
	A	103466	103477	103489	103497	103507	103517	103529	103540	103551	103560	103632	103642
	B	103467	103479	103490	103498	103508	103518	103530	103546	103552	103561	103634	103644
	C	103468	103480	103491	103499	103509	103519	103532	103547	103553	103562	103636	103645
Source Plate D-10033	D	103469	103482	103492	103500	103510	103520	103534	103548	103555	103563	103637	103646
	row B	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	103471	103483	103493	103501	103512	103523	103535	103549	103554	103564	103638	103647
	F	103472	103484	103494	103504	103514	103514	103536	103550	103556	103565	103639	103648
	G	103474	103485	103495	103505	103515	103525	103537	103629	103557	103566	103340	103626
	H	103475	103487	103496	103506	103516	103528	103538	103628	103558	103631	103341	103627
	row C	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	A	103567	103575	103583	103593	103601	103618						
	B	103568	103576	103584	103594	103602	103619						
	C	103569	103577	103585	103595	103603	103620						
Source Plate D-10034	D	103570	103578	103586	103596	103604	103621						
	row D	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	103571	103579	103588	103597	103609	103622						
	F	103572	103580	103590	103598	103611	103623						
	G	103573	103582	103592	103599	103612	103625						
	H	103574	103581	103592	103600	103616	103539						

Pool Plate 10

	row A	1	2	3	4	5	6	7	8	9	10	11	12
	A	103647	103661	103682	103773	103790	103743	103691	103706	103725	103649	103674	103783
	B	103652	103666	103683	103777	103791	103744	103694	103708	103727	103650	103675	103787
	C	103655	103667	103684	103779	103792	103749	103695	103709	103729	103654	103685	103788
Source Plate D-10035	D	103656	103668	103688	103780	103796	103751	103699	103710	103730	103662	103687	103795
	row B	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	103657	103669	103689	103781	103797	103752	103702	103712	103803	103664	103774	103798
	F	103658	103679	103690	103784	103732	103753	103703	103714	103807	103665	103775	103800
	G	103659	103680	103771	103785	103759	103759	103704	103718	103808	103670	103776	103731
	H	103660	103681	103772	103789	103764	103764	103705	103719	103809	103672	103782	103733
	row C	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	A	103842	103853	103862	103882	103891	103901	103910	104021	104040	104056	103971	103986
	B	103843	103854	103866	103883	103892	103902	103911	104024	104043	104057	103974	103988
	C	103845	103855	103867	103884	103893	103903	103912	104026	104044	104059	103975	103989
Source Plate D-10036	D	103846	103856	103868	103885	103894	103904	103914	104028	104045	103961	103976	103990
	row D	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	103848	103857	103870	103887	103895	103905	103915	104029	104047	103962	103980	103991
	F	103849	103859	103871	103888	103898	103907	103916	104032	104048	103963	103981	103992
	G	103851	103860	103875	103889	103899	103908	103919	104037	104052	103964	103982	103997
	H	103852	103861	103881	103890	103900	103909	103920	104039	104055	103970	103985	103999
	row E	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	A	103740	103692	103722	103928	103938	103949	103958	104075	104083	104004	104106	104165
	B	103742	103697	103723	103929	103941	103950	103959	104076	104086	104011	104107	104166
	C	103748	103698	103921	103930	103942	103951	103960	104077	104087	104012	104108	104167
Source Plate D-10037	D	103754	103701	103922	103932	103943	103952	104062	104078	104090	104013	104111	104170
	row F	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	103756	103707	103923	103934	103944	103953	104065	104079	104091	104014	104112	104175
	F	103760	103713	103925	103935	103945	103954	104070	104080	104094	104016	104114	104180
	G	103767	103716	103926	103936	103946	103956	104072	104081	104095	104017	104161	104187
	H	103768	103720	103927	103937	103947	103957	104074	104082	104097	104019	104162	104188
	row G	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	A	104121	104135	104149	104201	104221	104238	104408	104418	104432	104324	104348	104359
	B	104124	104136	104150	104202	104227	104239	104409	104420	104436	104325	104349	104281
	C	104125	104137	104152	104206	104230	104240	104411	104421	104447	104326	104351	104282
Source Plate D-10038	D	104127	104139	104153	104207	104231	104401	104412	104422	104440	104330	104353	104283
	row H	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
	E	104130	104140	104154	104210	104232	104402	104413	104424	104331	104341	104354	104285
	F	104132	104141	104159	104215	104234	104403	104414	104426	104340	104343	104355	104286
	G	104133	104142	103958	104216	104235	104405	104415	104428	104321	104344	104356	104287
	H	104134	104148	103960	104219	104237	104406	104416	104430	104322	104346	104357	104288

Pool Plate 11

	row A	1	2	3	4	5	6	7	8	9	10	11	12
	A	104191	104269	104367	104382	104391	104400	104574	104586	104598	104529	104538	104549
	B	104192	104270	104368	104383	104392	104562	104576	104589	104600	104530	104539	104550
	C	104193	104271	104371	104384	104393	104563	104577	104590	104521	104531	104540	104551
Source Plate D-10039	D	104194	104274	104373	104386	104395	104565	104578	104592	104522	104532	104543	104553
	row B	1	2	3	4	5	6	7	8	9	10	11	12
	E	104197	104277	104374	104387	104396	104569	104579	104594	104525	104533	104544	104554
	F	104250	104361	104379	104388	104397	104570	104581	104595	104526	104535	104545	104555
	G	104253	104363	104380	104389	104398	104571	104583	104596	104527	104536	104546	104556
	H	104262	104364	104381	104390	104399	104572	104585	104597	104528	103537	104548	104557
	row C	1	2	3	4	5	6	7	8	9	10	11	12
	A	104289	104301	104315	104486	104496	104504	104517	104687	104701	104713	104609	104622
	B	104291	104307	104316	104487	104497	104507	104518	104688	104702	104714	104613	104623
	C	104292	104308	104317	104488	104498	104511	104519	104689	104703	104716	104614	104625
Source Plate D-10040	D	104294	104310	104318	104489	104499	104512	104520	104690	104704	104717	104615	104626
	row D	1	2	3	4	5	6	7	8	9	10	11	12
	E	104295	104311	104319	104490	104500	104513	104681	104693	104705	104719	104616	104633
	F	104297	104312	104320	104491	104501	104514	104682	104698	104708	104720	104617	104634
	G	104298	104313	104482	104494	104502	104515	104684	104699	104711	104605	104618	104635
	H	104300	104314	104485	104495	104503	104516	104686	104700	104712	104607	104621	104638
	row E	1	2	3	4	5	6	7	8	9	10	11	12
	A	104558	104729	104743	104754	104647	104658	104669	104845	104854	104866	104877	104811
	B	104560	104730	104744	104755	104649	104659	104671	104846	104855	104868	104878	104812
	C	104721	104733	104745	104756	104650	104660	104673	104847	104857	104869	104879	104813
Source Plate D-10041	D	104722	104734	104748	104758	104651	104662	104674	104849	104858	104870	104880	104814
	row F	1	2	3	4	5	6	7	8	9	10	11	12
	E	104723	104737	104749	104641	104652	104665	104676	104850	104861	104871	104802	104815
	F	104725	104738	104750	104642	104654	104666	104680	104851	104862	104872	104803	104816
	G	104726	104739	104752	104643	104655	104667	104643	104852	104864	104873	104805	104817
	H	104727	104741	104753	104644	104657	104668	104844	104853	104865	104874	104809	104818
	row G	1	2	3	4	5	6	7	8	9	3	11	12
	A	104639	104888	104902	104910	104761	104774	104786	105203	105212	105162	105172	105182
	B	104640	104889	104903	104911	104762	104775	104787	105204	105213	105163	105173	105183
	C	104881	104891	104904	104914	104764	104776	104788	105205	105214	105164	105174	105184
Source Plate D-10042	D	104882	104892	104905	104916	104765	104779	104790	105206	105216	105166	105175	105185
	row H	1	2	3	4	5	6	7	8	9	10	11	12
	E	104883	104893	104906	104917	104766	104780	104791	105207	105217	105167	105177	105186
	F	104884	104895	104907	104918	104767	104782	104794	105208	105218	105168	105178	105187
	G	104886	104899	104908	104919	104770	104783	104795	105209	105219	105170	105179	105188
	H	104887	104901	104909	104920	104773	104784	105202	105211	105222	105171	105181	105189

Pool Plate 12

	row A	1	2	3	4	5	6	7	8	9	10	11	12
	A	104819	104833	104924	104937	104948	104958	104967	104979	104988	104996	105006	105018
	B	104823	104834	104927	104938	104949	104959	104970	104980	104989	104997	105007	105019
	C	104824	104836	104928	104939	104950	104960	104971	104981	104990	104998	105009	105020
Source Plate D-10043	D	104825	104837	104930	104943	104951	104961	104973	104982	104991	104999	105010	105021
	row B	1	2	3	4	5	6	7	8	9	10	11	12
	E	104828	104839	104932	104944	104952	104962	104974	104983	104992	105000	105011	105022
	F	104829	104840	104933	104945	104954	104963	104975	104984	104993	105002	105012	105025
	G	104830	104921	104935	104946	104955	104964	104976	104985	104994	105003	105014	105026
	H	104831	104923	104936	104947	104957	104965	104977	104987	104995	105004	105017	105028
	row C	1	2	3	4	5	6	7	8	9	10	11	12
	A	105190	105200	105049	105057	105065	105075	105125	105134	105143	105154	105032	105081
	B	105191	105042	105050	105058	105066	105076	105126	105135	105145	105155	105033	105082
	C	105192	105043	105051	105059	105068	105077	105127	105137	105146	105156	105034	105083
Source Plate D-10044	D	105195	105044	105052	105060	105069	105079	105129	105138	105147	105157	105035	105084
	row D	1	2	3	4	5	6	7	8	9	10	11	12
	E	105195	105045	105053	105061	105070	105080	105130	105139	105148	105159	105036	105085
	F	105196	105046	105054	105062	105071	105121	105131	105140	105149	105160	105037	105086
	G	105197	105047	105055	105063	105073	105122	105132	105141	105150	105030	105038	105089
	H	105198	105048	105056	105064	105074	105124	105133	105142	105151	105031	105039	105091
	row E	1	2	3	4	5	6	7	8	9	10	11	12
	A	105094	105093	105115									
	B	105095	105105	105116									
	C	105096	105107	105118									
Source Plate D-10045	D	105097	105108	105119									
	row F	1	2	3	4	5	6	7	8	9	10	11	12
	E	105098	105109	105120									
	F	105100	105110										
	G	105102	105113										
	H	105103	105114										
	row G	1	2	3	4	5	6	7	8	9	10	11	12
	A												
	B												
	C												
	D												
	row H	1	2	3	4	5	6	7	8	9	10	11	12
	E												
	F												
	G												
	H												

Appendix B- CODM mutation screen

Pool plate	Well	Screening		Rescreening		Plant Sd numbers				DNA Source	Sequencing Result	CODM copy affected	Seed batch number of mutant	Population
		Frag 1 (bp)	Frag 2 (bp)	Frag 1 (bp)	Frag 2 (bp)									
One	A8	315	-	881	-	100057	100058	100059	100060	D-10001 A8-D8	Intron 2	-	-	-
	B1	326	-	-	-	100005	100006	100007	100008	D-10001 E1-H1	-	-	-	-
	B6	321	-	-	-	100045	100046	100047	100048	D-10001 E6-H6	-	-	-	-
	C2	355	-	-	-	100089	101206	100091	100092	D-10002 A2-D2	-	-	-	-
	C4	390	914	386	911	100105	100106	100107	100108	D-10002 A4-D4	Intron 3	-	-	-
	E1	450	-	-	-	100193	100194	100195	100196	D-10003 A1-D1	-	-	-	-
	E9	325	-	-	-	100337	100338	100339	100340	D-10003 A9-D9	-	-	-	-
	F2	416	-	-	-	100285	100286	100287	100288	D-10003 E2-H2	-	-	-	-
	F3	323	-	-	-	100293	100294	100295	100296	D-10003 E3-H3	-	-	-	-
	F4	322	-	-	-	100301	100302	100303	100304	D-10003 E4-H4	-	-	-	-
	F5	321	-	-	-	101229	100310	100311	100312	D-10003 E5-H5	-	-	-	-
	F6	323	-	-	-	100317	100318	100319	100320	D-10003 E5-H6	-	-	-	-
Two	A7	173	-	-	-	100457	101236	100459	100460	D-10005 A7-D7	-	-	-	-
	B4	778	-	795	-	100429	100430	100431	101234	D-10005 E4-H4	Y177Y silent	-	-	-
	C10	456	841	462	849	100569	100570	100571	100572	D-10006 A10-D10	Intron 2	-	-	-
	C11	426	-	-	-	100577	100578	100579	100580	D-10006 A11-D11	-	-	-	-
	H3	185	-	-	-	100765	100766	100767	100768	D-10008 E3-H3	-	-	-	-
Three	A12	392	-	-	-	100809	100810	100811	100812	D-10009 A12-D12	-	-	-	-
	D8	383	941	373	919	101195	100942	100943	100944	D-10010 E8-H8	Intron 3	-	-	-

Four	E10	912	-	406	896	101712	101714	101720	101721	D-10015 A10-D10	Intron 2	-	-	-
	F3	364	966	349	940	101493	101494	101495	101497	D-10015 E3-H3	T283I	CODMb	S-101497	B2
	F9	355	974	-	-	101701	101704	101706	101711	D-10015 E9-H9	-	-	-	-
	H1	475	-	-	-	101570	101573	101574	101575	D-10016 E1-H1	-	-	-	-
	H3	470	-	-	-	101590	101591	101592	101593	D-10016 E3-H3	-	-	-	-
	H6	328	-	-	-	101617	101618	101619	101620	D-10016 E6-H6	-	-	-	-
	H7	446	849	442	844	101626	101627	101628	101629	D-10016 E7-H7	Intron 2	-	-	-
Five	F2	802	-	528	786	101981	101982	101983	101984	D-10019 E2-H2	S176F	CODMb	S-101983	B2
	H1	493	826	483a 370b	813a 831b	101946	101947	101948	101949	D-10020 E1-H1	Intron 2, a Intron 3,b	-	-	-
	H11	135	-	-	-	102309	102310	102311	102312	D-10020 E11-H11	-	-	-	-
Six	C7	326	-	-	-	102372	102373	102374	102377	D-10022 A7-D7	-	-	-	-
	D7	327	-	-	-	102378	102379	102380	102381	D-10022 E7-H7	-	-	-	-
	F3	172	-	-	-	102436	102437	102438	102439	D-10023 E3-H3	-	-	-	-
	G7	530	788	536	784	102571	102572	102573	102574	D-10024 A7-D7	W261*	CODMa	S-102574	B2
	G8	323	-	-	-	102579	102580	102581	102582	D-10024 A8-D8	-	-	-	-
	H2	357	-	-	-	102494	102495	102496	102498	D-10024 E2-H2	-	-	-	-
	H3	785	-	535	783	102503	102504	102505	102506	D-10024 E3-H3	Y177Y silent	-	-	-
Seven	A6	323	-	-	-	102605	102606	102607	102608	D-10025 A6-D6	-	-	-	-
	B8	314	-	-	-	102625	102626	102627	102628	D-10025 E8-H8	-	-	-	-
	G11	495	813	500	807	103225	103226	103227	103228	D-10028 A11-D11	F271L	CODMb	S-103228	C1
	H6	360	-	-	-	103016	103017	103018	103019	D-10028 E6-H6	-	-	-	-

Eight	B8	336	-	426	878	103059	103060	103061	103062	D-10029 E8-H8	Intron 2	-	-	-
	B12	401	-	-	-	103091	103092	103093	103094	D-10029 E12-H12	-	-	-	-
	D4	317	1050	313	994	103123	103124	103125	103127	D-10030 E4-H4	K300K silent E301K	CODMb CODMa	S-103127	B2
	E6	372	955	336	957	103279	103282	103283	103284	D-10031 A6-D6	T283I	CODMa	S-103284	C1
	F2	442	939	586a 427b	745a 892b	103248a	103249	103250b	103251	D-10031 E2-H2	Intron 3,b	-	-	-
	F9	337		386	935	103312	103314	103315	103316	D-10031 E9-H9	Intron 3	-	-	-
	G2	523	847	505	813	103357	103358	103359	103360	D-10032 A2-D2	Intron 2	-	-	-
	G3	385	984	373	933	103366	103367	103369	103370	D-10032 A3-D3	Intron 3	-	-	-
	G4	554	814	528	769	103378	103379	103381	103382	D-10032 A4-D4	S176S silent	-	-	-
	G8	467	904	452	863	103419	103420	103421	103422	D-10032 A8-D8	Intron 2	-	-	-
	G12	556	823	533	785	103456	103458	103459	103460	D-10032 A12-D12	I262V	CODMb	S-103458	C1
	H2	604	833	489a 519b	823a 796b	103361	103362	103364	103365	D-10032 E2-H2	Intron 2, a Not found, b	-	-	-
	H4	552	817	538	788	103383	103385	103386	103387	D-10032 E4-H4	S176F	CODMb	S-103383	C1

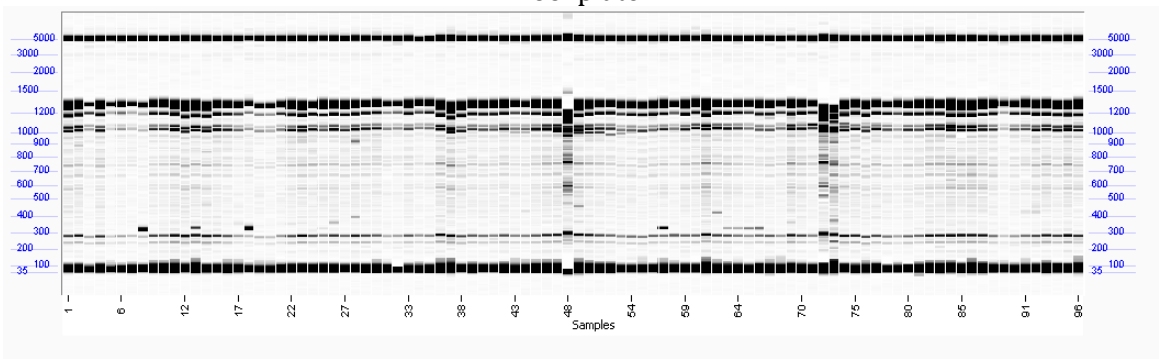
Nine	A1	444	-	422	900	103466	103467	103468	103469	D-10033 A1-D1	Intron 3	-	-	-
	A6	336	-	-	-	103517	103518	103519	103520	D-10033 A6-D6	-	-	-	-
	A7	537	817	520	786	103529	103530	103532	103534	D-10033 A7-D7	P266S	CODMb	S-103534	C1
	A9	338	-	-	-	103551	103552	103553	103555	D-10033 A9-D9	-	-	-	-
	B4	336	-	-	-	103501	103504	103505	103506	D-10033 E4-H4	-	-	-	-
	B7	336, 367,516	824	373a 513b	963a 822b	103535	103536	103537	103538	D-10033 E7-H7	Not found, a Not found, b	-	-	-
	B12	192	1166	506	816	103647	103648	103626	103627	D-10033 E12-H12	Intron 2	-	-	-
	C1	337, 435	1150	-	-	103567	103568	103569	103570	D-10034 A1-D1	-	-	-	-
	C3	450	902	436	873	103583	103584	103585	103586	D-10034 A3-D3	Intron 2	-	-	-
	C4	338	-	-	-	103593	103594	103595	103596	D-10034 A4-D4	-	-	-	-
	C6	424	934	400	875	103618	103619	103620	103621	D-10034 A6-D6	Intron 3	-	-	-
	D1	328,450,482	1009	-	-	103571	103572	103573	103574	D-10034 E1-H1	-	-	-	-
	D6	334	1019	-	-	103622	103623	103625	103539	D-10034 E6-H6	-	-	-	-
Ten	A3	414	-	-	-	103682	103683	103684	103688	D-10035 A3-D3	-	-	-	-
	B4	467	876	454	840	103781	103784	103785	103789	D-10035 E4-H4	Intron 2	-	-	-
	C2	166	-	-	-	103853	103854	103855	103856	D-10036 A2-D2	-	-	-	-
	C5	222	-	422	859	103891	103892	103893	103894	D-10036 A5-D5	Intron 3	-	-	-
	D1	129	163	-	-	103848	103849	103850	103851	D-10036 E1-H1	-	-	-	-
	D9	383	-	-	-	104047	104048	104052	104055	D-10036 E9-H9	-	-	-	-
	F3	828	-	506	780	103923	103925	103926	103927	D-10037 E3-H3	Not found	-	-	-
	F4	425	935	417	881	103934	103935	103936	103937	D-10037 E4-H4	Intron 3	-	-	-
	F5	195	-	-	-	103944	103945	103946	103947	D-10037 E5-H5	-	-	-	-
	G11	644	731	622	705	104359	104281	104282	104283	D-10038 A11-D11	L232F	CODMb	S-104281	B3
	H1	170	-	-	-	104130	104132	104133	104134	D-10038 E1-H1	-	-	-	-
	H5	186	386	-	-	104232	104234	104235	104237	D-10038 E5-H5	-	-	-	-
	H9	400	936	383	889	104331	104340	104321	104322	D-10038 E9-H9	Intron 3	-	-	-

Eleven	A1	483	849	472	821	104191	104192	104193	104194	D-10039 A1-D1	Intron 2	-	-	-
	A2	372	-	-	-	104269	104270	104271	104274	D-10039 A2-D2	-	-	-	-
	A5	330	-	-	-	104391	104392	104393	104395	D-10039 A5-D5	-	-	-	-
	A10	573	787	548	754	104529	104530	104531	104532	D-10039 A10-D10	Q254*	CODMb	S-104530	B3
	A12	370	971	362	939	104549	104550	104551	104553	D-10039 A12-D12	Intron 2	-	-	-
	B8	328	889	-	-	104594	104595	104596	104597	D-10039 E8-H8	-	-	-	-
	B12	434	919	421	884	104554	104555	104556	104557	D-10039 E12-H12	Intron 2	-	-	-
	C2	327	-	-	-	104301	104307	104308	104310	D-10040 A2-D2	-	-	-	-
	C6	223	334	-	-	104504	104507	104511	104512	D-10040 A6-D6	-	-	-	-
	D5	355	-	-	-	104500	104501	104502	104503	D-10040 E5-H5	-	-	-	-
	D8	336	-	-	-	104693	104698	104699	104700	D-10040 E8-H8	-	-	-	-
	F2	393	971	378	930	104737	104738	104739	104741	D-10041 E2-H2	Intron 3	-	-	-
	F6	404	-	346	952	104665	104666	104667	104668	D-10041 E6-H6	G285G silent	-	-	-
	F9	336	-	-	-	104861	104862	104864	104865	D-10041 E9-H9	-	-	-	-
	G2	333	-	414	888	104888	104889	104891	104892	D-10042 A2-D2	Not found	-	-	-
	H5	342,456	599	-	-	104766	104767	104770	104773	D-10042 E5-H5	-	-	-	-
Twelve	A6	329	-	-	-	104958	104959	104960	104961	D-10043 A6-D6	-	-	-	-
	B3	515	-	-	-	104932	104933	104935	104936	D-10043 E3-H3	-	-	-	-
	B6	265	-	-	-	104962	104963	104964	104965	D-10043 E6-H6	-	-	-	-
	B9	334,393	-	-	-	104992	104993	104994	104995	D-10043 E9-H9	-	-	-	-
	C5	467	895	451	863	105065	105066	105068	105069	D-10044 A5-D5	Intron 2	-	-	-
	C8	673	-	-	-	105134	105135	105137	105138	D-10044 A8-D8	-	-	-	-
	D12	599	696	-	-	105085	105086	105089	105091	D-10044 E12-H12	-	-	-	-

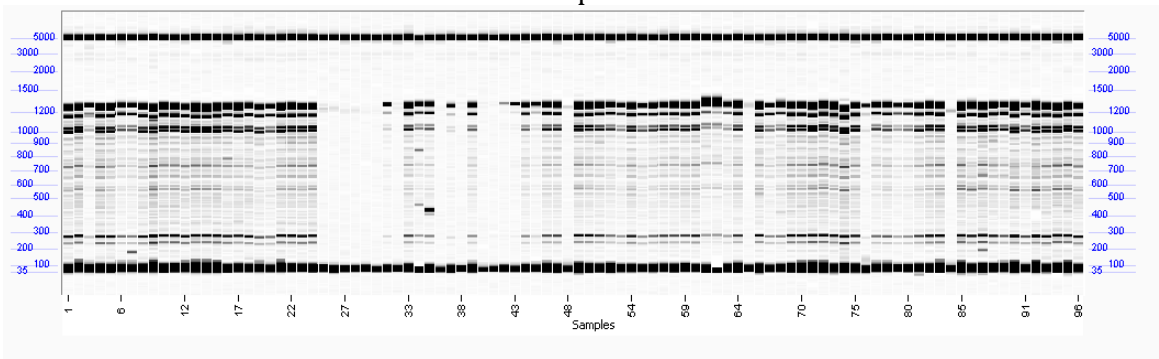
TILLING Gel images

For each gel image lanes 1-12 = wells A1-A12, lanes 13-24=wells B1-B12 etc.

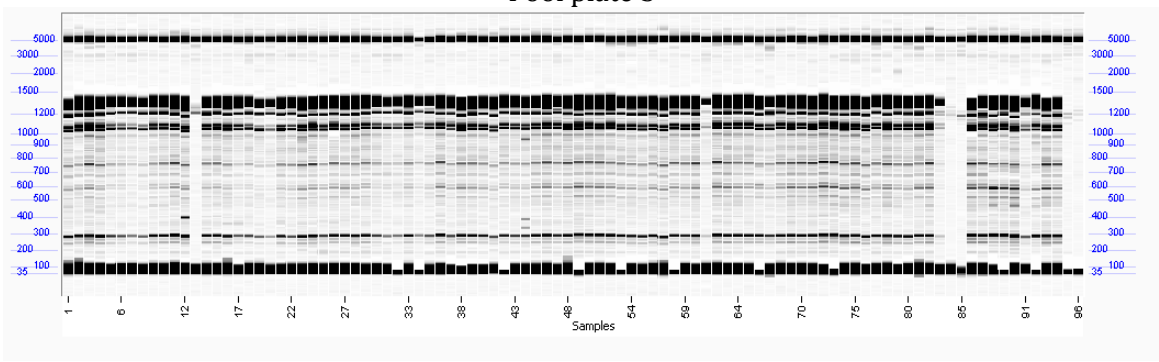
Pool plate 1



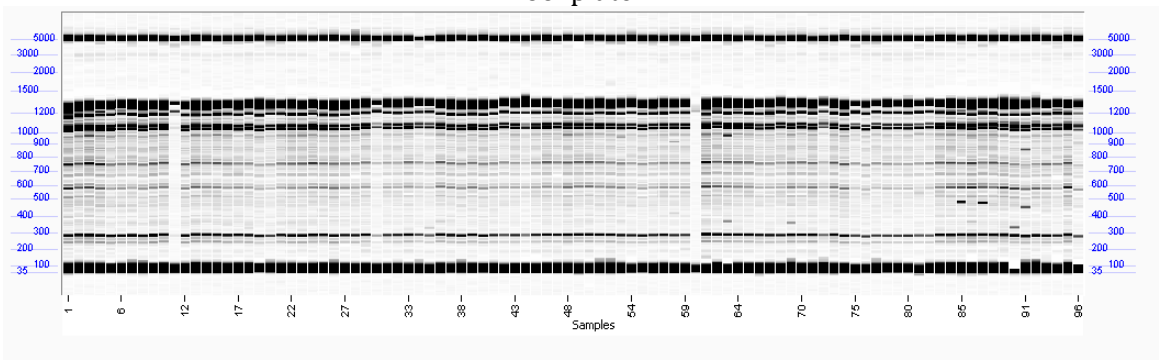
Pool plate 2



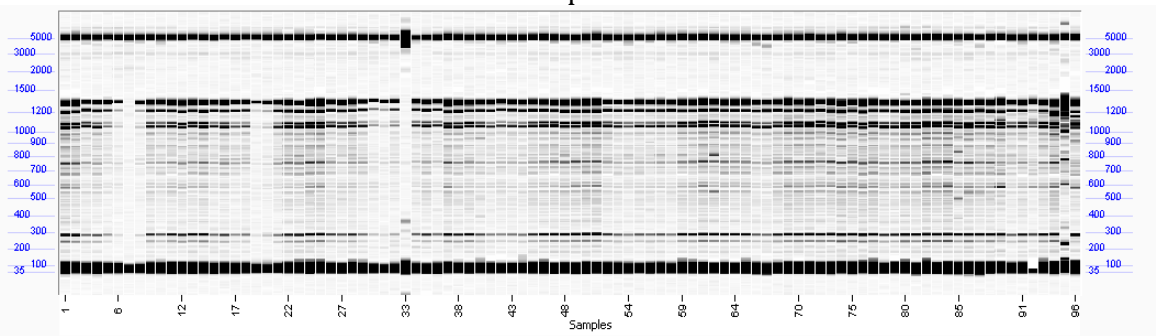
Pool plate 3



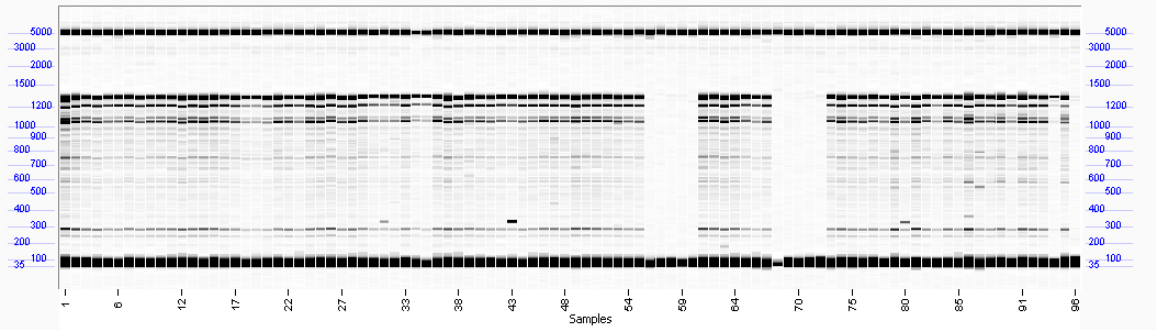
Pool plate 4



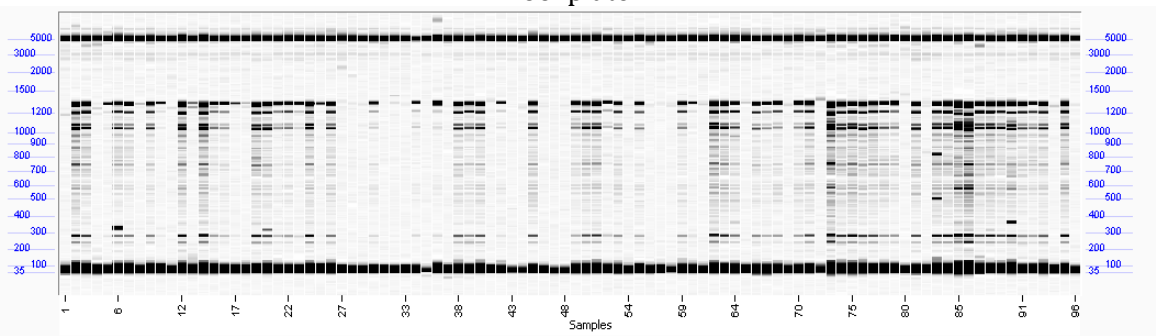
Pool plate 5



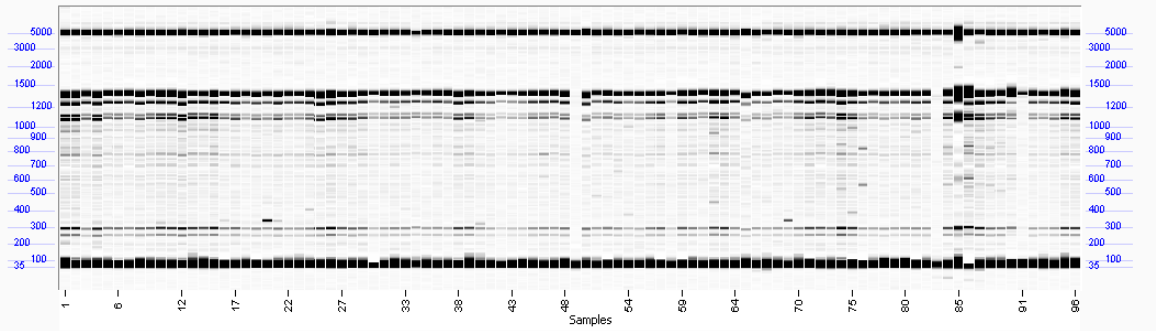
Pool plate 6



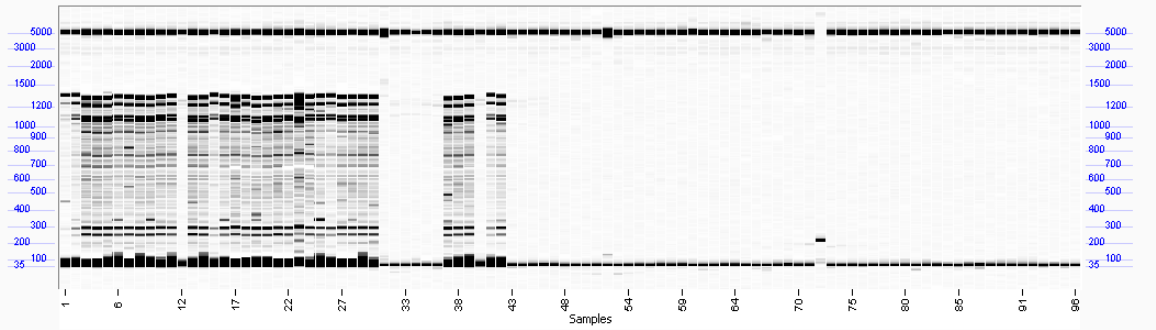
Pool plate 7



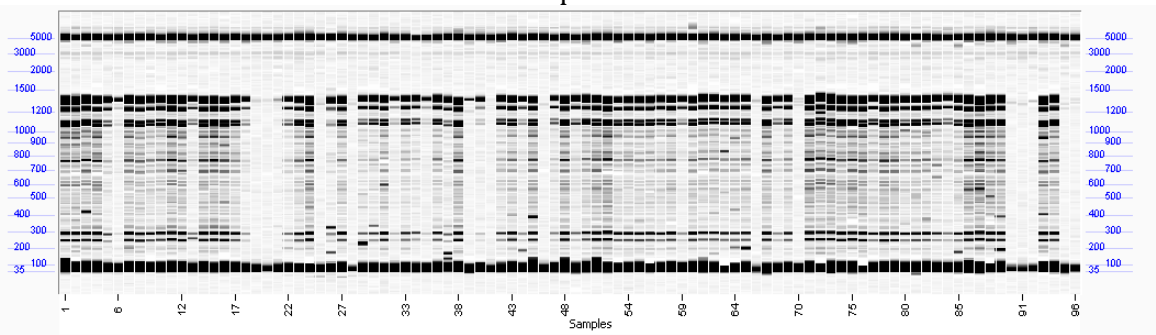
Pool plate 8



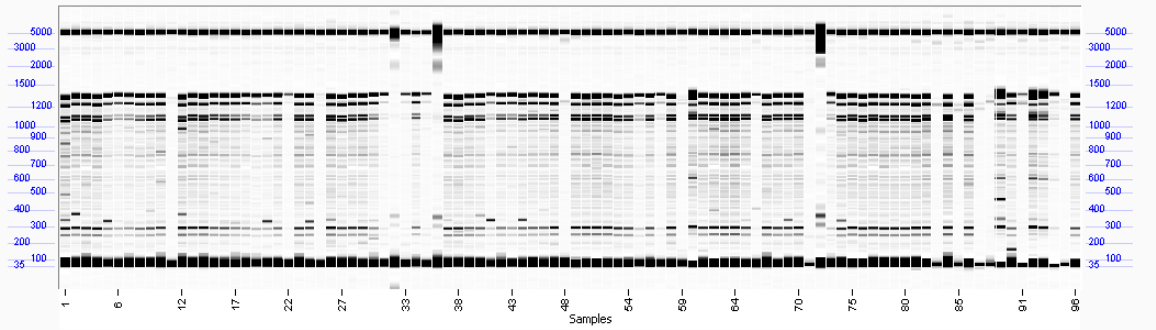
Pool plate 9



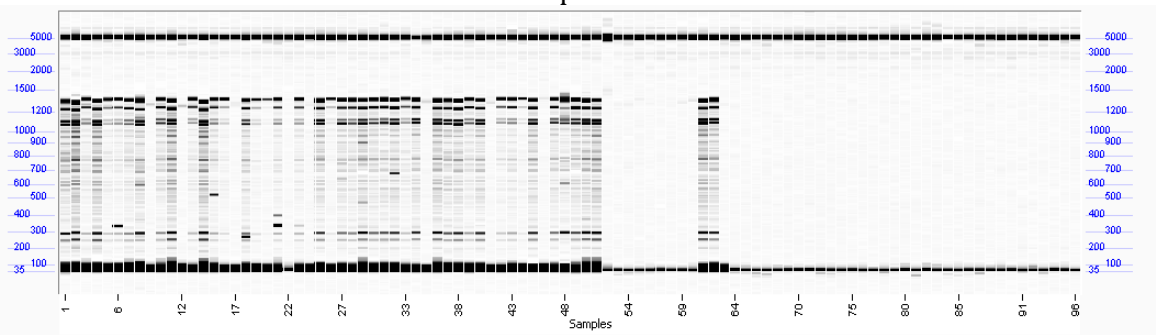
Pool plate 10



Pool plate 11

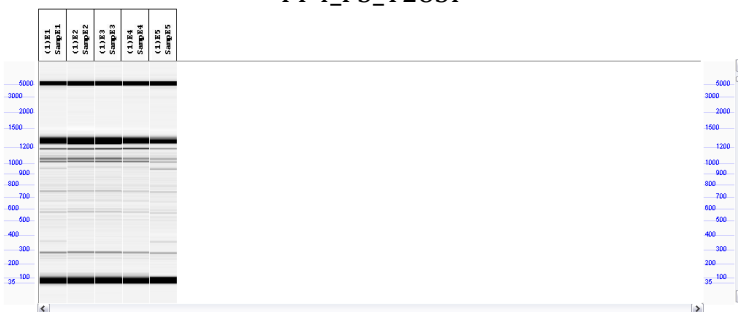


Pool plate 12



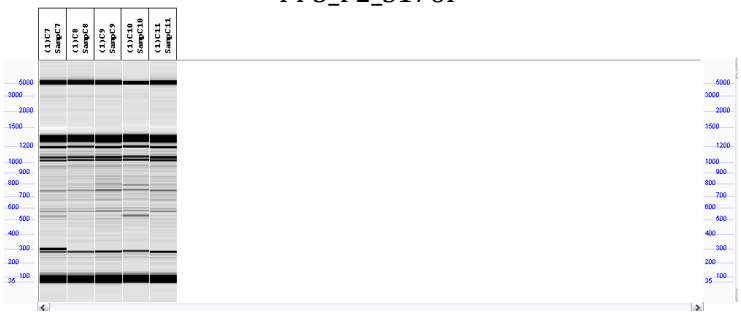
Rescreen of mutations that would later be identified as missense and nonsense mutations

PP4_F3_T283I



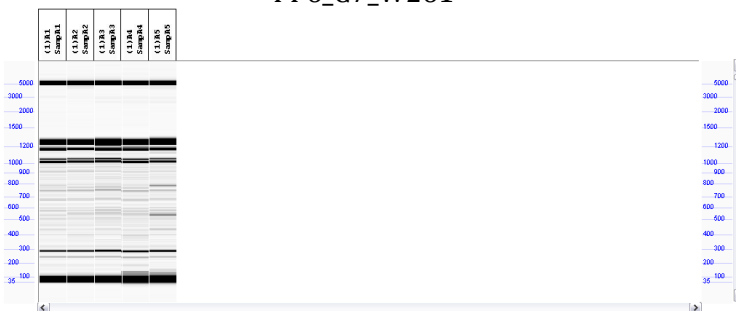
349 bp and 940 bp bands in the fourth individual (Sd-101497).

PP5_F2_S176F



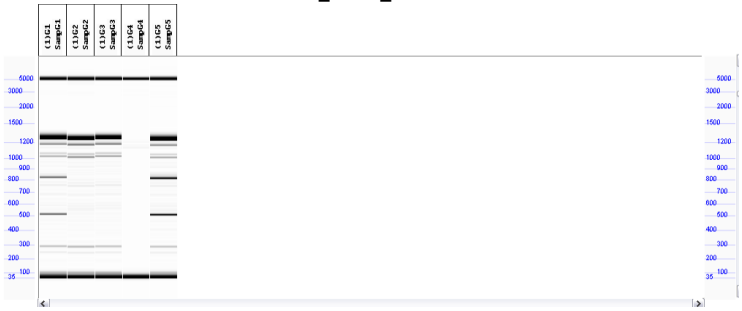
528 bp and 786 bp bands in third individual (Sd-101983).

PP6_G7_W261*



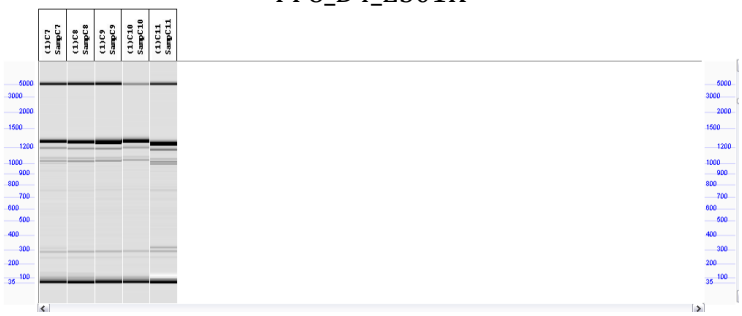
536 bp and 784 bp bands in fourth individual (Sd-102574)

PP7_G11_F271L



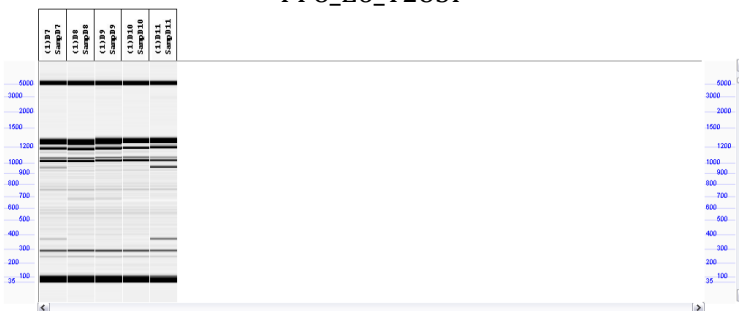
500 bp and 807 bp bands in fourth individual (Sd-103228)

PP8_D4_E301K



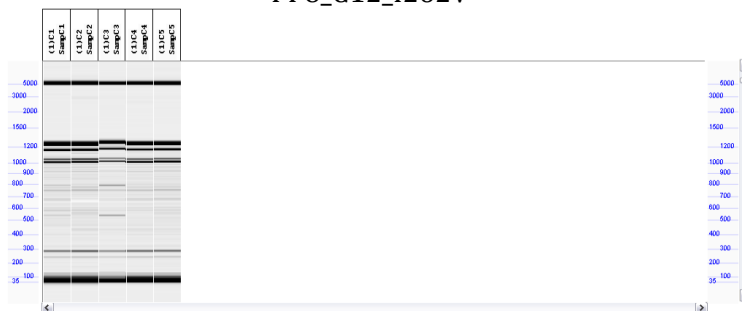
313 bp and 994 bp bands in the fourth individual (Sd-103127)

PP8_E6_T283I



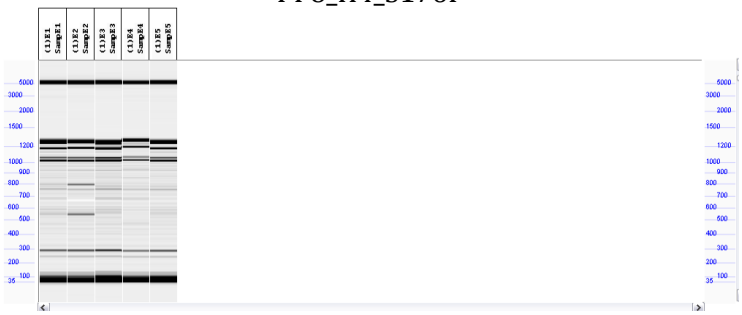
336 bp and 957 bp bands in the fourth individual (Sd-103284)

PP8_G12_I262V



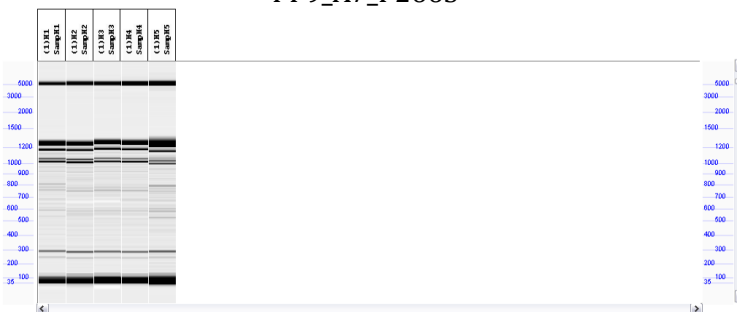
533 bp and 785 bp bands in second individual (Sd-103458)

PP8_H4_S176F



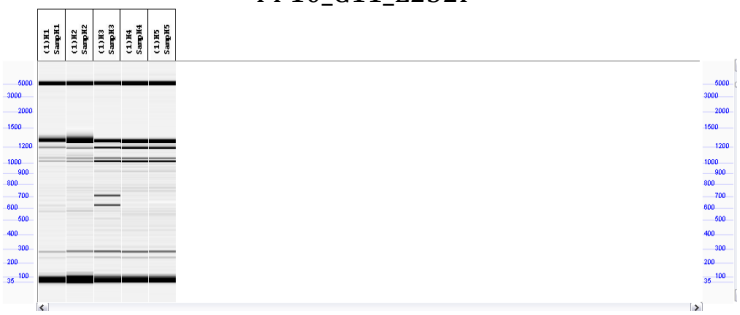
538 bp and 788 bp bands in the first individual (Sd-103383)

PP9_A7_P266S



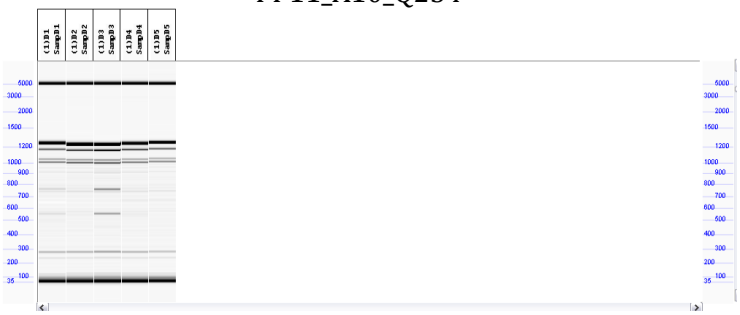
520 bp and 786 bp bands in fourth individual (Sd-103534)

PP10_G11_L232F



622 bp and 705 bp bands in second individual (Sd-104281)

PP11_A10_Q254*



548 bp and 754 bp bands in second individual (Sd-104530)

Location of CODM mutations in nucleotide and amino acid sequences

```
M   E   T   P   I   L   I   K   L   G   N   G   L   S   I   P   S   V      18
atg gag aca cca ata ctt atc aag cta ggc aat ggt ttg tca ata cca agt gtt    54

Q   E   L   A   K   L   T   L   A   E   I   P   S   R   Y   T   C   T      36

cag gaa ttg gct aaa ctc acg ctt gca gaa att cca tct cga tac aca tgc acc   108

G   E   S   P   L   N   N   I   G   A   S   V   T   D   D   E   T   V      54

ggt gaa agc ccg ttg aat aat att ggt gcg tct gta aca gat gat gaa aca gtt   162

P   V   I   D   L   Q   N   L   L   S   P   E   P   V   V   G   K   L      72

cct gtc atc gat ttg caa aat tta cta tct cca gaa ccc gta gtt gga aag tta   216
```

IPB002283A (4.5e-13) IC 2.46

E L D K L H S A C K E W G F F Q | L 89

gaa ttg gat aag ctt cat tct gct tgc aaa gaa tgg ggt ttc ttt cag | ctg 267

IPB002283B (4.7e-05) IC 1.94

V N H G V D A L L M D N I K S E I K 107

gtt aac cat gga gtc gac gct tta ctg atg gac aat ata aaa tca gaa att aaa 321

G F F N L P M N E K T K Y G Q Q D G 125

ggt ttc ttt aac ctt cca atg aat gag aaa act aaa tac gga cag caa gat gga 375

D F E G F G Q P Y I E S E D Q R L D 143

gat ttt gaa gga ttt gga caa ccc tat att gaa tcg gag gac caa aga ctt gat 429

W T E V F S M L S L P L H L R K P H 161

tgg act gaa gtg ttt agc atg tta agt ctt cct ctc cat tta agg aag cct cat 483

L F P E L P L P F R | E T L E S Y L 178

ttg ttt cca gaa ctc cct ctg cct ttc ag | g gag aca ctg gaa tcc tac cta 534

tS176F

tS176F

S K M K K L S T V V F E M L E K S L 196

tca aaa atg aaa aaa cta tca acg gtt gtc ttt gag atg ttg gaa aaa tct cta 588

Q L V E I K G M T D L F E D G L Q T 214

caa tta gtt gag att aaa ggt atg aca gac tta ttt gaa gat ggg ttg caa aca 642

M R M N Y Y P P C P R P E L V L G L 232

atg agg atg aac tat tat cct cct tgt cct cga cca gag ctt gta ctt ggt ctt 696

tL232F

IPB002283F (5.8e-06) IC 2.81

T S H
S D F S G L T I L L Q L N E V 250

acg tca cac tcg gat ttt agc ggt ttg aca att ctc ctt caa ctt aat gaa gtt 750

E G L Q I R K E E R W I S I K P L P 268

gaa gga tta caa ata aga aaa gaa gag agg tgg att tca atc aaa cct cta cct 804

tQ254* **gI262V** **tP266S**

aW261*

IPB002283G (1.6e-16) IC 2.76

D A F I V N V G D I L E |I M T N G 285

gat gcg ttc ata gtg aat gtt gga gac att ttg gag | ata atg act aat ggg 855

gF271L **tT283I**

tT283I

IPB002283H (6.0e-08) IC 2.25

I Y R S V E H R A V V N S T K E R L 303

att tac cgt agc gtc gag cac cgg gca gta gta aac tca aca aag gag agg ctc 909

aE301K

S I A T F H D S K L E S E I G P I S 321

tca atc gcg aca ttt cat gac tct aaa cta gag tca gaa ata ggc cca att tcg 963

S L V T P E T P A L F K R G R Y E D 339

agc ttg gtc aca cca gag aca cct gct ttg ttc aaa aga ggt agg tat gag gat 1017

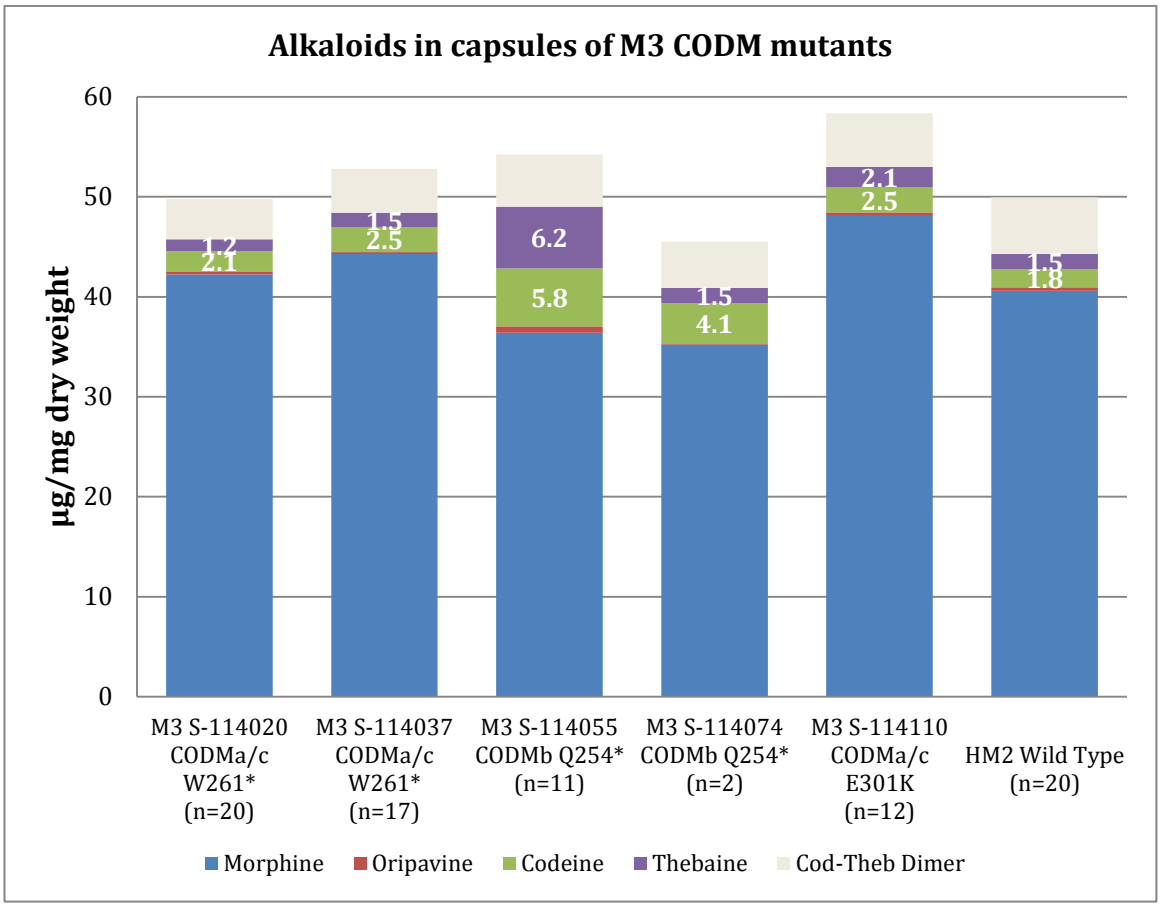
I L K E N L S R K L D G K S F L D Y 357

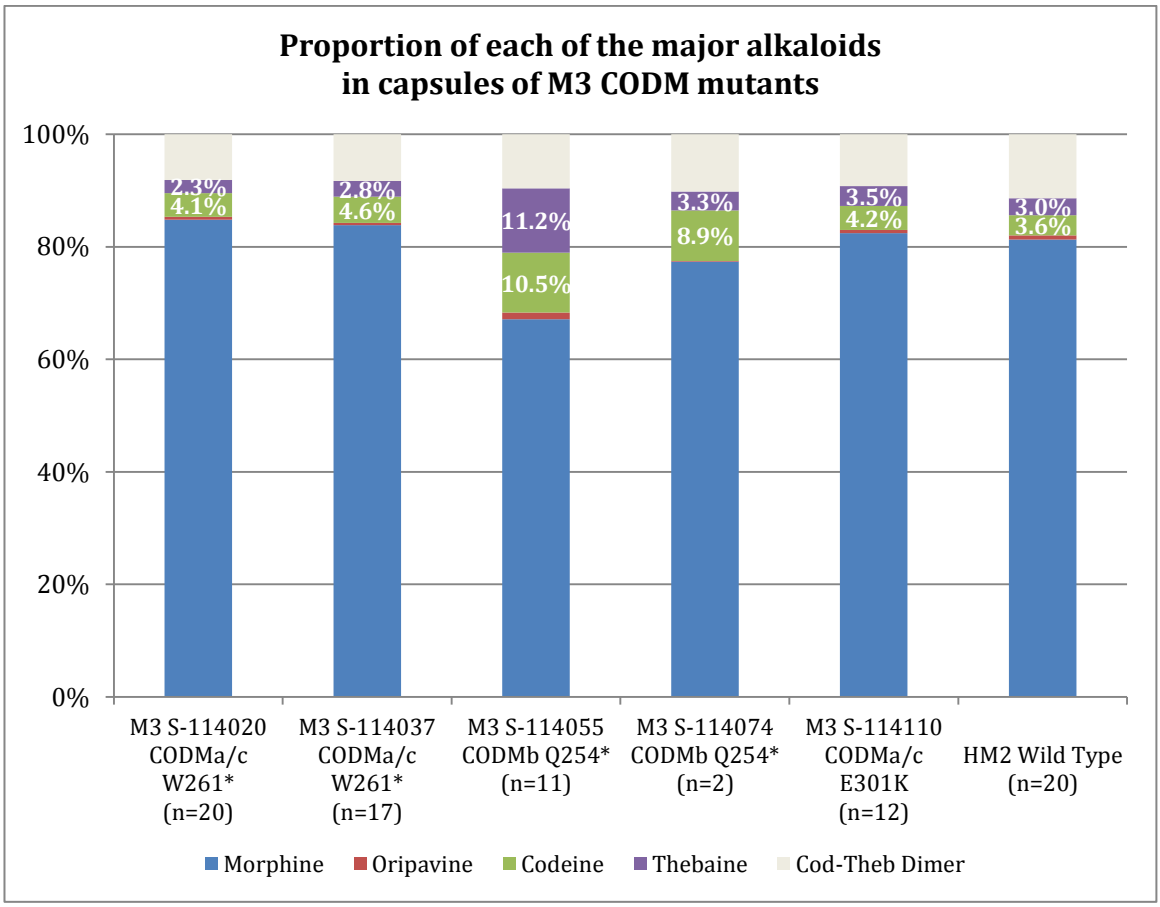
att ttg aag gaa aat ctt tca agg aag ctt gat gga aaa tca ttt ctc gac tac 1071

M R M * 361

atg agg atg tga 1083

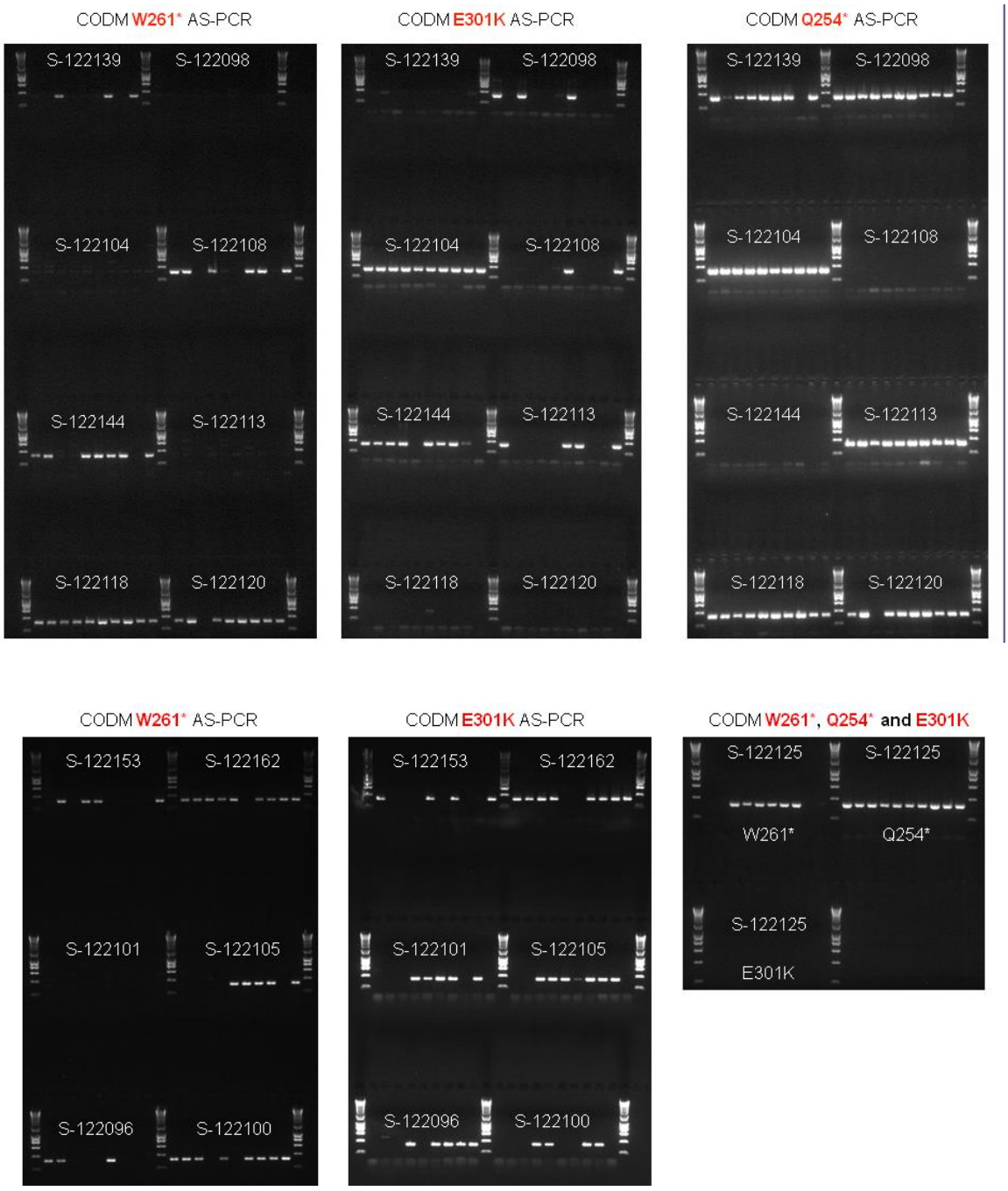
Alkaloid content of CODM M3 mutants





Genotyping of F1s resulting from crosses between M3 CODM mutants

Bands indicate the presence of the appropriate CODM mutation.



F1 plants carrying both W261* and E301K CODM mutations			F1 plants carrying both Q254* and E301K CODM mutations			F1 plants carrying both W261* and Q254* CODM mutations		
F1 Seed batch id	Seedling id	F2 seed batch id	F1 Seed batch id	Seedling id	F2 seed batch id	F1 Seed batch id	Seedling id	F2 seed batch id
S-122153	Sd-831374	S-123026	S-122098	Sd-831390	S-123047*	S-122139	Sd-831452	S-123064
	Sd-831379	S-123027*		Sd-831392	S-123048		Sd-831457	S-123065
S-122162	Sd-831380	S-123028		Sd-831396	S-123049		Sd-831459	S-123066
	Sd-831381	S-123029	S-122104	Sd-831400	S-123051	S-122118	Sd-831420	S-123067
	Sd-831382	**		Sd-831401	S-123050		Sd-831421	S-123068
	Sd-831383	S-123030		Sd-831402	S-123052		Sd-831422	S-123069
	Sd-831386	S-123031		Sd-831403	S-123053*		Sd-831423	S-123070*
	Sd-831387	S-123032		Sd-831404	S-123054*		Sd-831424	**
	Sd-831388	S-123033		Sd-831405	S-123055		Sd-831425	S-123071
	Sd-831389	S-123034		Sd-831406	S-123056		Sd-831426	S-123072
S-122101	None***	**		Sd-831407	S-123057		Sd-831427	S-123073*
S-122105	Sd-831344	S-123035		Sd-831408	S-123058		Sd-831428	**
	Sd-831346	S-123036		Sd-831409	S-123059		Sd-831429	S-123074
	Sd-831347	S-123037	S-122113	Sd-831410	S-123060	S-122120	Sd-831430	S-123075
S-122096	Sd-831316	S-123038		Sd-831415	S-123061		Sd-831431	S-123076
S-122100	Sd-831324	S-123039		Sd-831416	S-123062		Sd-831433	S-123077*
	Sd-831327	S-123040		Sd-831419	S-123063		Sd-831434	S-123078
	Sd-831328	S-123041					Sd-831435	S-123079
S-122108	Sd-831359	S-123042					Sd-831436	S-123080
S-122144	Sd-831360	**					Sd-831437	S-123081*
	Sd-831361	S-123043					Sd-831438	**
	Sd-831365	S-123044					Sd-831439	S-123082
	Sd-831366	S-123045				S-122125	Sd-831442	S-123083
	Sd-831367	S-123046					Sd-831443	S-123084*
							Sd-831444	S-123085
* Very little or poor quality seed recovered							Sd-831445	S-123086
** No seed recovered							Sd-831446	S-123087*
*** One parent WT CODM							Sd-831447	S-123088

Genotyping F2s resulting from W261 x Q254* crosses*

W261* screen Q254* screen

Plate 1

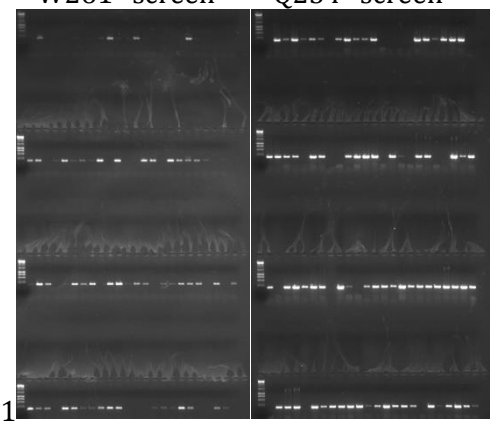
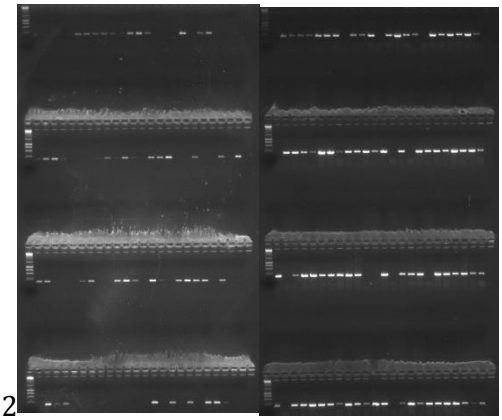
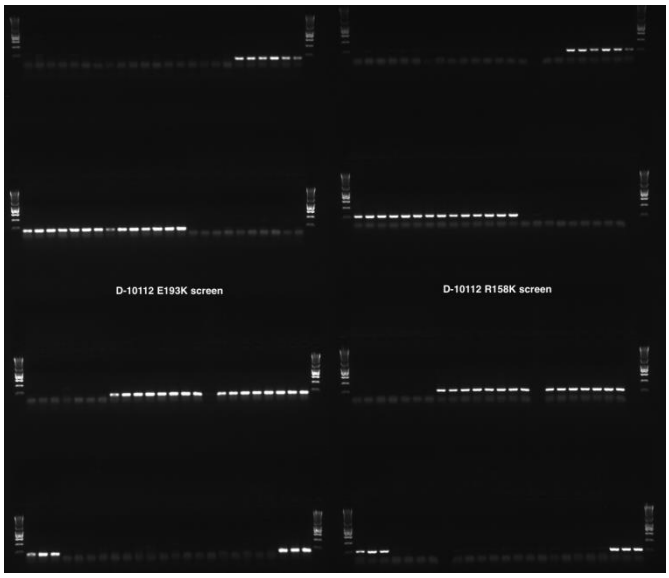


Plate 2



Detection of E193K and R158K CODM polymorphisms in remaining two high codeine forward screen lines S-114247 and S-114248

Plate D-10112 contained DNA from the remaining two (of five) high codeine forward screen lines plus DNA from two high thebaine forward screen lines (S-114249 and S-114250), wild type HM2 DNA and three positive controls i.e. one example of each a plant from the other three high codeine forward screen lines where the E193K and R158K polymorphisms were first identified. The products generated after E193K and R158K AS-PCR on the entire plate are displayed. The final 6 samples of row 1 and the first 14 on row 2 from plants generated with S-114248. The end of row 3 and the first 3 of row 4 are samples from plants generated with S-114247. Both these high codeine lines, then possess the same CODM polymorphisms detected in S-122066, S-122067 and S-122068. The final 3 samples in row 4 are the positive controls. No AS-PCR products were generated with samples from the high thebaine forward screen lines (S-114249 and S-114250) and HM2 wild type indicating the specificity of the reaction for plants with these polymorphisms.



Genealogies of high codeine forward screen lines

Phenotypes expressed as percentage dry weight of morphine (M), codeine (C), oripavine (O) and thebaine (T).

MAS 2235 D

S-114247

popn	season	old ID	seed ID		plot no	M	C	O	T
				OP row	**HM2**	**3.1117**			
C1M2	07_08	**25c**	**S-103234**	**OP row**	MUT 437	2.9664	0.2557	0.0150	0.2019
				SP caps	MUTS 437 G	4.4067	0.5943	0.0060	0.0769
C1M3	08_09			**OP row**	**HM3**	**3.67**			
				OP row	RMUU 314	3.7286	0.6138	0.0011	0.2387
				SP caps	RMUUS 314 E	4.2614	0.6006	0.0136	0.0813
C1M4	09_10			**OP row**	**HM1**	**2.918**	**0.332**	**0.027**	**0.121**
				OP row	MUV 107	2.352	1.391	0.051	1.105
				SP caps	MUVS 107 A	2.315	2.316	0.024	1.209
C1M5	2010	**Seed from MUVS 107 A sent to Tom Davies to grow over Australian winter. Seed harvested and returned labelled UKMUVS 107 A**							
					Seed sown into 5 replicated rows in TAS				

						M+pM	C	O	T	M
C1M6	10_11			**OP row**	**HT4**	**1.361**	**0.541**	**0.23**	**1.862**	**1.163**
			REP 2/5	**OP row**	**HM4**	**2.569**	**0.144**	**0.015**	**0.105**	**2.352**
			ex ukmuvs 107a	**OP row**	MUWT 1926	2.264	1.221	0.030	0.915	2.053
			(C1M5 2010)	**SP caps**	MUWTS 1926 C	2.607	2.200	0.065	2.086	2.345

					M	C	O	T
C1M7	11_12							
		REP 2/2	**OP row**	**HM5**	**3.205**	**0.134**	**0.016**	**0.116**
		ex MUWTS 1926 C	**OP row**	**HT5**	**1.233**	**0.222**	**0.654**	**2.413**
		(C1M6 10-11)	**OP row**	MA 2235	2.868	1.369	0.035	0.981
			SP caps	**MAS 2235 D**	3.305	2.174	0.045	1.434

MAS 2237 C

S-114248

popn	season	old ID	seed ID		plot no	M	C	O	T
				OP row	**HM2**	**3.1117**			
C1M2	07_08	**25c**	**S-103234**	**OP row**	MUT 437	2.9664	0.2557	0.0150	0.2019
				SP caps	MUTS 437 G	4.4067	0.5943	0.0060	0.0769
C1M3	08_09			**OP row**	**HM3**	**3.67**			
				OP row	RMUU 314	3.7286	0.6138	0.0011	0.2387
				SP caps	RMUUS 314 E	4.2614	0.6006	0.0136	0.0813
C1M4	09_10			**OP row**	**HM1**	**2.918**	**0.332**	**0.027**	**0.121**
				OP row	MUV 107	2.352	1.391	0.051	1.105
				SP caps	MUVS 107 A	2.315	2.316	0.024	1.209
C1M5	2010	**Seed from MUVS 107 A sent to Tom Davies to grow over Australian winter. Seed harvested and returned labelled UKMUVS 107 A**			Seed sown into 5 replicated rows in TAS				

						M+pM	C	O	T	M
C1M6	10_11			**OP row**	**HT4**	**1.361**	**0.541**	**0.23**	**1.862**	**1.163**
			REP 3/5	**OP row**	**HM4**	**2.569**	**0.144**	**0.015**	**0.105**	**2.352**
			ex ukmuvs 107a	**OP row**	MUWT 1945	1.755	1.485	0.027	1.307	1.577
			(C1M5 2010)	**SP caps**	MUWTS 1945 D	1.967	2.274	0.066	1.669	1.742

					M	C	O	T
C1M7	11_12							
		REP 2/2	**OP row**	**HM5**	**3.205**	**0.134**	**0.016**	**0.116**
		ex MUWTS 1945 D	**OP row**	**HT5**	**1.233**	**0.222**	**0.654**	**2.413**
		(C1M6 10-11)	**OP row**	MA 2237	2.800	1.409	0.027	1.035
			SP caps	**MAS 2237 C**	3.332	2.270	0.034	0.908

MAS 1833 B

S-122067

popn	season	old ID	seed ID		plot no	M	C	O	T
				OP row	**HM2**	**3.1117**			
C1M2	07_08	**25c**	**S-103234**	**OP row**	MUT 437	2.9664	0.2557	0.0150	0.2019
				SP caps	MUTS 437 G	4.4067	0.5943	0.0060	0.0769

						M	C	O	T
C1M3	08_09			**OP row**	**HM3**	**3.67**			
				OP row	RMUU 314	3.7286	0.6138	0.0011	0.2387
				SP caps	RMUUS 314 E	4.2614	0.6006	0.0136	0.0813

C1M4 **2009** **Seed from RMUUS 314 E sent to Tom Davies to grow over Australian winter. Seed harvested and returned labelled UK MUUS 21**

					M	C	O	T
C1M5	09_10	ex uk muus 21	**OP row**	**HM1**	**3.12**	**0.36**	**0.04**	**0.17**
		(C1M4 2009)	**OP row**	TMUW 125	2.201	1.711	0.068	1.327
			SP caps	TMUWS 125 A	3.276	1.792	0.027	0.486

					M+pM	C	O	T	M
C1M6	10_11	ex tmuws 125a	**OP row**	**HM4**	**2.569**	**0.144**	**0.015**	**0.105**	**2.352**
		(C1M5 09-10)	**OP row**	**HT4**	**1.361**	**0.541**	**0.23**	**1.862**	**1.163**
			OP row	MUX 1836	2.411	1.193	0.052	1.135	2.267
			SP caps	MUXS 1836 B	3.918	1.266	0.061	0.296	3.744

					M	C	O	T	Total alk
C1M7	11_12	ex muxs 1836 B	**OP row**	**HM5**	**3.205**	**0.134**	**0.016**	**0.116**	**3.472**
		(C1M6 10-11)	**OP row**	**HT5**	**1.233**	**0.222**	**0.654**	**2.413**	**4.521**
			OP row	MA 1833	2.498	0.973	0.070	1.613	5.155
			SP caps	**MAS 1833 B**	3.212	1.444	0.115	2.558	7.329

MAS 1835 D

S-122068

popn	season	old ID	seed ID		plot no	M	C	O	T
				OP row	**HM2**	**3.1117**			
C1M2	07_08	**25c**	**S-103234**	**OP row**	MUT 437	2.9664	0.2557	0.0150	0.2019
				SP caps	MUTS 437 G	4.4067	0.5943	0.0060	0.0769
C1M3	08_09			**OP row**	**HM3**	**3.67**			
				OP row	RMUU 314	3.7286	0.6138	0.0011	0.2387
				SP caps	RMUUS 314 E	4.2614	0.6006	0.0136	0.0813
C1M4	09_10			**OP row**	**HM1**	**2.918**	**0.332**	**0.027**	**0.121**
				OP row	MUV 107	2.352	1.391	0.051	1.105
				SP caps	MUVS 107 A	2.315	2.316	0.024	1.209

C1M5 **2010** **Seed from MUVS 107 A sent to Tom Davies to grow over Australian winter. Seed harvested and returned labelled UKMUVS 107 A**

Seed sown into 5 replicated rows in TAS

						M+pM	C	O	T	M
		REP 2/5								
C1M6	10_11	ex ukmuvs 107a		**OP row**	**HM4**	**2.569**	**0.144**	**0.015**	**0.105**	**2.352**
		(C1M5 2010)		**OP row**	**HT4**	**1.361**	**0.541**	**0.23**	**1.862**	**1.163**
				OP row	MUWT 1926	2.264	1.221	0.030	0.915	2.053
				SP caps	MUWTS 1926 C	2.607	2.200	0.065	2.086	2.345
C1M7	11_12	ex muwts 1926 C				M	C	O	T	Total alk
		(C1M6 10-11)		**OP row**	**HM5**	**3.205**	**0.134**	**0.016**	**0.116**	**3.472**
				OP row	**HT5**	**1.233**	**0.222**	**0.654**	**2.413**	**4.521**
				OP row	MA 1835	2.806	1.477	0.046	1.558	5.886
				SP caps	**MAS 1835 D**	2.661	1.944	0.037	1.230	5.872

MAS 1983 C

S-122066

popn	season	old ID	seed ID			plot no	M	C	O	T
				OP row		**HM2**	**3.1117**			
C1M2	07_08	**25c**	**S-103234**	**OP row**		MUT 437	2.9664	0.2557	0.0150	0.2019
				SP caps		MUTS 437 G	4.4067	0.5943	0.0060	0.0769
C1M3	08_09			**OP row**		**HM3**	**3.67**			
				OP row		RMUU 314	3.7286	0.6138	0.0011	0.2387
				SP caps		RMUUS 314 E	4.2614	0.6006	0.0136	0.0813
C1M4	09_10			**OP row**		**HM1**	**2.918**	**0.332**	**0.027**	**0.121**
				OP row		MUV 107	2.352	1.391	0.051	1.105
				SP caps		MUVS 107 A	2.315	2.316	0.024	1.209

C1M5 **2010** **Seed from MUVS 107 A sent to Tom Davies to grow over Australian winter. Seed harvested and returned labelled UKMUVS 107 A**

Seed sown into 5 replicated rows in TAS

						M+pM	C	O	T	M
C1M6	10_11	**REP 4/5**		**OP row**	**HM4**	**2.569**	**0.144**	**0.015**	**0.105**	**2.352**
		ex ukmuvs 107a		**OP row**	**HT4**	**1.361**	**0.541**	**0.23**	**1.862**	**1.163**
		(C1M5 2010)		**OP row**	MUWT 1964	2.356	0.709	0.017	0.429	2.073
				SP caps	MUWTS 1964 C	2.266	2.001	0.059	0.696	2.154

						M	C	O	T	Total alk
C1M7	11_12	ex muwts 1964 C		**OP row**	**HM5**	**3.373**	**0.129**	**0.017**	**0.116**	**3.634**
		(C1M6 10-11)		**OP row**	**HT5**	**1.001**	**0.189**	**0.709**	**2.473**	**4.372**
				OP row	MA 1983	3.005	1.691	0.040	1.316	6.052
				SP caps	**MAS 1983 C**	3.766	2.089	0.076	1.379	7.309

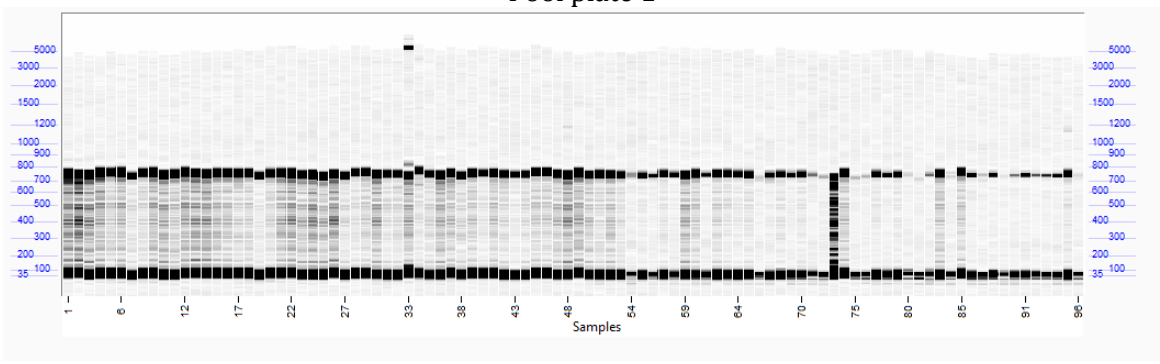
Appendix C- T6ODMa mutation screen

Pool plate	Well	Screening		Rescreening		Plant Sd numbers				DNA Source	Sequencing Result	Seed batch number of mutant	Population
		Fragment 1 (bp)	Fragment 2 (bp)	Fragment 1 (bp)	Fragment 2 (bp)								
One	C9	80	615	-	-	100145	100146	100147	101212	D-10002 A9-D9	-	-	-
Two	-	-	-	-	-	-	-	-	-	-	-	-	-
Three	-	-	-	-	-	-	-	-	-	-	-	-	-
Four	B1	612	-	-	-	101258	101259	101261	101263	D-10013 E1-H1	-	-	-
	B12	615	-	-	-	101352	101353	101357	101358	D-10013 E12-H12	-	-	-
	D11	253	-	-	-	101460	101461	101462	101463	D-10014 E11-H11	-	-	-
	D12	276	-	-	-	101468	101469	101470	101471	D-10014 E12-H12	-	-	-
	E2	219	611	-	-	101480	101481	101482	101483	D-10015 A2-D2	-	-	-
	F3	253	-	-	-	101493	101494	101495	101497	D-10015 E3-H3	-	-	-
Five	B4	612	-	-	-	101776	101777	101778	101779	D-10017 E4-H4	-	-	-
Six	F7	132	-	-	-	102474	102475	102476	102477	D-10023 E7-H7	-	-	-
Seven	B11	316	611	161	450	102690	102691	102692	102693	D-10025 E11-H11	Q141*	S-102693	B2
Eight	A12	58	612	103	509	103087	103088	103089	103090	D-10029 A12-D12	H162Y	S-103090	B2
	C9	43	609	147	468	103160	103161	103162	103163	D-10030 A9-D9	A146T	S-103161	B2
	E10	471	584	254	367	103317	103318	103319	103320	D-10031 A10-D10	F110F silent	-	C1
	F3	563	606	181	438	103257	103258	103259	103260	D-10031 E3-H3	G134G silent	-	C1
	G6	107	611	115	490	103397	103398	103399	103402	D-10032 A6-D6	L156L silent	-	C1
	G8	167	609	156	473	103419	103420	103421	103422	D-10032 A8-D8	A146T	S-103421	C1
Nine	-	-	-	-	-	-	-	-	-	-	-	-	-
Ten	D12	504	565	283	345	103991	103992	103997	103999	D-10036 E12-H12	Intron 1	-	B3
	F12	498	-	-	-	104175	104180	104187	104188	D-10037 E12-H12	-	-	-
Eleven	A12	219	609	151	468	104549	104550	104551	104553	D-10039 A12-D12	W145*	S-104553	B3
	H12	54	617	-	-	105186	105187	105188	105189	D-10042 E12-H12	-	-	-
Twelve	-	-	-	-	-	-	-	-	-	-	-	-	-

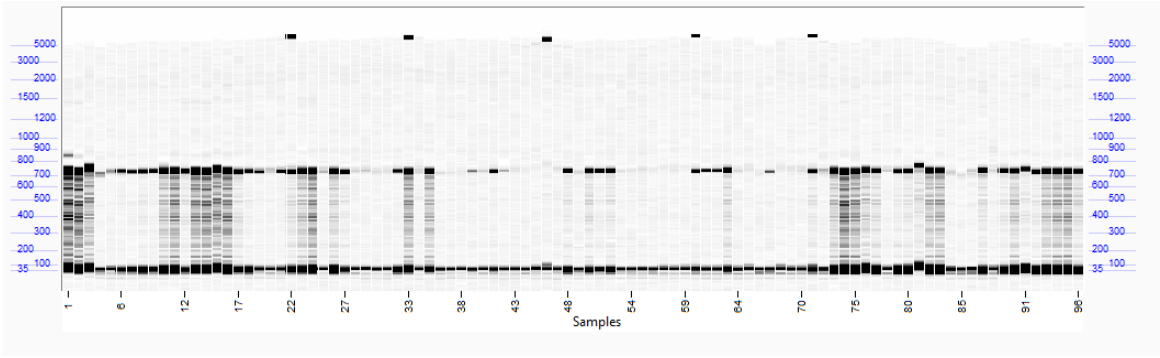
TILLING gel images

For each gel image lanes 1-12 = wells A1-A12, lanes 13-24=wells B1-B12 etc.

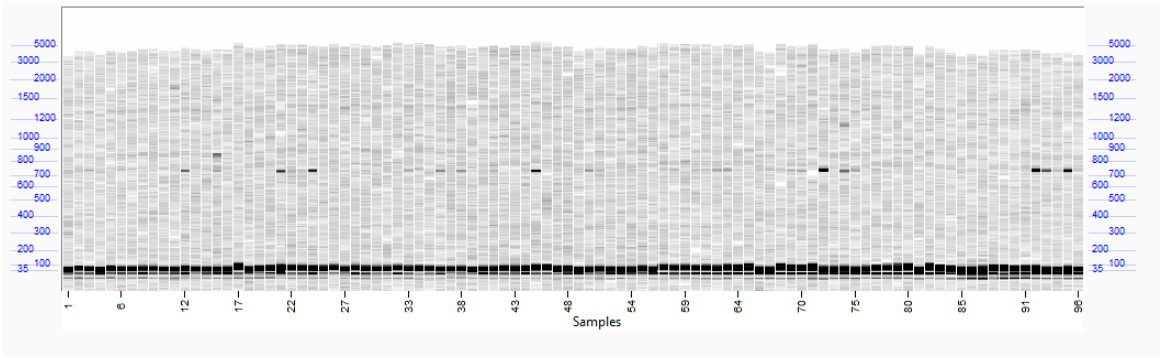
Pool plate 1



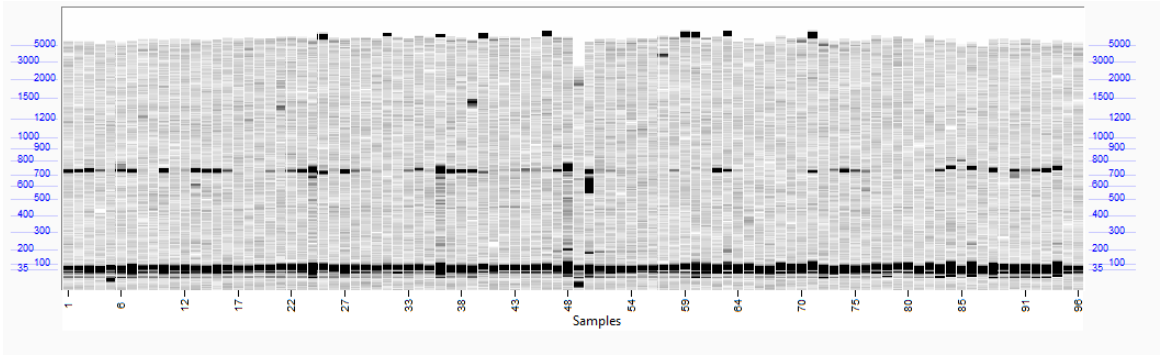
Pool plate 2



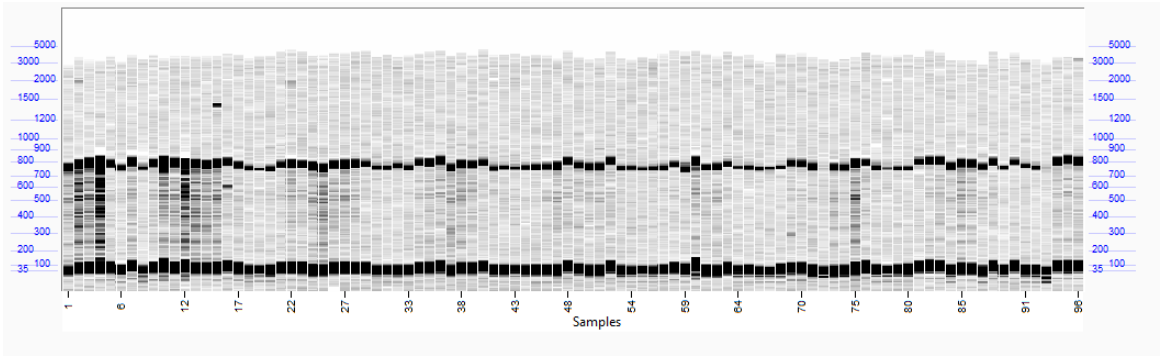
Pool plate 3



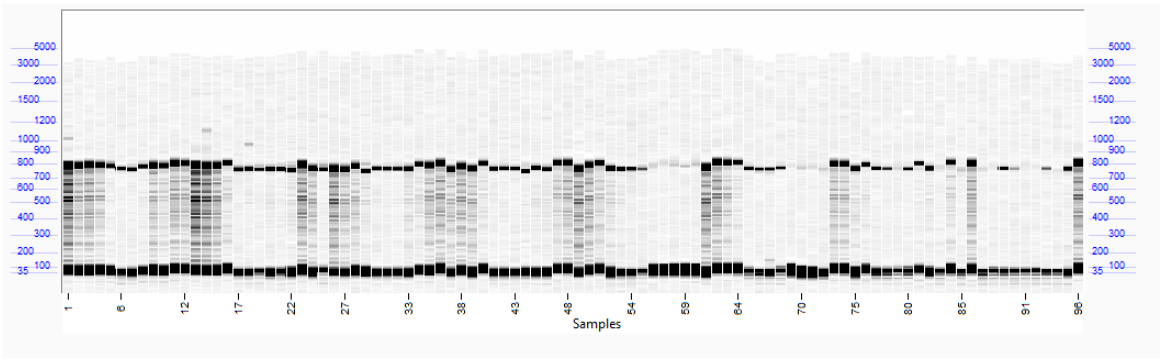
Pool plate 4



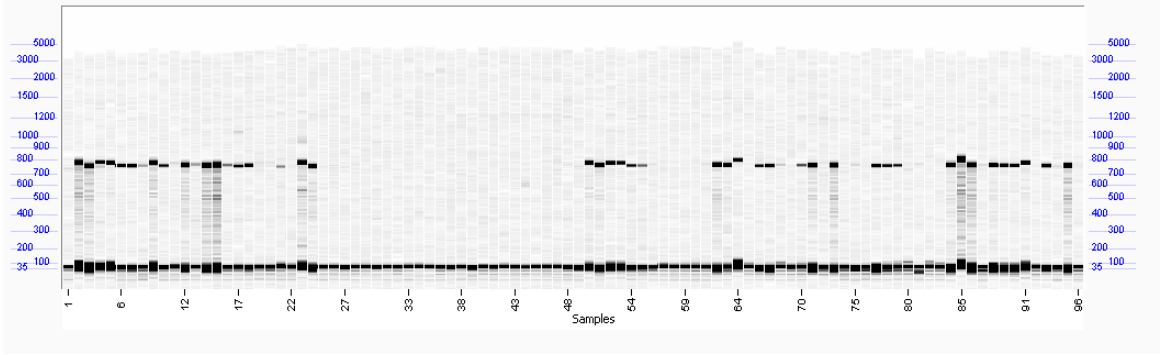
Pool plate 5



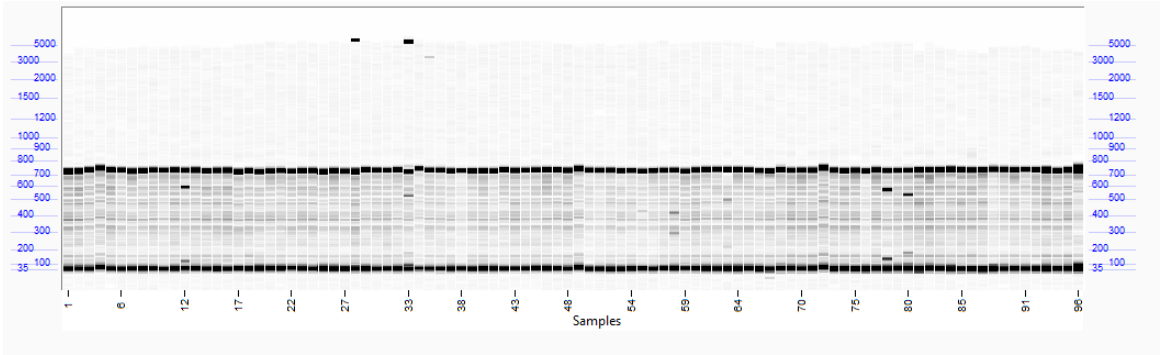
Pool plate 6



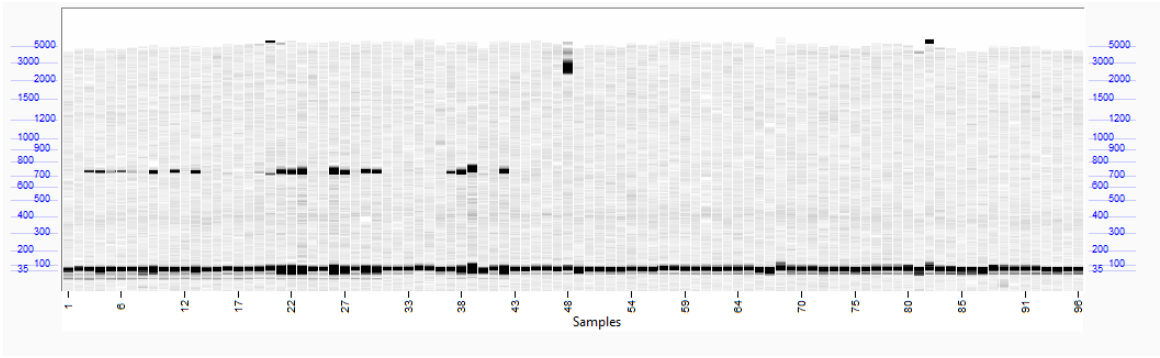
Pool plate 7



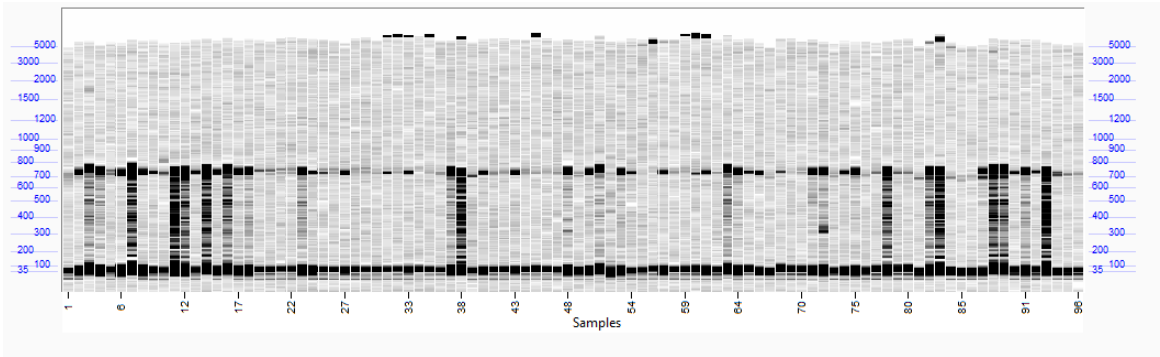
Pool plate 8



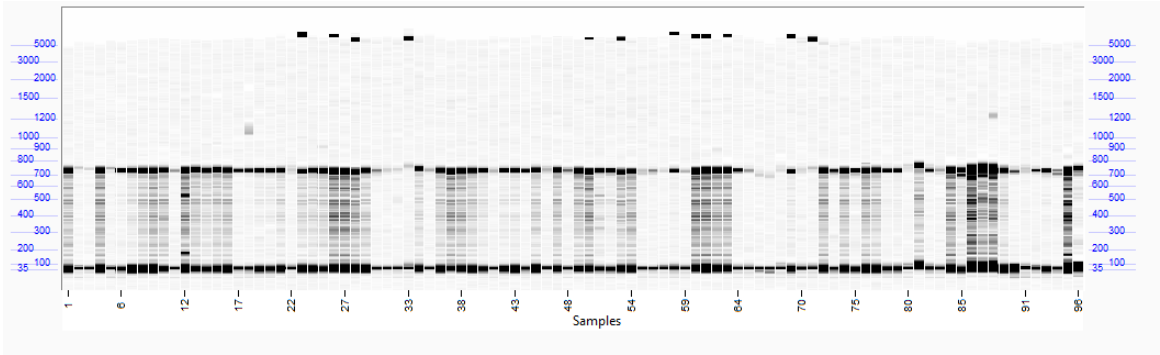
Pool plate 9



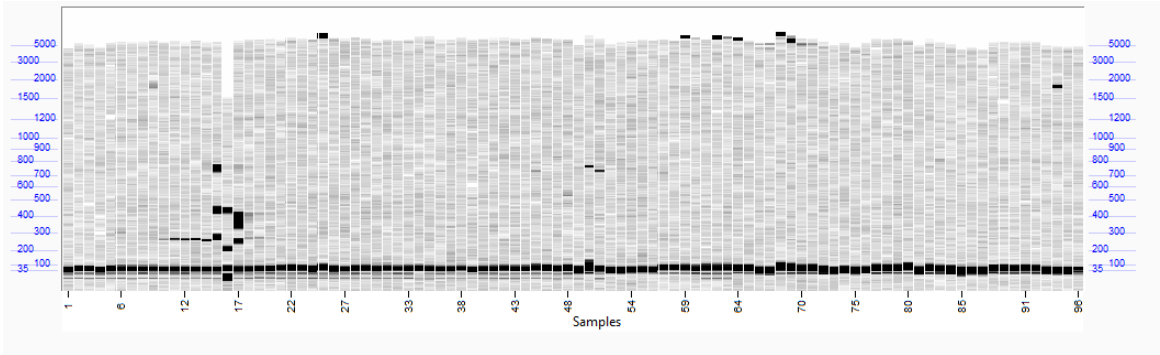
Pool plate 10



Pool plate 11

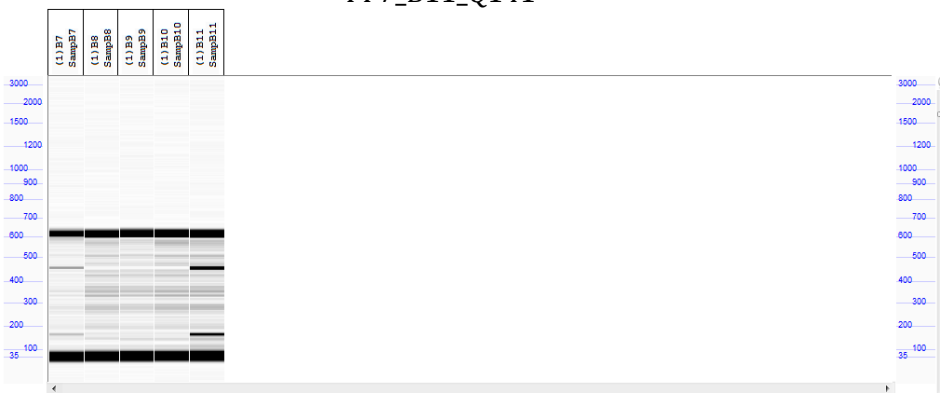


Pool plate 12



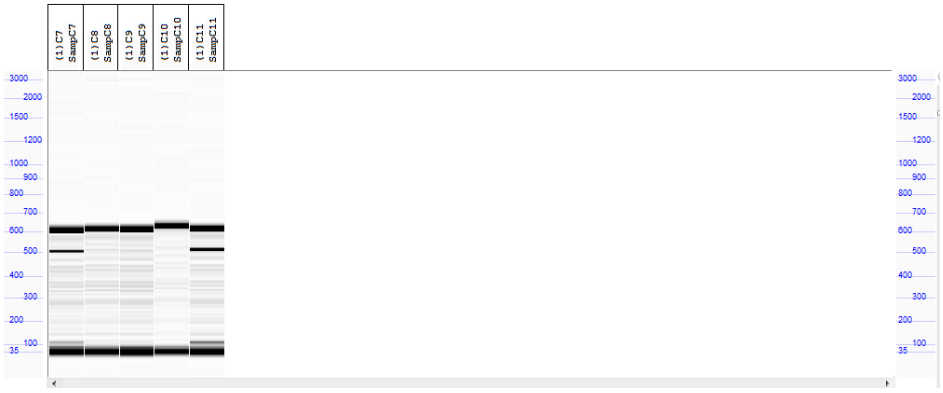
Rescreen of mutations that would later be identified as missense and nonsense mutations

PP7_B11_Q141*



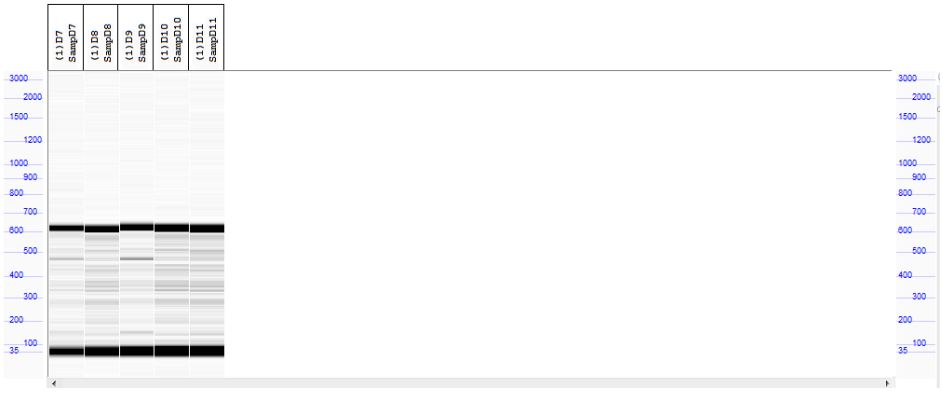
161 bp and 450 bp fragments in fourth individual (Sd-102693)

PP8_A12_H162Y



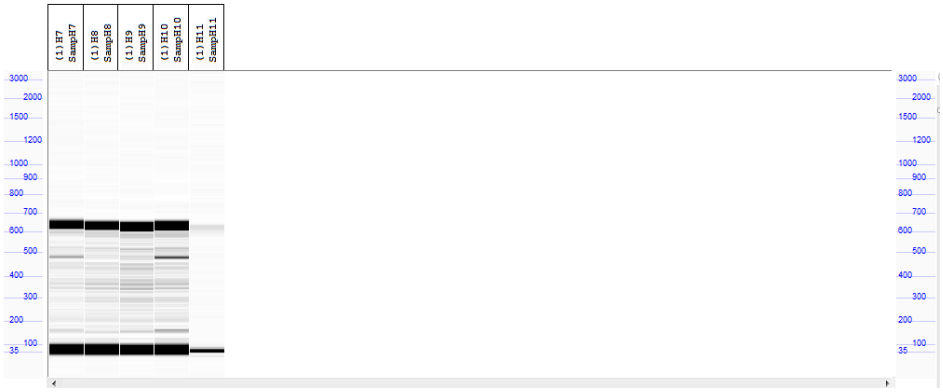
103 bp and 509 bp fragments in fourth individual (Sd-103090)

PP8_C9_A146T



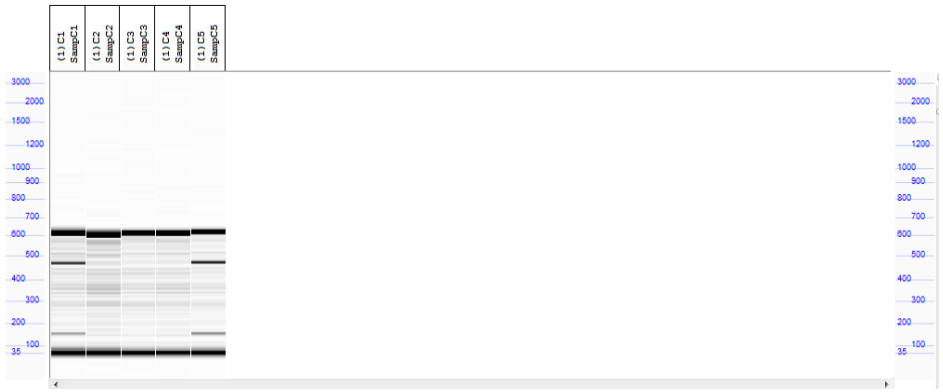
147 bp and 468 bp fragments in second individual (Sd-103161)

PP8_G8_A146T



156 bp and 473 bp fragments in third individual (Sd-103421)

PP11_A12_W145*



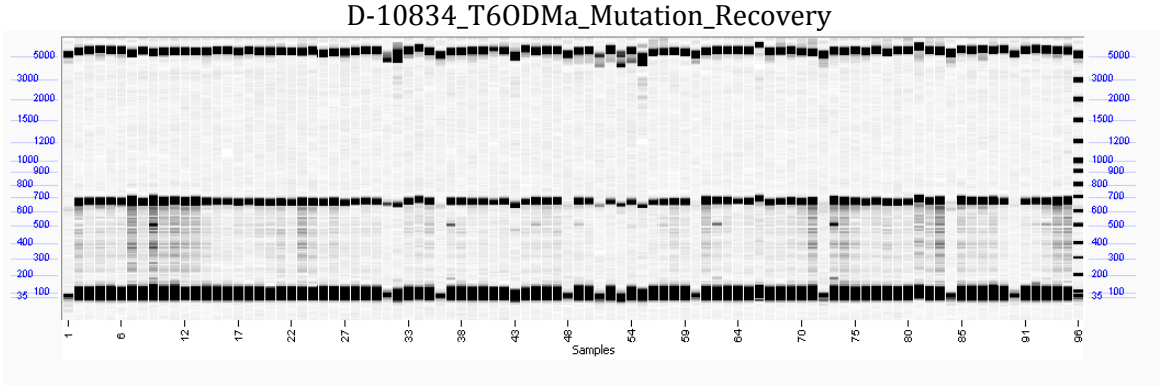
151 bp and 468 bp fragments in fourth individual (Sd-104553)

Location of T6ODM mutations in nucleotide and amino acid sequences

```
M   E   K   A   K   L   M   K   L   G   N   G   M   E   I   P   S   V     18
 atg gag aaa gca aaa ctt atg aag cta ggt aat ggt atg gaa ata cca agt gtt    54


Q   E   L   A   K   L   T   L   A   E   I   P   S   R   Y   V   C   A     36
caa gaa ttg gct aaa ctc acg ctt gcc gaa att cca tct cga tac gta tgc gcc    108


N   E   N   L   L   L   P   M   G   A   S   V   I   N   D   H   E   T     54
aat gaa aac ctt ttg ttg cct atg ggt gca tct gtc ata aat gat cat gaa acc    162


I   P   V   I   D   I   E   N   L   L   S   P   E   P   I   I   G   K     72
att cct gtc atc gat ata gaa aat tta tta tct cca gaa cca ata atc gga aag    216

                    IPB002283A (4.2e-13) IC 2.46
L   E   L   D   R   L   H   F   A   C   K   E   W   G   F   F   Q   | V    90
tta gaa tta gat agg ctt cat ttt gct tgc aaa gaa tgg ggt ttt ttt cag | gt   269


  V   N   H   G   V   D   A   S   L   V   D   S   V   K   S   E   I       107
a gtg aac cat gga gtc gac gct tca ttg gtg gat agt gta aaa tca gaa att      321

IPB002283B (4.2e-05) IC 1.94
Q   G   F   F   N   L   S   M   D   E   K   T   K   Y   E   Q   E   D     125
caa ggt ttc ttt aac ctt tct atg gat gag aaa act aaa tat gaa cag gaa gat    375


G   D   V   E   G   F   G   Q   G   F   I   E   S   E   D   Q   T   L     143
gga gat gtg gaa gga ttt gga caa ggc ttt att gaa tca gag gac caa aca ctt    429
                                                    tQ141*


D   W   A   D   I   F   M   M   F   T   L   P   L   H   L   R   K   P     161
gat tgg gca gat ata ttt atg atg ttc act ctt cca ctc cat tta agg aag cct    483
      aA146T
   aW145*
```

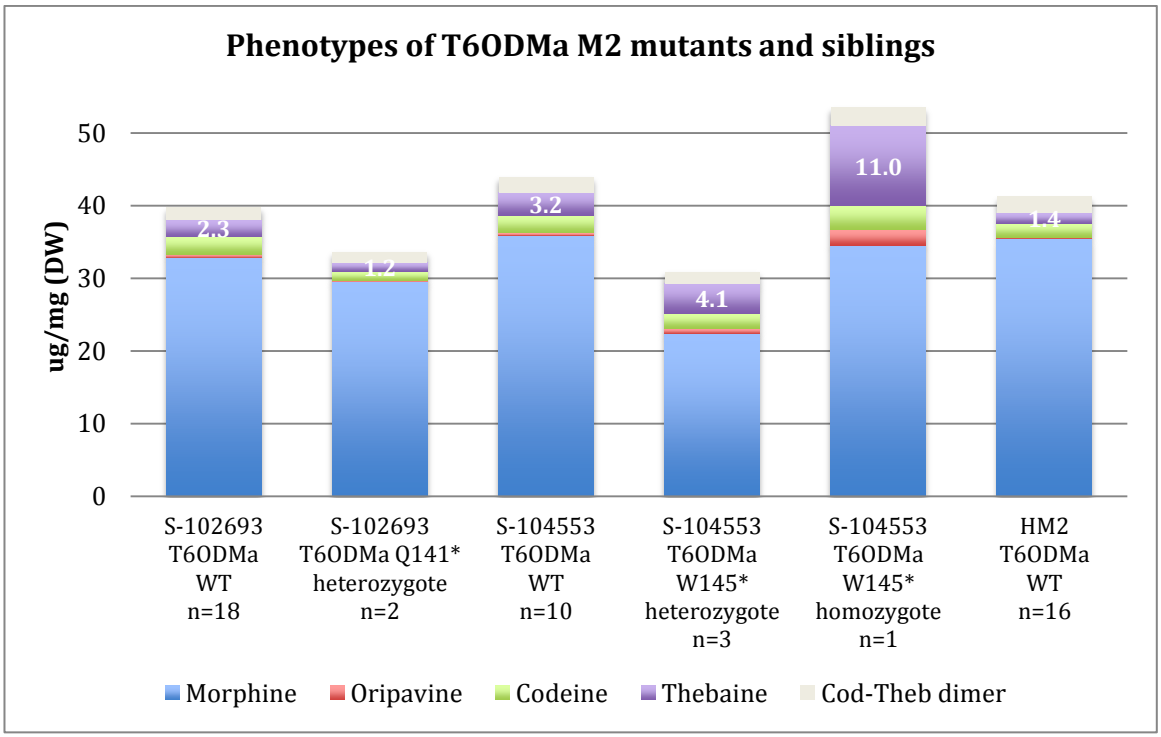
```
H   L   F   S   K   L   P   V   P   L   R   |   E   T   I   E   S   Y       178
cac tta ttt tca aaa ctc cca gtg cct ctc ag  | g gag aca atc gaa tcc tac     534
tH162Y


S   S   E   M   K   K   L   S   M   V   L   F   N   K   M   E   K   A       196
tca tca gaa atg aaa aag tta tcc atg gtt ctc ttt aat aag atg gaa aaa gct     588


L   Q   V   Q   A   A   E   I   K   G   M   S   E   V   F   I   D   G       214
cta caa gta caa gca gcc gag att aag ggt atg tca gag gtg ttt ata gat ggg     642


T   Q   A   M   R   M   N   Y   Y   P   P   C   P   Q   P   N   L   A       232
aca caa gca atg agg atg aac tat tat ccc cct tgt cct caa cca aat ctc gcc     696

                        IPB002283F (9.2e-06) IC 2.81
I   G   L   T   S   H   S   D   F   G   G   L   T   I   L   L   Q   I       250
atc ggt ctt acg tcg cac tcg gat ttt ggc ggt ttg aca atc ctc ctt caa atc     750


N   E   V   E   G   L   Q   I   K   R   E   G   T   W   I   S   V   K       268
aac gaa gtg gaa gga tta cag ata aaa aga gag ggg aca tgg att tca gtc aaa     804

                    IPB002283G (3.0e-15) IC 2.76
P   L   P   N   A   F   V   V   N   V   G   D   I   L   E   | I   M       285
cct cta cct aat gcg ttc gta gtg aat gtt gga gat att ttg gag | ata atg     855

                                                          IPB002283H (2.4e-09) IC 2.25
T   N   G   I   Y   H   S   V   D   H   R   A   V   V   N   S   T   N       303
act aat gga att tac cat agt gtc gat cac cgg gca gta gta aac tca aca aat     909


E   R   L   S   I   A   T   F   H   D   P   S   L   E   S   V   I   G       321
gag agg ctc tca atc gca aca ttt cat gac cct agt cta gag tcg gta ata ggc     963


P   I   S   S   L   I   T   P   E   T   P   A   L   F   K   S   G   S       339
cca ata tca agc ttg att act cca gag aca cct gct ttg ttt aaa agt gga tct    1017


T   Y   G   D   L   V   E   E   C   K   T   R   K   L   D   G   K   S       357
aca tat ggg gat ctt gtg gag gaa tgt aaa aca agg aag ctc gat gga aaa tca    1071


F   L   D   S   M   R   I   *                                               365
ttt ctt gac tcc atg agg att tga                                            1095
```

Recovery of T6ODMa mutations in M2 individuals

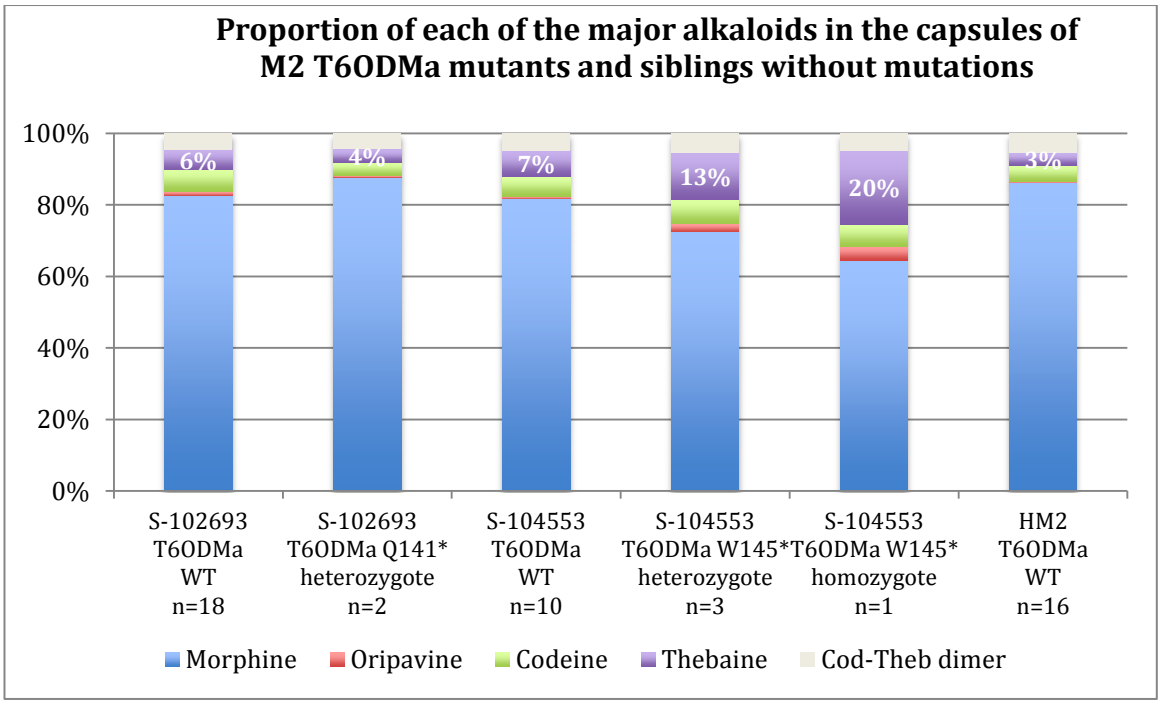


W145* mutations observed in wells D1, E1, F2 and G1 (lanes 37, 49, 62 and 73 respectively). Q141* mutation observed in well D7 (lane 43).

Alkaloid content of the capsules of T6ODMa M2 mutants and siblings



Absolute alkaloid content in the capsules of T6ODMa mutants and siblings with WT T6ODMa.



Relative alkaloid content in the capsules of T6ODMa mutants and siblings with WT T6ODMa.

Genealogies of high thebaine forward screen lines derived from HM2

Phenotypes expressed as percentage dry weight of morphine (M), codeine (C), oripavine (O) and thebaine (T).

MAS 2787 D

popn	season	seed ID		plot no	M	C	O	T
			OP row	**HM5**	**3.429**	**0.172**	**0.009**	**0.051**
B6M2	11_12	**S-109141**	**OP row**	**HT5**	**1.153**	**0.202**	**0.679**	**2.606**
			OP row	MA 2787	3.438	0.285	0.001	0.508
			SP cap	MAS 2787 4	3.561	1.053	0.356	3.243

MAS 2150 B C

popn	season	seed ID			plot no	M	C	O	T	
				OP row	**HM1**	**3.75**	**0.21**	**0.04**	**0.16**	
B4M2	09_10	**S-107166**		**OP row**	MUA 2642	1.167	0.295	0.576	2.136	
				SP caps	MUAS 2642 A	1.168	0.249	0.919	3.321	
B4M3	10_11			**OP row**	**HT4**	**0.926**	**0.187**	**0.520**	**2.029**	**0.926**
			REP 1/2	**OP row**	**HM4**	**2.889**	**0.139**	**0.022**	**0.098**	**2.735**
			ex MUAS 2642a	**OP row**	MUB 1013	0.832	0.231	0.305	2.194	0.683
			(B4M2 09-10)	**SP caps**	MUBS 1013 B	0.785	0.150	0.597	1.536	0.777
B4M4	11_12									
			REP 2/2	**OP row**	**HM5**	**3.587**	**0.107**	**0.023**	**0.081**	**3.799**
			ex MUBS 1013 B	**OP row**	**HT5**	**1.188**	**0.150**	**0.864**	**2.248**	**4.450**
			(B4M3 10-11)	**OP row**	MA 2150	0.970	0.127	0.799	2.042	3.938
				SP caps	**MAS 2150 B C**	0.845	0.353	1.056	3.182	5.437

MAS 1956-4

popn	season	seed ID			plot no	M	C	O	T	
				OP row	**HM1**	**3.65**	**0.15**	**0.03**	**0.14**	
B4M2	09_10	**S-106367**		**OP row**	MUA 1723	3.136	0.294	0.122	1.335	
				SP caps	MUAS 1723 A	1.344	0.240	0.145	3.576	
						M+pM	**C**	**O**	**T**	**M**
B4M3	10_11			**OP row**	**HT4**	**0.789**	**0.192**	**0.417**	**1.877**	**0.735**
			Single Row	**OP row**	**HM4**	**2.432**	**0.145**	**0.035**	**0.186**	**2.248**
			ex MUAS 1723a	**OP row**	MUB 1336	0.507	0.453	0.074	3.279	0.475
			(B4M2 09-10)	**SP caps**	MUBS 1336 D	0.424	0.616	0.078	4.374	0.391
						M	**C**	**O**	**T**	
B4M4	11_12									
			JIFFY (low seed)	**OP row**	**HM5**	**4.112**	**0.152**	**0.017**	**0.089**	
			ex MUBS 1336 D	**OP row**	**HT5**	**1.163**	**0.161**	**0.799**	**2.244**	
			(B4M3 10-11)	**OP row**	MA OP 1956- 4	0.537	0.228	0.002	6.218	
				SP caps	**MAS 1956-4**	0.873	0.214	0.091	4.229	

Appendix D- Design of primers for AS-PCR

CODMa/c

E193K_WASP

```
AAAGATTGATATATGATCTGAAGATCTGACAAGAAAGTTCATCAAATATAGAGTTCATGGAGACACCAATACTTATCAAGCTAGGCAATGGTTTGTCAATACCA
AGTGTTCAGGAATTGGCTAAACTCACGCTTGCAGAAATTCCATCTCGATACACATGCACCGGTGAAAGCCCGTTGAATAATATTGGTGCGTCTGTAACAGATGA
TGAAACAGTTCCTGTCATCGATTTGCAAAATTTACTATCTCCAGAACCCGTAGTTGGAAAGTTAGAATTGGATAAGCTTCATTCTGCTTGCAAAGAATGGGGTT
TCTTTCAGGTATACAAATTTTGATTATACGTAGTTGATTGACGACTATTTAGTGTTTTACTAGTTCATGCTAATCTGCTTTATGTTTCAGCTGGTTAACCATGG
AGTCGACGCTTTACTGATGGACAATATAAAATCAGAAATTAAAGGTTTCTTTAACCTTCCAATGAATGAGAAAACTAAATACGGACAGCAAGATGGAGATTTTG
AAGGATTTGGACAACCCTATATTGAATCGGAGGACCAAAGACTTGATTGGACTGAAGTGTTTAGCATGTTAAGTCTTCCTCTCCATTTAAGGAAGCCTCATTTG
TTTCCAGAACTCCCTCTGCCTTTCAGGTTATTGTTTCTCTCCGTTTGCTTTTTCTCAATGCCGCACGTGATATTTTTTTTTTTTTAATGTTATTTTGGTTGGTTC
AGAATAGTTTCATCAATATCGTCTAGGTTGCTAGGATAAACCATGGAAAATCCTGTCTAGGTTGCCTAAAATTCACATTCGGAGGCAGTATCAAGCAATCCGGG
TTTGTGAATTGCATGGTTTCCAAAATATATCGAATCCAAACATCATGCATCCCTAAAATCTGTCAGATGGAAAATATACTGAAGTTGCATTTGATCCCTAATTT
ACATCCTTTATGCAGGGAGACACTGGAATCCTACCTATCAAAAATGAAAAAACTATCAACGGTTGTCTTTGAGATGTTG[G/A]AAAAATCTCTACAATTAGTT
GAGATTAAAGGTATGACAGACTTATTTGAAGATGGGTTGCAAACAATGAGGATGAACTATTATCCTCCTTGTCCTCGACCAGAGCTTGTACTTGGTCTTACGTC
ACACTCGGATTTTAGCGGTTTGACAATTCTCCTTCAACTTAATGAAGTTGAAGGATTACAAATAAGAAAAGAAGAGAGGTGGATTTCAATCAAACCTCTACCTG
ATGCGTTCATAGTGAATGTTGGAGACATTTTGGAGGTATTTAGGCTAAGGGCATATATGGCCTGCCTACTGGTCCTTGTTGTGAGTACCTGTTTTGACTGTACC
CCGAAAAGTTGCTTAGTTACTTAATGTTTCCCTTTTACAGATAATGACTAATGGGATTTACCGTAGCGTCGAGCACCGGGCAGTAGTAAACTCAACAAAGGAGA
GGCTCTCAATCGCGACATTTCATGACTCTAAACTAGAGTCAGAAATAGGCCCAATTTCGAGCTTGGTCACACCAGAGACACCTGCTTTGTTCAAAAGAGGTAGG
TATGAGGATATTTTGAAGGAAAATCTTTCAAGGAAGCTTGATGGAAAATCATTTCTCGACTACATGAGGATGTGAGAAAGTGTGAACATACATTATACTCCACA
TTGTGTTTAATATATGATGAAATAAGTTGCTTTTGAAGTATGATGAAATAAGTTGGTTTTGAAGAATTCATATTGTGCTTAAATTTCGTGGATGACTGAGAGAT
TTATTATGTAATAATAATGTATTGGTTTGAAGATTCTCGTCTCACTATATGTAAGACTCTGTTTGGG
```

```
Oligo 1 :                                                    Pos.    Len.     Tm
GC%   Self Any   Self End
Wildtype Forward Primer 5': TCAACGGTTGTCTTTGAGATGTTTG        992     25     64.49
40.00     7.00     0.00
Mutant Forward Primer 5': TCAACGGTTGTCTTTGAGATGTTTA          992     25     62.03
36.00     7.00     2.00
Common Reverse Primer 5': AAGTACAAGCTCTGGTCGAG               1129     20     54.76
50.00     4.00     3.00
Product Size: 138
```

*W261*_WASP*

```
AAAGATTGATATATGATCTGAAGATCTGACAAGAAAGTTCATCAAATATAGAGTTCATGGAGACACCAATACTTATCAAGCTAGGCAATGGTTTGTCAAT
ACCAAGTGTTCAGGAATTGGCTAAACTCACGCTTGCAGAAATTCCATCTCGATACACATGCACCGGTGAAAGCCCGTTGAATAATATTGGTGCGTCTGTA
ACAGATGATGAAACAGTTCCTGTCATCGATTTGCAAAATTTACTATCTCCAGAACCCGTAGTTGGAAAGTTAGAATTGGATAAGCTTCATTCTGCTTGCA
AAGAATGGGGTTTCTTTCAGGTATACAAATTTTGATTATACGTAGTTGATTGACGACTATTTAGTGTTTTACTAGTTCATGCTAATCTGCTTTATGTTTC
AGCTGGTTAACCATGGAGTCGACGCTTTACTGATGGACAATATAAAATCAGAAATTAAAGGTTTCTTTAACCTTCCAATGAATGAGAAAACTAAATACGG
ACAGCAAGATGGAGATTTTGAAGGATTTGGACAACCCTATATTGAATCGGAGGACCAAAGACTTGATTGGACTGAAGTGTTTAGCATGTTAAGTCTTCCT
CTCCATTTAAGGAAGCCTCATTTGTTTCCAGAACTCCCTCTGCCTTTCAGGTTATTGTTTCTCTCCGTTTGCTTTTTCTCAATGCCGCACGTGATATTTT
TTTTTTTTAATGTTATTTTGGTTGGTTCAGAATAGTTTCATCAATATCGTCTAGGTTGCTAGGATAAACCATGGAAAATCCTGTCTAGGTTGCCTAAAAT
TCACATTCGGAGGCAGTATCAAGCAATCCGGGTTTGTGAATTGCATGGTTTCCAAAATATATCGAATCCAAACATCATGCATCCCTAAAATCTGTCAGAT
GGAAAATATACTGAAGTTGCATTTGATCCCTAATTTACATCCTTTATGCAGGGAGACACTGGAATCCTACCTATCAAAAATGAAAAAACTATCAACGGTT
GTCTTTGAGATGTTGGAAAAATCTCTACAATTAGTTGAGATTAAAGGTATGACAGACTTATTTGAAGATGGGTTGCAAACAATGAGGATGAACTATTATC
CTCCTTGTCCTCGACCAGAGCTTGTACTTGGTCTTACGTCACACTCGGATTTTAGCGGTTTGACAATTCTCCTTCAACTTAATGAAGTTGAAGGATTACA
AATAAGAAAAGAAGAGAGGT[G/A]GATTTCAATCAAACCTCTACCTGATGCGTTCATAGTGAATGTTGGAGACATTTTGGAGGTATTTAGGCTAAGGGC
ATATATGGCCTGCCTACTGGTCCTTGTTGTGAGTACCTGTTTTGACTGTACCCCGAAAAGTTGCTTAGTTACTTAATGTTTCCCTTTTACAGATAATGAC
TAATGGGATTTACCGTAGCGTCGAGCACCGGGCAGTAGTAAACTCAACAAAGGAGAGGCTCTCAATCGCGACATTTCATGACTCTAAACTAGAGTCAGAA
ATAGGCCCAATTTCGAGCTTGGTCACACCAGAGACACCTGCTTTGTTCAAAAGAGGTAGGTATGAGGATATTTTGAAGGAAAATCTTTCAAGGAAGCTTG
ATGGAAAATCATTTCTCGACTACATGAGGATGTGAGAAAGTGTGAACATACATTATACTCCACATTGTGTTTAATATATGATGAAATAAGTTGCTTTTGA
AGTATGATGAAATAAGTTGGTTTTGAAGAATTCATATTGTGCTTAAATTTCGTGGATGACTGAGAGATTTATTATGTAATAATAATGTATTGGTTTGAAG
ATTCTCGTCTCACTATATGTAAGACTCTGTTTGGG
```

```
Oligo 3 :                                                    Pos.    Len.     Tm       GC%    Self Any
Self End
Wildtype Reverse Primer 5': AGGTAGAGGTTTGATTGAAATAC          1243     23     53.54     34.78     5.00
2.00
Mutant Reverse Primer 5': AGGTAGAGGTTTGATTGAAATAT            1243     23     53.14     30.44     5.00
4.00
Common Forward Primer 5': TTGCATTTGATCCCTAATTT               917     20     54.84     30.00     4.00
3.00
Product Size: 327
```

E301K_WASP

```
AAAGATTGATATATGATCTGAAGATCTGACAAGAAAGTTCATCAAATATAGAGTTCATGGAGACACCAATACTTATCAAGCTAGGCAATGGTTTGTCAAT
ACCAAGTGTTCAGGAATTGGCTAAACTCACGCTTGCAGAAATTCCATCTCGATACACATGCACCGGTGAAAGCCCGTTGAATAATATTGGTGCGTCTGTA
ACAGATGATGAAACAGTTCCTGTCATCGATTTGCAAAATTTACTATCTCCAGAACCCGTAGTTGGAAAGTTAGAATTGGATAAGCTTCATTCTGCTTGCA
AAGAATGGGGTTTCTTTCAGGTATACAAATTTTGATTATACGTAGTTGATTGACGACTATTTAGTGTTTTACTAGTTCATGCTAATCTGCTTTATGTTTC
```

AGCTGGTTAACCATGGAGTCGACGCTTTACTGATGGACAATATAAAATCAGAAATTAAAGGTTTCTTTAACCTTCCAATGAATGAGAAAACTAAATACGG
ACAGCAAGATGGAGATTTTGAAGGATTTGGACAACCCTATATTGAATCGGAGGACCAAAGACTTGATTGGACTGAAGTGTTTAGCATGTTAAGTCTTCCT
CTCCATTTAAGGAAGCCTCATTTGTTTCCAGAACTCCCTCTGCCTTTCAGGTTATTGTTTCTCTCCGTTTGCTTTTTCTCAATGCCGCACGTGATATTTT
TTTTTTTTAATGTTATTTTGGTTGGTTCAGAATAGTTTCATCAATATCGTCTAGGTTGCTAGGATAAACCATGGAAAATCCTGTCTAGGTTGCCTAAAAT
TCACATTCGGAGGCAGTATCAAGCAATCCGGGTTTGTGAATTGCATGGTTTCCAAAATATATCGAATCCAAACATCATGCATCCCTAAAATCTGTCAGAT
GGAAAATATACTGAAGTTGCATTTGATCCCTAATTTACATCCTTTATGCAGGGAGACACTGGAATCCTACCTATCAAAAATGAAAAAACTATCAACGGTT
GTCTTTGAGATGTTGGAAAAATCTCTACAATTAGTTGAGATTAAAGGTATGACAGACTTATTTGAAGATGGGTTGCAAACAATGAGGATGAACTATTATC
CTCCTTGTCCTCGACCAGAGCTTGTACTTGGTCTTACGTCACACTCGGATTTTAGCGGTTTGACAATTCTCCTTCAACTTAATGAAGTTGAAGGATTACA
AATAAGAAAAGAAGAGAGGTGGATTTCAATCAAACCTCTACCTGATGCGTTCATAGTGAATGTTGGAGACATTTTGGAGGTATTTAGGCTAAGGGCATAT
ATGGCCTGCCTACTGGTCCTTGTTGTGAGTACCTGTTTTGACTGTACCCCGAAAAGTTGCTTAGTTACTTAATGTTTCCCTTTTACAGATAATGACTAAT
GGGATTTACCGTAGCGTCGAGCACCGGGCAGTAGTAAACTCAACAAAG[G/A]AGAGGCTCTCAATCGCGACATTTCATGACTCTAAACTAGAGTCAGAA
ATAGGCCCAATTTCGAGCTTGGTCACACCAGAGACACCTGCTTTGTTCAAAAGAGGTAGGTATGAGGATATTTTGAAGGAAAATCTTTCAAGGAAGCTTG
ATGGAAAATCATTTCTCGACTACATGAGGATGTGAGAAAGTGTGAACATACATTATACTCCACATTGTGTTTAATATATGATGAAATAAGTTGCTTTTGA
AGTATGATGAAATAAGTTGGTTTTGAAGAATTCATATTGTGCTTAAATTTCGTGGATGACTGAGAGATTTATTATGTAATAATAATGTATTGGTTTGAAG
ATTCTCGTCTCACTATATGTAAGACTCTGTTTGGG

Oligo 4 :	Pos.	Len.	Tm	GC%	Self Any Self End
Wildtype Forward Primer 5': GCAGTAGTAAACTCAACAAATG	1428	22	52.08	36.36	4.00 3.00
Mutant Forward Primer 5': GCAGTAGTAAACTCAACAAATA	1428	22	49.46	31.82	4.00 2.00
Common Reverse Primer 5': CTCTCAGTCATCCACGAAAT	1762	20	55.19	45.00	3.00 2.00

Product Size: 335

CODMb

R158K_WASP

AAAGATTGATATATGATCTGAAGATCTGACAAGAAAGTTCATCAAATATAGAGTTCATGGAGACACCAATACTTATCAAGCTAGGCAATGGTTTGTCAAT
ACCAAGTGTTCAGGAATTGGCTAAACTCACGCTTGCAGAAATTCCATCTCGATACACATGCACCGGTGAAAGCCCGTTGAATAATATTGGTGCGTCTGTA
ACAGATGATGAAACAGTTCCTGTCATCGATTTGCAAAATTTACTATCTCCAGAACCCGTAGTTGGAAAGTTAGAATTGGATAAGCTTCATTCTGCTTGCA
AAGAATGGGGTTTCTTTCAGGTATACAAATTTTGATTATACGTAGTTGATTGACGACTATTTAGTGTTTTACTAGTTCATGCTAATCTGCTTTATGTTTC
AGCTGGTTAACCATGGAGTCGACGCTTTACTGATGGACAATATAAAATCAGAAATTAAAGGTTTCTTTAACCTTCCAATGAATGAGAAAACTAAATACGG
ACAGCAAGATGGAGATTTTGAAGGATTTGGACAACCCTATATTGAATCGGAGGACCAAAGACTTGATTGGACTGAAGTGTTTAGCATGTTAAGTCTTCCT
CTCCATTTAA[G/A]GAAGCCTCATTTGTTTCCAGAACTCCCTCTGCCTTTCAGGTTATTGTTTCTCTCCGTTTGCTTTTTCTCAATGCCGCACGTGATA
TTTTTTTTTTTTAATGTTATTTTGGTTGGTTCAGAATAGTTTCATCAATATCGTCTAGGTTGCTAGGATAAACCATGGAAAATCCTGTCTAGGTTGCCTA
AAATTCACATTCGGAGGCAGTATCAAGCAATCCGGGTTTGTGAATTGCATGGTTTCCAAAATATATCGAATCCAAACATCATGCATCCCTAAAATCTGTC
AGATGGAAAATATACTGAAGTTGCATTTGATCCCTAATTTACATCCTTTATGCAGGGAGACACTGGAATCCTACCTATCAAAAATGAAAAAACTATCAAC
GGTTGTCTTTGAGATGTTGGAAAAAATCTCTACAATTAGTTGAGATTAAAGGTATGACAGACTTATTTGAAGATGGGTTGCAAACAATGAGGATGAACTA
TTATCCTCCTTGTCCTCGACCAGAGCTTGTACTTGGTCTTACGTCACACTCGGATTTTAGCGGTTTGACAATTCTCCTTCAACTTAATGAAGTTGAAGGA
TTACAAATAAGAAAAGAAGAGAGGTGGATTTCAATCAAACCTCTACCTGATGCGTTCATAGTGAATGTTGGAGACATTTTGGAGGTATTTAGGCTAAGGG
CATATATGGCCTGCCTACTGGTCCTTGTTGTGAGTACCTGTTTTGACTGTACCCCGAAAAGTTGCTTAGTTACTTAATGTTTCCCTTTTACAGATAATGA
CTAATGGGATTTACCGTAGCGTCGAGCACCGGGCAGTAGTAAACTCAACAAAGGAGAGGCTCTCAATCGCGACATTTCATGACTCTAAACTAGAGTCAGA
AATAGGCCCAATTTCGAGCTTGGTCACACCAGAGACACCTGCTTTGTTCAAAAGAGGTAGGTATGAGGATATTTTGAAGGAAAATCTTTCAAGGAAGCTT
GATGGAAAATCATTTCTCGACTACATGAGGATGTGAGAAAGTGTGAACATACATTATACTCCACATTGTGTTTAATATATGATGAAATAAGTTGCTTTTG
AAGTATGATGAAATAAGTTGGTTTTGAAGAATTCATATTGTGCTTAAATTTCGTGGATGACTGAGAGATTTATTATGTAATAATAATGTATTGGTTTGAA
GATTCTCGTCTCACTATATGTAAGACTCTGTTTGGG

Oligo 3 :	Pos.	Len.	Tm	GC%	Self Any Self End
Wildtype Forward Primer 5': TTAAGTCTTCCTCTCCATTTACG	589	23	56.83	39.13	5.00 2.00
Mutant Forward Primer 5': TTAAGTCTTCCTCTCCATTTACA	589	23	55.02	34.78	5.00 2.00
Common Reverse Primer 5': TTAGGGATCAAATGCAACTT	933	20	54.90	35.00	4.00 3.00

Product Size: 345

*Q254*_WebSNAPER*

AAAGATTGATATATGATCTGAAGATCTGACAAGAAAGTTCATCAAATATAGAGTTCATGGAGACACCAATACTTATCAAGCTAGGCAATGGTTTGTCAATACCA
AGTGTTCAGGAATTGGCTAAACTCACGCTTGCAGAAATTCCATCTCGATACACATGCACCGGTGAAAGCCCGTTGAATAATATTGGTGCGTCTGTAACAGATGA
TGAAACAGTTCCTGTCATCGATTTGCAAAATTTACTATCTCCAGAACCCGTAGTTGGAAAGTTAGAATTGGATAAGCTTCATTCTGCTTGCAAAGAATGGGGTT
TCTTTCAGGTATACAAATTTTGATTATACGTAGTTGATTGACGACTATTTAGTGTTTTACTAGTTCATGCTAATCTGCTTTATGTTTCAGCTGGTTAACCATGG
AGTCGACGCTTTACTGATGGACAATATAAAATCAGAAATTAAAGGTTTCTTTAACCTTCCAATGAATGAGAAAACTAAATACGGACAGCAAGATGGAGATTTTG
AAGGATTTGGACAACCCTATATTGAATCGGAGGACCAAAGACTTGATTGGACTGAAGTGTTTAGCATGTTAAGTCTTCCTCTCCATTTAAGGAAGCCTCATTTG
TTTCCAGAACTCCCTCTGCCTTTCAGGTTATTGTTTCTCTCCGTTTGCTTTTTCTCAATGCCGCACGTGATATTTTTTTTTTTTTAATGTTATTTTGGTTGGTTC
AGAATAGTTTCATCAATATCGTCTAGGTTGCTAGGATAAACCATGGAAATCCTGTCTAGGTTGCCTAAAATTCACATTCGGAGGCAGTATCAAGCAATCCGGGT
TTGTGAATTGCATGGTTTCCAAAATATATCGAATCCAAACATCATGCATCCCTAAAATCTGTCAGATGGAAAATATACTGAAGTTGCATTTGATCCCTAATTTA
CATCCTTTATGCAGGGAGACACTGGAATCCTACCTATCAAAAATGAAAAAACTATCAACGGTTGTCTTTGAGATGTTGGAAAAATCTCTACAATTAGTTGAGAT
TAAAGGTATGACAGACTTATTTGAAGATGGGTTGCAAACAATGAGGATGAACTATTATCCTCCTTGTCCTCGACCAGAGCTTGTACTTGGTCTTACGTCACACT
CGGATTTTAGCGGTTTGACAATTCTCCTTCAACTTAATGAAGTTGAAGGATTA[C/T]AAATAAGAAAAGAAGAGAGGTGGATTTCAATCAAACCTCTACCTGA
TGCGTTCATAGTGAATGTTGGAGACATTTTGGAGGTATTTAGGCTAAGGGCATATATGGCCTGCCTACTGGTCCTTGTTGTGAGTACCTGTTTTGACTGTACCC
CGAAAAGTTGCTTAGTTACTTAATGTTTCCCTTTTACAGATAATGACTAATGGGATTTACCGTAGCGTCGAGCACCGGGCAGTAGTAAACTCAACAAAGGAGAG
GCTCTCAATCGCGACATTTCATGACTCTAAACTAGAGTCAGAAATAGGCCCAATTTCGAGCTTGGTCACACCAGAGACACCTGCTTTGTTCAAAAGAGGTAGGT
ATGAGGATATTTTGAAGGAAAATCTTTCAAGGAAGCTTGATGGAAAATCATTTCTCGACTACATGAGGATGTGAGAAAGTGTGAACATACATTATACTCCACAT
TGTGTTTAATATATGATGAAATAAGTTGCTTTTGAAGTATGATGAAATAAGTTGGTTTTGAAGAATTCATATTGTGCTTAAATTTCGTGGATGACTGAGAGATT
TATTATGTAATAATAATGTATTGGTTTGAAGATTCTCGTCTCACTATATGTAAGACTCTGTTTGGG

Primer Id	Primer Sequence	Tm	Primer Length	Product Size	Warnings
Q254*_R_REF_1	GAAATCCACCTCTCTTCTTTTCTTATTCG	61.925	29		
Q254*_R_ALT_1	GATTGAAATCCACCTCTCTTCTTTTCTTATTAA	62.420	33		
Q254*_R_ALT_1_REVERSE	TGTCAGATGGAAAATATACTGAAGTTGCA	62.003	29	339	

T6ODMa

Intron 1 polymorphism in GSK thebaine lines (Tint)_WASP

Intron 1 sequence in lower case

GTTCTTAATTCATTAATTAATTTAGAAAAATCATGGAGAAAGCAAAACTTATGAAGCTAGGTAATGGTATGGAAATACCAAGTGTTCAAGAATTGGCTAA
ACTCACGCTTGCCGAAATTCCATCTCGATACGTATGCGCCAATGAAAACCTTTTGTTGCCTATGGGTGCATCTGTCATAAATGATCATGAAACCATTCCT
GTCATCGATATAGAAAATTTATTATCTCCAGAACCAATAATCGGAAAGTTAGAATTAGATAGGCTTCATTTTGCTTGCAAAGAATGGGGTTTTTTTCAGg
tatatatgtcgtaaatacaagcattgtgattgcatccagtagttgtactttgttacactaaatattagc[t/a]agtgaatcttactttacatatatatg
ttgcagGTAGTGAACCATGGAGTCGACGCTTCATTGGTGGATAGTGTAAAATCAGAAATTCAAGGTTTCTTTAACCTTTCTATGGATGAGAAAACTAAAT
ATGAACAGGAAGATGGAGATGTGGAAGGATTTGGACAAGGCTTTATTGAATCAGAGGACCAAACACTTGATTGGGCAGATATATTTATGATGTTCACTCT
TCCACTCCATTTAAGGAAGCCTCACTTATTTTCAAAACTCCCAGTGCCTCTCAGGTTATTATTAATTCATAATCGGTTTCTATGCACTTGACTTTTTAAT
CTCACATTTTCGTAGTCGGTTTCTGTGCACCTCACTCAGTTTATTCTTATTCTTTTTTTTGTTTGGCCCACACATATTTAACATTTAGTCCACCTTGTTA
GAATTTCAGACTATAAATCATTCTTTACAATTCCTTATTAGCTGTAACCATGTTGTATGTGAACATATTTAATATTTTATGCTGATATATAATTATGTAT
ATGCATGCTCTAGAATTATATATTCTACCACTTGTAGGAAGATACTTCCAATTTTATCTTCTCTTATTCTTTAATGATGGAAAATTCCTTTATAAACAAG
AGCTAGAACGGAGTACACCTCACCTCATGCTAGGAAAGACATATGTCATTCAACCTTCCATCAAGAATTGAAAACAACCTCACATAGGATTATCATAACT
AGTATGTCATTCCAAATAGTGTAAGAGATATTTGTCCTTATGTACCCTTAAGCATACATAATTAACACAAACAAACCTTAATTCATTGATTAGGATTAGT
ACTAATCAAGTTTAAAGGAATACAAATTCTGTTTTACTGCTGAACCAACATAAAAAACTGACACACCTAAGTCTCCCTATTTCTATAACAAAAATTATCC
CTCTTGATATTATTTTCTGGTTCCCCCATTGACCAATAAACATGCAAAATTAACACTCTTTAAAAGAGTACGTATGATAGATCATAATATTTTTCTCTCT
TTGAATTTTCTCAGTGATGTGATTGGATACATTAATTTGGTTGGTTTATTATAGTTTTTCATATGATTATCTAATTTACGTACAGGGAGACAATCGAATC
CTACTCATCAGAAATGAAAAAGTTATCCATGGTTCTCTTTAATAAGATGGAAAAAGCTCTACAAGTACAAGCAGCCGAGATTAAGGGTATGTCAGAGGTG
TTTATAGATGGGACACAAGCAATGAGGATGAACTATTATCCCCCTTGTCCTCAACCAAATCTCGCCATCGGTCTTACGTCGCACTCGGATTTTGGCGGTT
TGACAATCCTCCTTCAAATCAACGAAGTGGAAGGATTACAGATAAAAAGAGAGGGGACATGGATTGCAGTCGAACCTCTACCTAATGCGTTCGTAGTGAA
TGTTGGAGATATTTTGGAGGTACTGGTCCTTGTTTCGAATACTGGTTTTGCCTGTACCCAGAAAATTTCTTATTGTTTTTTCTTTTACAGATAATGACTA
ATGGAATTTACCATAGTGTCGATCACCGGGCAGTAGTAAACTCAACAAATGAGAGGCTCTCAATCGCAATATTTCATGACCCTAGTCTAGAGTCGGTAAT
AGGCCCAATATCAAGCTTGATTACTCCAGAGACACCTGCTTTGTTTAAAAGTGGATCTACATATGGGGATCTTGTGGAGGAATGTAAAACAAGGAAGCTC
GATGGAAAATCATTTCTTGACTCCATGAGGATTTGAAAACTCAAGAAAAAATAATACGACGTGATTGCATGTCAGATTCAACTATCCTTTTGTTGTTTTT
TGGTGCTCGAGTCCTTAATTGTTTTGATCATTGCTTTTGATTCTAATTAATAAGACTTTTCTCAAGAACCACATGTAATGTACCTTTACTTTCAGAAAAT
AAAAAGTATTGAGGCACAAATGAGAAAATTGA

Oligo 3 :	Pos.	Len.	Tm	GC%	Self Any	Self End
Wildtype Reverse Primer 5': TATATGTAAAGTAAGATTCACCA	392	23	49.01	26.09	4.00	0.00
Mutant Reverse Primer 5': TATATGTAAAGTAAGATTCACCT	392	23	47.51	26.09	4.00	2.00
Common Forward Primer 5': CAGAACCAATAATCGGAAAG	229	20	54.91	40.00	5.00	0.00

Product Size: 164

*Q141*_WASP*

GTTCTTAATTCATTAATTAATTTAGAAAAATCATGGAGAAAGCAAAACTTATGAAGCTAGGTAATGGTATGGAAATACCAAGTGTTCAAGAATTGGCTAA
ACTCACGCTTGCCGAAATTCCATCTCGATACGTATGCGCCAATGAAAACCTTTTGTTGCCTATGGGTGCATCTGTCATAAATGATCATGAAACCATTCCT
GTCATCGATATAGAAAATTTATTATCTCCAGAACCAATAATCGGAAAGTTAGAATTAGATAGGCTTCATTTTGCTTGCAAAGAATGGGGTTTTTTTCAGG
TATATATGTCGTAAATACAAGCATTGTGATTGCATCCAGTAGTTGTACTTTGTTACACTAAATATTAGCTAGTGAATCTTACTTTACATATATATGTTGC
AGGTAGTGAACCATGGAGTCGACGCTTCATTGGTGGATAGTGTAAAATCAGAAATTCAAGGTTTCTTTAACCTTTCTATGGATGAGAAAACTAAATATGA
ACAGGAAGATGGAGATGTGGAAGGATTTGGACAAGGCTTTATTGAATCAGAGGAC[C/T]AAACACTTGATTGGGCAGATATATTTATGATGTTCACTCT
TCCACTCCATTTAAGGAAGCCTCACTTATTTTCAAAACTCCCAGTGCCTCTCAGGTTATTATTAATTCATAATCGGTTTCTATGCACTTGACTTTTTAAT
CTCACATTTTCGTAGTCGGTTTCTGTGCACCTCACTCAGTTTATTCTTATTCTTTTTTTTGTTTGGCCCACACATATTTAACATTTAGTCCACCTTGTTA
GAATTTCAGACTATAAATCATTCTTTACAATTCCTTATTAGCTGTAACCATGTTGTATGTGAACATATTTAATATTTTATGCTGATATATAATTATGTAT
ATGCATGCTCTAGAATTATATATTCTACCACTTGTAGGAAGATACTTCCAATTTTATCTTCTCTTATTCTTTAATGATGGAAAATTCCTTTATAAACAAG
AGCTAGAACGGAGTACACCTCACCTCATGCTAGGAAAGACATATGTCATTCAACCTTCCATCAAGAATTGAAAACAACCTCACATAGGATTATCATAACT
AGTATGTCATTCCAAATAGTGTAAGAGATATTTGTCCTTATGTACCCTTAAGCATACATAATTAACACAAACAAACCTTAATTCATTGATTAGGATTAGT
ACTAATCAAGTTTAAAGGAATACAAATTCTGTTTTACTGCTGAACCAACATAAAAAACTGACACACCTAAGTCTCCCTATTTCTATAACAAAAATTATCC
CTCTTGATATTATTTTCTGGTTCCCCCATTGACCAATAAACATGCAAAATTAACACTCTTTAAAAGAGTACGTATGATAGATCATAATATTTTTCTCTCT
TTGAATTTTCTCAGTGATGTGATTGGATACATTAATTTGGTTGGTTTATTATAGTTTTTCATATGATTATCTAATTTACGTACAGGGAGACAATCGAATC
CTACTCATCAGAAATGAAAAAGTTATCCATGGTTCTCTTTAATAAGATGGAAAAAGCTCTACAAGTACAAGCAGCCGAGATTAAGGGTATGTCAGAGGTG
TTTATAGATGGGACACAAGCAATGAGGATGAACTATTATCCCCCTTGTCCTCAACCAAATCTCGCCATCGGTCTTACGTCGCACTCGGATTTTGGCGGTT
TGACAATCCTCCTTCAAATCAACGAAGTGGAAGGATTACAGATAAAAAGAGAGGGGACATGGATTGCAGTCGAACCTCTACCTAATGCGTTCGTAGTGAA
TGTTGGAGATATTTTGGAGGTACTGGTCCTTGTTTCGAATACTGGTTTTGCCTGTACCCAGAAAATTTCTTATTGTTTTTTCTTTTACAGATAATGACTA
ATGGAATTTACCATAGTGTCGATCACCGGGCAGTAGTAAACTCAACAAATGAGAGGCTCTCAATCGCAATATTTCATGACCCTAGTCTAGAGTCGGTAAT
AGGCCCAATATCAAGCTTGATTACTCCAGAGACACCTGCTTTGTTTAAAAGTGGATCTACATATGGGGATCTTGTGGAGGAATGTAAAACAAGGAAGCTC
GATGGAAAATCATTTCTTGACTCCATGAGGATTTGAAAACTCAAGAAAAAATAATACGACGTGATTGCATGTCAGATTCAACTATCCTTTTGTTGTTTTT
TGGTGCTCGAGTCCTTAATTGTTTTGATCATTGCTTTTGATTCTAATTAATAAGACTTTTCTCAAGAACCACATGTAATGTACCTTTACTTTCAGAAAAT
AAAAAGTATTGAGGCACAAATGAGAAAATTGA

```
Oligo 5 :                                                    Pos.    Len.     Tm      GC%   Self Any
Self End
    Wildtype Forward Primer 5': GCTTTATTGAATCAGAGGAAC       536     21     53.32    38.10     5.00
1.00
    Mutant Forward Primer 5': GCTTTATTGAATCAGAGGAAT         536     21     52.88    33.33     5.00
2.00
    Common Reverse Primer 5': AGCATGAGGTGAGGTGTACT         1027     20     54.68    50.00     4.00
2.00
    Product Size: 492
```

*W145*_WebSNAPER*

```
GTTCTTAATTCATTAATTAATTTAGAAAAATCATGGAGAAAGCAAAACTTATGAAGCTAGGTAATGGTATGGAAATACCAAGTGTTCAAGAATTGGCTAAACTC
ACGCTTGCCGAAATTCCATCTCGATACGTATGCGCCAATGAAAACCTTTTGTTGCCTATGGGTGCATCTGTCATAAATGATCATGAAACCATTCCTGTCATCGA
TATAGAAAATTTATTATCTCCAGAACCAATAATCGGAAAGTTAGAATTAGATAGGCTTCATTTTGCTTGCAAAGAATGGGGTTTTTTTCAGGTATATATGTCGT
AAATACAAGCATTGTGATTGCATCCAGTAGTTGTACTTTGTTACACTAAATATTAGCTAGTGAATCTTACTTTACATATATATGTTGCAGGTAGTGAACCATGG
AGTCGACGCTTCATTGGTGGATAGTGTAAAATCAGAAATTCAAGGTTTCTTTAACCTTTCTATGGATGAGAAAACTAAATATGAACAGGAAGATGGAGATGTGG
AAGGATTTGGACAAGGCTTTATTGAATCAGAGGACCAAACACTTGATTG[G/A]GCAGATATATTTATGATGTTCACTCTTCCACTCCATTTAAGGAAGCCTCA
CTTATTTTCAAAACTCCCAGTGCCTCTCAGGTTATTATTAATTCATAATCGGTTTCTATGCACTTGACTTTTTAATCTCACATTTTCGTAGTCGGTTTCTGTGC
ACCTCACTCAGTTTATTCTTATTCTTTTTTTTGTTTGGCCCACACATATTTAACATTTAGTCCACCTTGTTAGAATTTCAGACTATAAATCATTCTTTACAATT
CCTTATTAGCTGTAACCATGTTGTATGTGAACATATTTAATATTTTATGCTGATATATAATTATGTATATGCATGCTCTAGAATTATATATTCTACCACTTGTA
GGAAGATACTTCCAATTTTATCTTCTCTTATTCTTTAATGATGGAAAATTCCTTTATAAACAAGAGCTAGAACGGAGTACACCTCACCTCATGCTAGGAAAGAC
ATATGTCATTCAACCTTCCATCAAGAATTGAAAACAACCTCACATAGGATTATCATAACTAGTATGTCATTCCAAATAGTGTAAGAGATATTTGTCCTTATGTA
CCCTTAAGCATACATAATTAACACAAACAAACCTTAATTCATTGATTAGGATTAGTACTAATCAAGTTTAAAGGAATACAAATTCTGTTTTACTGCTGAACCAA
CATAAAAAACTGACACACCTAAGTCTCCCTATTTCTATAACAAAAATTATCCCTCTTGATATTATTTTCTGGTTCCCCCATTGACCAATAAACATGCAAAATTA
ACACTCTTTAAAAGAGTACGTATGATAGATCATAATATTTTTCTCTCTTTGAATTTTCTCAGTGATGTGATTGGATACATTAATTTGGTTGGTTTATTATAGTT
TTTCATATGATTATCTAATTTACGTACAGGGAGACAATCGAATCCTACTCATCAGAAATGAAAAAGTTATCCATGGTTCTCTTTAATAAGATGGAAAAAGCTCT
ACAAGTACAAGCAGCCGAGATTAAGGGTATGTCAGAGGTGTTTATAGATGGGACACAAGCAATGAGGATGAACTATTATCCCCCTTGTCCTCAACCAAATCTCG
CCATCGGTCTTACGTCGCACTCGGATTTTGGCGGTTTGACAATCCTCCTTCAAATCAACGAAGTGGAAGGATTACAGATAAAAAGAGAGGGGACATGGATTGCA
GTCGAACCTCTACCTAATGCGTTCGTAGTGAATGTTGGAGATATTTTGGAGGTACTGGTCCTTGTTTCGAATACTGGTTTTGCCTGTACCCAGAAAATTTCTTA
TTGTTTTTTCTTTTACAGATAATGACTAATGGAATTTACCATAGTGTCGATCACCGGGCAGTAGTAAACTCAACAAATGAGAGGCTCTCAATCGCAATATTTCA
TGACCCTAGTCTAGAGTCGGTAATAGGCCCAATATCAAGCTTGATTACTCCAGAGACACCTGCTTTGTTTAAAAGTGGATCTACATATGGGGATCTTGTGGAGG
AATGTAAAACAAGGAAGCTCGATGGAAAATCATTTCTTGACTCCATGAGGATTTGAAAACTCAAGAAAAAATAATACGACGTGATTGCATGTCAGATTCAACTA
TCCTTTTGTTGTTTTTTGGTGCTCGAGTCCTTAATTGTTTTGATCATTGCTTTTGATTCTAATTAATAAGACTTTTCTCAAGAACCACATGTAATGTACCTTTA
CTTTCAGAAAATAAAAAGTATTGAGGCACAAATGAGAAAATTGA
```

Primer ID	Primer Sequence	Tm	Primer Len.	Product Len.	Warnings
W145*_L_ALT_2	TTTATTGAATCAGAGGAC CAAACACTTGAGTGA	67.0	33		
W145*_L_ALT_2_REV ERSE	ATGTTCACATACAACATG GTTACAGCTAATAAGGAA	65.7	36	325	

A146T_WebSNAPER

```
GTTCTTAATTCATTAATTAATTTAGAAAAATCATGGAGAAAGCAAAACTTATGAAGCTAGGTAATGGTATGGAAATACCAAGTGTTCAAGAATTGGCTAAACTC
ACGCTTGCCGAAATTCCATCTCGATACGTATGCGCCAATGAAAACCTTTTGTTGCCTATGGGTGCATCTGTCATAAATGATCATGAAACCATTCCTGTCATCGA
TATAGAAAATTTATTATCTCCAGAACCAATAATCGGAAAGTTAGAATTAGATAGGCTTCATTTTGCTTGCAAAGAATGGGGTTTTTTTCAGGTATATATGTCGT
AAATACAAGCATTGTGATTGCATCCAGTAGTTGTACTTTGTTACACTAAATATTAGCTAGTGAATCTTACTTTACATATATATGTTGCAGGTAGTGAACCATGG
AGTCGACGCTTCATTGGTGGATAGTGTAAAATCAGAAATTCAAGGTTTCTTTAACCTTTCTATGGATGAGAAAACTAAATATGAACAGGAAGATGGAGATGTGG
AAGGATTTGGACAAGGCTTTATTGAATCAGAGGACCAAACACTTGATTGG[G/A]CAGATATATTTATGATGTTCACTCTTCCACTCCATTTAAGGAAGCCTCA
CTTATTTTCAAAACTCCCAGTGCCTCTCAGGTTATTATTAATTCATAATCGGTTTCTATGCACTTGACTTTTTAATCTCACATTTTCGTAGTCGGTTTCTGTGC
ACCTCACTCAGTTTATTCTTATTCTTTTTTTTGTTTGGCCCACACATATTTAACATTTAGTCCACCTTGTTAGAATTTCAGACTATAAATCATTCTTTACAATT
CCTTATTAGCTGTAACCATGTTGTATGTGAACATATTTAATATTTTATGCTGATATATAATTATGTATATGCATGCTCTAGAATTATATATTCTACCACTTGTA
GGAAGATACTTCCAATTTTATCTTCTCTTATTCTTTAATGATGGAAAATTCCTTTATAAACAAGAGCTAGAACGGAGTACACCTCACCTCATGCTAGGAAAGAC
ATATGTCATTCAACCTTCCATCAAGAATTGAAAACAACCTCACATAGGATTATCATAACTAGTATGTCATTCCAAATAGTGTAAGAGATATTTGTCCTTATGTA
CCCTTAAGCATACATAATTAACACAAACAAACCTTAATTCATTGATTAGGATTAGTACTAATCAAGTTTAAAGGAATACAAATTCTGTTTTACTGCTGAACCAA
CATAAAAAACTGACACACCTAAGTCTCCCTATTTCTATAACAAAAATTATCCCTCTTGATATTATTTTCTGGTTCCCCCATTGACCAATAAACATGCAAAATTA
ACACTCTTTAAAAGAGTACGTATGATAGATCATAATATTTTTCTCTCTTTGAATTTTCTCAGTGATGTGATTGGATACATTAATTTGGTTGGTTTATTATAGTT
TTTCATATGATTATCTAATTTACGTACAGGGAGACAATCGAATCCTACTCATCAGAAATGAAAAAGTTATCCATGGTTCTCTTTAATAAGATGGAAAAAGCTCT
ACAAGTACAAGCAGCCGAGATTAAGGGTATGTCAGAGGTGTTTATAGATGGGACACAAGCAATGAGGATGAACTATTATCCCCCTTGTCCTCAACCAAATCTCG
CCATCGGTCTTACGTCGCACTCGGATTTTGGCGGTTTGACAATCCTCCTTCAAATCAACGAAGTGGAAGGATTACAGATAAAAAGAGAGGGGACATGGATTGCA
GTCGAACCTCTACCTAATGCGTTCGTAGTGAATGTTGGAGATATTTTGGAGGTACTGGTCCTTGTTTCGAATACTGGTTTTGCCTGTACCCAGAAAATTTCTTA
TTGTTTTTTCTTTTACAGATAATGACTAATGGAATTTACCATAGTGTCGATCACCGGGCAGTAGTAAACTCAACAAATGAGAGGCTCTCAATCGCAATATTTCA
TGACCCTAGTCTAGAGTCGGTAATAGGCCCAATATCAAGCTTGATTACTCCAGAGACACCTGCTTTGTTTAAAAGTGGATCTACATATGGGGATCTTGTGGAGG
AATGTAAAACAAGGAAGCTCGATGGAAAATCATTTCTTGACTCCATGAGGATTTGAAAACTCAAGAAAAAATAATACGACGTGATTGCATGTCAGATTCAACTA
TCCTTTTGTTGTTTTTTGGTGCTCGAGTCCTTAATTGTTTTGATCATTGCTTTTGATTCTAATTAATAAGACTTTTCTCAAGAACCACATGTAATGTACCTTTA
CTTTCAGAAAATAAAAAGTATTGAGGCACAAATGAGAAAATTGA
```

Primer ID	Primer Sequence	Tm	Primer Len.	Product Len.	Warnings
A146T_R_ALT_1	ATGGAGTGGAAGAGTGAA CATCATAAATATATCCGT	67.2	36		
A146T_R_ALT_1_REV ERSE	AATCGGAAAGTTAGAATT AGATAGGCTTCATTTTGC	67.0	36	368	

Appendix E- Record of crosses carried out

Cross	Mother plant		Father plant		Date of emasc-ulation	Date of cross	Assigned seed_ batch_id	Description
	plant_id	seed_batch _id	plant_id	seed_batch_ id				
X001	Sd-833543	HM6	Sd-833401	S-114248	11-Sep-13	12-Sep-13	S-123957	HM6 x CODM E193K/R158K
X002	Sd-833557	HM6	Sd-833411	S-114248	11-Sep-13	12-Sep-13	S-123954	HM6 x CODM E193K/R158K
X003	Sd-833509	S-122112	Sd-833401	S-114248	9-Sep-13	12-Sep-13	S-124007	CODMb Q254* x CODM E193K/R158K
X004	Sd-833595	HN4	Sd-833411	S-114248	11-Sep-13	12-Sep-13	S-123974	HN4 x CODM E193K/R158K
X005	Sd-833452	S-122067	Sd-833868	S-125004	9-Sep-13	12-Sep-13	S-124004	CODM E193K/R158K x T/O only from HN3 mut
X006	Sd-833421	S-122066	Sd-833868	S-125004	9-Sep-13	12-Sep-13	S-123959	CODM E193K/R158K x T/O only from HN3 mut
X007	Sd-833598	HN4	Sd-833517	S-122112	11-Sep-13	14-Sep-13	S-123973	HN4 x CODMb Q254*
X008	Sd-833541	HM6	Sd-833508	S-122112	12-Sep-13	14-Sep-13	S-123958	HM6 x CODMb Q254*
X009	Sd-833514	S-122112	Sd-833385	S-114247	11-Sep-13	14-Sep-13	S-124006	CODMb Q254* x CODM E193K/R158K
X010	Sd-833617	S-125001	Sd-833418	S-114248	12-Sep-13	14-Sep-13	S-123988	T/O only from HN3 mut x CODM E193K/R158K
X011	Sd-833596	HN4	Sd-833402	S-114248	11-Sep-13	14-Sep-13	S-123975	HN4 x CODM E193K/R158K
X012	Sd-833386	S-114247	Sd-833542	HM6	11-Sep-13	14-Sep-13	S-123984	CODM E193K/R158K x HM6
X013	Sd-833569	HT6	Sd-833464	S-114249	14-Sep-13	16-Sep-13	S-123970	HT6 x CODMb/T6ODMa mutant
X014	Sd-833498	S-114250	Sd-833464	S-114249	14-Sep-13	16-Sep-13	S-124001	High T x CODMb/T6ODMa mutant
X015	Sd-833503	S-122112	Sd-833464	S-114249	14-Sep-13	16-Sep-13	S-124009	CODMb Q254* x CODMb/T6ODMa mutant
X016	Sd-833580	HN4	Sd-833464	S-114249	14-Sep-13	16-Sep-13	S-123972	HN4 x CODMb/T6ODMa mutant
X017	Sd-833488	S-114250	Sd-833545	HM6	14-Sep-13	16-Sep-13	S-123997	High T x HM6
X018	Sd-833483	S-114250	Sd-833850	S-125003	16-Sep-13	18-Sep-13	S-123998	High T x T/O only from HN3 mut
X019	Sd-833619	S-125001	Sd-833697	S-114139	16-Sep-13	18-Sep-13	S-123989	T/O only from HN3 mut x CODMb/T6ODMa mutant
X020	Sd-833620	S-114248	Sd-833566	HT6	16-Sep-13	18-Sep-13	S-124038	CODM E193K/R158K x HT6
X021	Sd-833789	HN4	Sd-833827	S-114250	16-Sep-13	18-Sep-13	S-124071	HN4 x High T
X022	Sd-833825	S-114250	Sd-833649	S-122067	16-Sep-13	18-Sep-13	S-124018	High T x CODM E193K/R158K
X023	Sd-833585	HN4	Sd-833873	S-125004	16-Sep-13	18-Sep-13	S-123976	HN4 x T/O only from HN3 mut
X024	Sd-833520	S-122152	Sd-833380	S-114247	16-Sep-13	18-Sep-13	S-123965	CODMa/c W261* x CODM E193K/R158K
X025	Sd-833481	S-114250	Sd-833532	S-122152	17-Sep-13	19-Sep-13	S-124002	High T x CODMa/c W261*
X026	Sd-833538	S-122152	Sd-833553	HM6	17-Sep-13	19-Sep-13	S-123963	CODMa/c W261* x HM6
X027	Sd-833559	HM6	Sd-833434	S-122066	16-Sep-13	19-Sep-13	S-123955	HM6 x CODM E193K/R158K
X028	Sd-833815	S-114250	Sd-833806	S-122066	17-Sep-13	19-Sep-13	S-124021	High T x CODM E193K/R158K
X029	Sd-833583	HN4	Sd-833842	S-125002	16-Sep-13	19-Sep-13	S-123977	HN4 x T/O only from HN3 mut
X030	Sd-833578	HT6	Sd-833390	S-114247	16-Sep-13	19-Sep-13	S-123971	HT6 x CODM E193K/R158K
X031	Sd-833555	HM6	Sd-833853	S-125003	18-Sep-13	20-Sep-13	S-123956	HM6 x T/O only from HN3 mut
X032	Sd-833536	S-122152	Sd-833599	HN4	18-Sep-13	20-Sep-13	S-123962	CODMa/c W261* x HN4
X033	Sd-833714	S-125001	Sd-833818	S-114250	18-Sep-13	20-Sep-13	S-124073	T/O only from HN3 mut x High T
X034	Sd-833755	S-122152	Sd-833895	S-114247	18-Sep-13	20-Sep-13	S-124025	CODMa/c W261* x CODM E193K/R158K

Cross	Mother plant		Father plant		Date of emasc-ulation	Date of cross	Assigned seed_ batch_id	Description
	plant_id	seed_batch _id	plant_id	seed_batch_ id				
X035	Sd-833720	S-122122	Sd-833489	S-114250	23-Sep-13	25-Sep-13	S-124057	CODMb Q254* x High T
X036	Sd-833725	S-122122	Sd-833851	S-125003	23-Sep-13	25-Sep-13	S-124052	CODMb Q254* x T/O only from HN3 mut
X037	Sd-833737	S-122122	Sd-833564	HT6	23-Sep-13	25-Sep-13	S-124043	CODMb Q254* x HT6
X038	Sd-833681	S-114139	Sd-833759	S-122152	23-Sep-13	25-Sep-13	S-124067	CODMb/T6ODMa mutant x CODMa/c W261*
X039	Sd-833778	HT6	Sd-833533	S-122152	23-Sep-13	25-Sep-13	S-124075	HT6 x CODMa/c W261*
X040	Sd-833770	HT6	Sd-833619	S-125001	23-Sep-13	25-Sep-13	S-124076	HT6 x T/O only from HN3 mut
X041	Sd-833680	S-114139	Sd-833604	S-125001	23-Sep-13	25-Sep-13	S-124064	CODMb/T6ODMa mutant x T/O only from HN3 mut
X042	Sd-833560	HT6	Sd-833486	S-114250	23-Sep-13	25-Sep-13	S-123968	HT6 x High T
X043	Sd-833577	HT6	Sd-833816	S-114250	23-Sep-13	25-Sep-13	S-123966	HT6 x High T
X044	Sd-833690	S-114139	Sd-833824	S-114250	23-Sep-13	25-Sep-13	S-124066	CODMb/T6ODMa mutant x High T
X045	Sd-833702	S-125001	Sd-833728	S-122122	25-Sep-13	27-Sep-13	S-124074	T/O only from HN3 mut x CODMb Q254*
X046	Sd-833839	S-125002	Sd-833746	S-122152	25-Sep-13	27-Sep-13	S-124077	T/O only from HN3 mut x CODMa/c W261*
X047	Sd-833753	S-122152	Sd-833719	S-125001	25-Sep-13	27-Sep-13	S-124037	CODMa/c W261* x T/O only from HN3 mut
X048	Sd-833750	S-122152	Sd-833773	HT6	25-Sep-13	27-Sep-13	S-124031	CODMa/c W261* x HT6
X049	Sd-833962	HM6	Sd-833695	S-114139	1-Oct-13	3-Oct-13	S-124069	HM6 x CODMb/T6ODMa
X050	Sd-833459	S-122067	Sd-833684	S-114139	1-Oct-13	3-Oct-13	S-124003	CODM E193K/R158K x CODMb/T6ODMa
X051	Sd-834053	S-114139	Sd-834078	S-114156	7-Oct-13	9-Oct-13	S-124013	CODMb/T6ODMa x CODM E193K/R158K
X052	Sd-833966	S-122152	Sd-834039	S-114118	7-Oct-13	9-Oct-13	S-124040	CODMa/c W261* x High T
X053	Sd-833989	S-122122	Sd-833950	HM6	7-Oct-13	9-Oct-13	S-124014	CODM Q254* x HM6
X054	Sd-833985	S-122122	Sd-834043	S-114139	7-Oct-13	9-Oct-13	S-124015	CODMb Q254* x CODMb F271L/T6ODMa
X055	Sd-833942	HT6	Sd-833628	S-114248	7-Oct-13	9-Oct-13	S-124068	HT6 x CODM E193K/R158K
X056	Sd-833920	HN4	Sd-833751	S-122152	7-Oct-13	9-Oct-13	S-124061	HN4 x CODMa/c W261*
X057	Sd-835248	S-114118	Sd-835193	S-123361	14-Jan-14	16-Jan-14	S-122954	HighT x T6ODMa W145* (homo)
X058	Sd-835269	S-122122	Sd-835198	S-123361	14-Jan-14	16-Jan-14	S-122968	CODMb Q254* x T6ODMa W145* (homo)
X059	Sd-835262	S-122122	Sd-835207	S-123349	14-Jan-14	16-Jan-14	S-122966	CODMb Q254* x T6ODMa Q141* (het)
~~X060~~	~~Sd-835205~~	~~S-123349~~	~~Sd-835280~~	~~S-114248~~	~~14-Jan-14~~	~~16-Jan-14~~	~~S-122944~~	~~T6ODMa Q141* (wt) x CODM E193K/R158K~~
X061	Sd-835261	S-122122	Sd-835187	S-123361	16-Jan-14	18-Jan-14	S-122969	CODMb Q254* x T6ODMa W145* (homo)
X062	Sd-835266	S-122122	Sd-835210	S-123349	16-Jan-14	18-Jan-14	S-122970	CODMb Q254* x T6ODMa Q141* (het)
X063	Sd-835250	S-114118	Sd-835189	S-123361	16-Jan-14	18-Jan-14	S-122952	HighT x T6ODMa W145* (homo)
X064	Sd-835292	S-114248	Sd-835193	S-123361	18-Jan-14	20-Jan-14	S-122956	CODM E193K/R158K x T6ODMa W145* (homo)
X065	Sd-835221	HT5	Sd-835184	S-123361	18-Jan-14	20-Jan-14	S-122962	HT5 x T6ODMa W145* (homo)
X066	Sd-835279	S-122122	Sd-835192	S-123361	18-Jan-14	20-Jan-14	S-122967	CODMb Q254* x T6ODMa W145* (homo)
X067	Sd-835230	HT5	Sd-835182	S-123361	20-Jan-14	22-Jan-14	S-122963	HT5 x T6ODMa W145* (homo)
~~X068~~	~~Sd-835233~~	~~HT5~~	~~Sd-835212~~	~~S-123349~~	~~20-Jan-14~~	~~22-Jan-14~~	~~S-122694~~	~~HT5 x T6ODMa Q141* (wt)~~
X069	Sd-835238	HT5	Sd-835204	S-123349	20-Jan-14	22-Jan-14	S-122965	HT5 x T6ODMa Q141* (het)

Cross	Mother plant		Father plant		Date of emasc-ulation	Date of cross	Assigned seed_ batch_id	Description
	plant_id	seed_batch _id	plant_id	seed_batch_ id				
X070	Sd-835288	S-114248	Sd-835182	S-123361	20-Jan-14	22-Jan-14	No seed	CODM E193K/R158K x T6ODMa W145* (homo)
X071	Sd-835583	HM6	Sd-835195	S-123361	20-Jan-14	22-Jan-14	S-122979	HM6 x T6ODMa W145* (homo)
~~X072~~	~~Sd-835247~~	~~S-114118~~	~~Sd-835218~~	~~S-123349~~	~~20-Jan-14~~	~~23-Jan-14~~	~~S-122953~~	~~High T Fwd x T6ODMa Q141* (wt)~~
~~X073~~	~~Sd-835284~~	~~S-114248~~	~~Sd-835218~~	~~S-123349~~	~~20-Jan-14~~	~~23-Jan-14~~	~~S-122957~~	~~CODM E193K/R158K x T6ODMa Q141* (wt)~~
X074	Sd-835253	S-114118	Sd-835201	S-123349	25-Jan-14	27-Jan-14	S-122955	High T Fwd x T6ODMa Q141* (het)
X075	Sd-835694	HT6	Sd-835673	S-124115	04-Mar-14	06-Mar-14	No seed	HT6 x T6ODMa A146T (het)
~~X076~~	~~Sd-835701~~	~~HT6~~	~~Sd-835682~~	~~S-124115~~	~~04-Mar-14~~	~~06-Mar-14~~	~~N/A~~	~~HT6 x T6ODMa A146T (wt)~~
Crosses involving F1s, BC1s and F2s								
X077	Sd-835585	HM6	Sd-835304	S-124006	18-Jan-14	20-Jan-14	S-122980	HM6 x X009
X078	Sd-835587	HM6	Sd-835413	S-123955	20-Jan-14	22-Jan-14	S-122978	HM6 x X027
X079	Sd-835580	HN4	Sd-835445	S-123972	20-Jan-14	22-Jan-14	S-122972	HN4 x X016
X080	Sd-835568	S-114139	Sd-835330	S-124013	20-Jan-14	22-Jan-14	S-122973	High T Fwd x X051
X081	Sd-835552	HT6	Sd-835427	S-123971	22-Jan-14	24-Jan-14	S-122977	HT6 x X030
~~X082~~	~~Sd-831987~~	~~HT5~~	~~Sd-837839~~	~~S-122977~~	~~28-May-14~~	~~29-May-14~~	~~TBD~~	~~HT5 x X081 E193K/R158K (wt)~~
~~X083~~	~~Sd-831956~~	~~HM6~~	~~Sd-837806~~	~~S-122978~~	~~28-May-14~~	~~29-May-14~~	~~TBD~~	~~HM6 x X078 E193K/R158K (wt)~~
~~X084~~	~~Sd-831930~~	~~HN4~~	~~Sd-837759~~	~~S-122972~~	~~28-May-14~~	~~29-May-14~~	~~TBD~~	~~HN4 x X079 Tint (wt)~~
X085	Sd-831981	HT5	Sd-831873	S-122962	04-Jun-14	05-Jun-14	S-186688	HT5 x X065

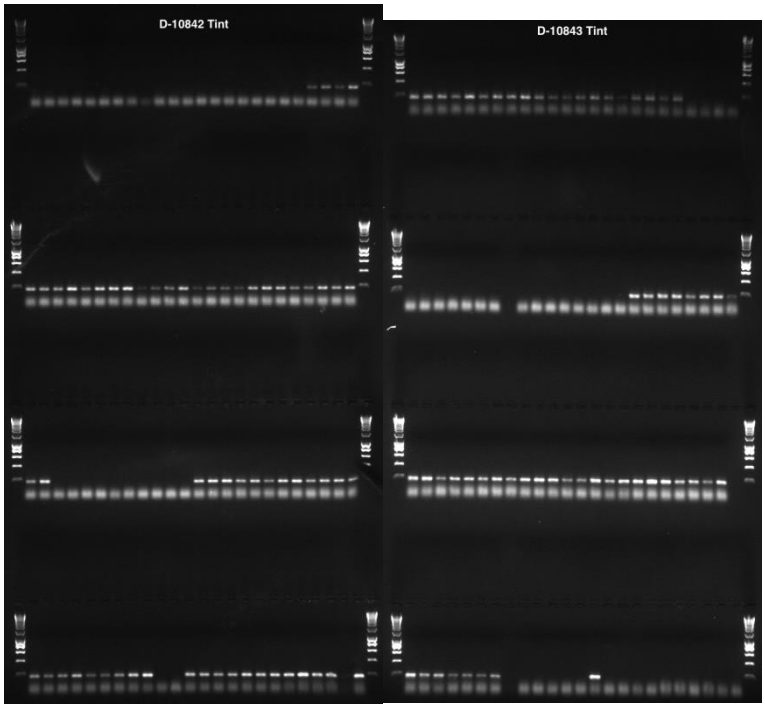
Crosses struckthrough indicate instances where the parent intended to be the transmitter of a CODM/T6ODM mutation was subsequently genotyped as wild type. This happened when the mutations were still segregating in the line sown to provide the parent in a cross. The plants flowered before genotyping was completed.

Appendix F- Genotyping of F1 individuals to confirm presence of alleles contributed by mutants

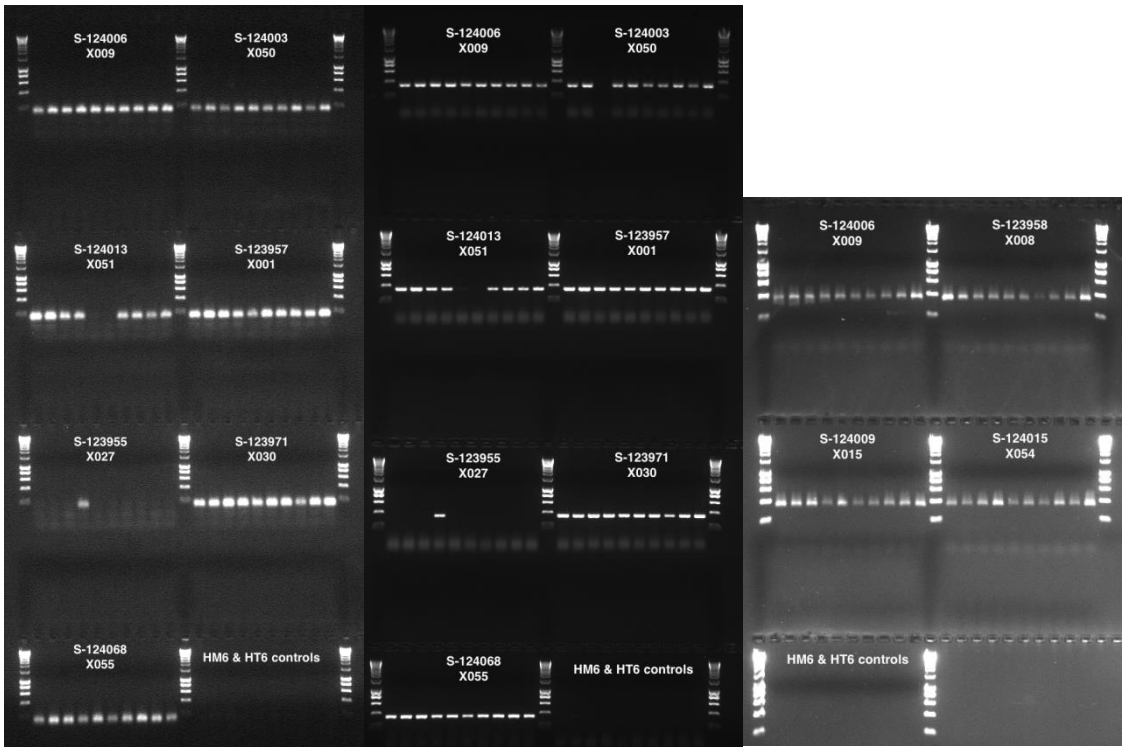
First batch of F1s-agarose gel visualisation

All samples of both D-10842 and D-10843 DNA plates were screened for the T6ODM Tint polymorphism. Only those samples relevant for the CODM E193K, R158K and Q254* polymorphisms were screened for these markers. Presence of an appropriately sized PCR product indicated possession the marker.

Tint 164 bp
E193K 138 bp
R158K 345 bp
Q254* 339 bp



D-10842 Tint D-10843 Tint

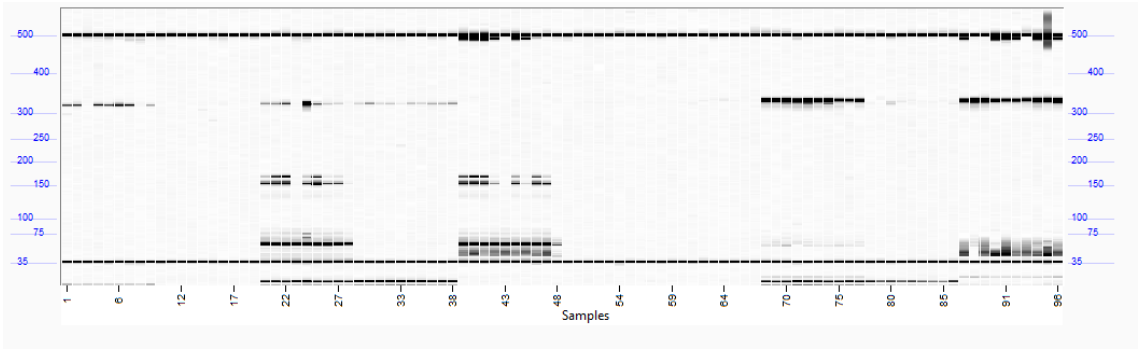


E193K R158K Q254*

Second batch of F1s-visualisation on Fragment Analyser

DNA plate D-10863 contained DNA from approximately ten plants of the following F1 seed batches:
S-123963 X026
S-123956 X031
S-122962 X065
S-122979 X071
S-122965 X069
S-123976 X023
S-123977 X029
S-122969 X061
S-122956 X064
S-122966 X059

F1s were screened for the Q254* (339 bp), R158K (345 bp), W261* (327 bp), Q141* (492 bp), W145* (325 bp), A146T (368 bp) and Tint (164 bp) SNPs where appropriate. AS-PCR products generated with each DNA sample are displayed on the TILLING gel below. Well A1 = lane 1, well B1 = lane 2 etc. Both these samples contain the CODMa/c W261* SNP (327 bp PCR product). Well D3 (lane 20) contains Tint (164 bp) and W145* (325 bp) SNPs.



D-10863

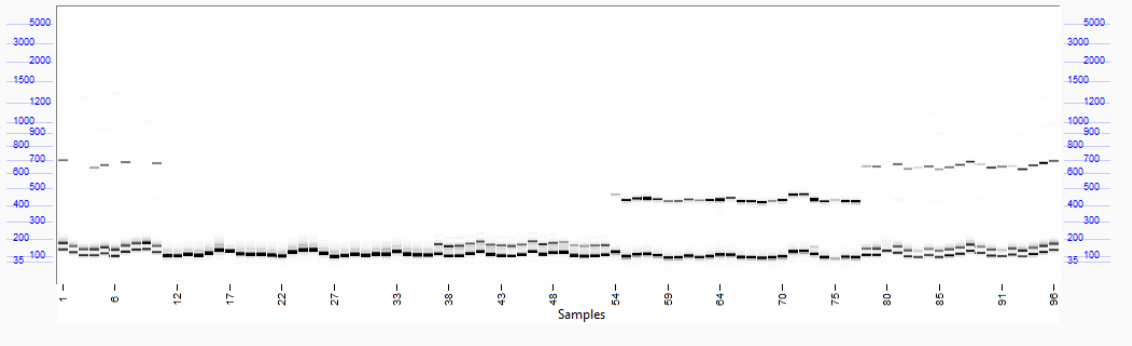
	1	2	3	4	5	6	7	8	9	10	11	12
A	W261*	W261*		W145* Tint	W145*	Q141* Tint				Q254* W145*		Q254*
B	W261*			W145* Tint	W145*	Q141* Tint				Q254* W145*		Q254* Q141*
C				W145* Tint	W145*					Q254* W145*		Q254* Q141*
D	W261*		W145* Tint		W145*	Q141* Tint			Q254* W145*	Q254* W145*		Q254* Q141*
E	W261*		W145* Tint	W145*	W145*	Q141*			Q254* W145*	Q254* W145*		Q254*
F	W261*		W145* Tint	W145*	W145*	Tint			Q254* W145*			Q254* Q141*
G	W261*			W145*	Q141* Tint	Tint			Q254* W145*		Q254* Q141*	Q254* Q141*
H			W145* Tint	W145*	Q141* Tint				Q254* W145*	**R158K het** W145*	Q254*	Q254*

DNA plate D-10864 contained DNA from approximately ten plants of the following F1 seed batches:
S-122955 X074 (A1-B2)
S-122997 X075 (E5-E7)

The plate also contained DNA from M4 T6ODMa TILLING mutants S-124986 (A146T), S-124789 (W145*) and S-122951 (Q141*).

For X074 F1s well A1 (lane 1) contained the Q141* mutation but B1 (lane 2) did not. The SNP was expected to be present in half of the F1s because the Q141* parent in the cross was a heterozygote (Appendix E). None of the F1s from the X075 cross carried the A146T SNP (lanes 37-53) so work with this cross was terminated at this stage.

Lanes 54 to 77 and 78 to 96 show the detection (in M4s) of the W145* and Q141* SNPs, respectively.



D-10864

	1	2	3	4	5	6	7	8	9	10	11	12
A	Q141*					Tint	Tint	W145*	W145*	W145*	Q141*	Q141*
B		Q141*				Tint	Tint	W145*	W145*	W145*	Q141*	Q141*
C						Tint	Tint	W145*	W145*	W145*	Q141*	Q141*
D	Q141*					Tint	Tint	W145*	W145*	W145*	Q141*	
E	Q141*				Tint	Tint	Tint	W145*	W145*	W145*	Q141*	Q141*
F					Tint	Tint	W145*	W145*	W145*	Q141*	Q141*	Q141*
G	Q141*				Tint	Tint	W145*	W145*	W145*	Q141*	Q141*	Q141*
H					Tint	Tint	W145*	W145*	W145*		Q141*	Q141*

F2 seed collection

Mother Seed batch id Seedling id Known mutations	Father Seed batch id Seedling id Known mutations	F1 seed batch id Cross id	F1 Seedling id	F2 seed batch id sown
S-122122 Sd-835261 **CODMb Q254***	S-123361 Sd-835187 **T6ODMa W145***	S-122969 X061	Sd-832352	S-186714
S-114248 Sd-835292 **CODMa/c E193K** **CODMb R158K**	S-123361 Sd-835193 **T6ODMa W145***	S-122956 X064	Sd-832263	S-186711
HT5 Sd-835221 T line **T6ODMa intron SNP**	S-123361 Sd-835184 **T6ODMa W145***	S-122962 X065	Sd-831875	S-186696
S-122122 Sd-835262 **CODMb Q254***	S-123349 Sd-835207 **T6ODMa Q141***	S-122966 X059	Sd-832370	S-186710
HT5 Sd-835238 T line **T6ODMa intron SNP**	S-123349 Sd-835204 **T6ODMa Q141***	S-122965 X069	Sd-831891	S-186698
S-114118 Sd-835253 High T Fwd	S-123349 Sd-835201 **T6ODMa Q141***	S-122955 X074	Sd-832380	S-186704
S-122152 Sd-833538 **CODMa/c W261***	HM6 Sd-833553 M line	S-123963 X026	Sd-831850	S-186689

Mother Seed batch id Seedling id Known mutations	Father Seed batch id Seedling id Known mutations	F1 seed batch id Cross id	F1 Seedling id	F2 seed batch id sown
HM6 Sd-833553 M line	S-125003 Sd-833853 T/O line	S-123956 X031	Sd-831867	S-186693
HM6 Sd-835583 M line	S-123361 Sd-835195 **T6ODMa W145***	S-122979 X071	Sd-831888	S-186703
HN4 Sd-833585 N line	S-125004 Sd-833873 T/O line	S-123976 X023	Sd-831907 Sd-831906	S-186682 (Tas) S-186684 (York)
HN4 Sd-833583 N line	S-125002 Sd-833842 T/O line	S-123977 X029	Sd-831918 Sd-831913	S-186685 (Tas) S-186686 (York)

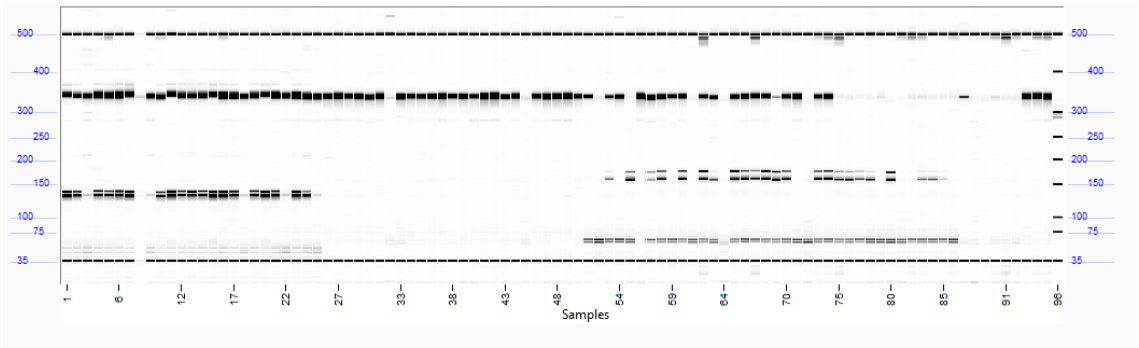
Appendix G- Genotyping of F2, F3 and BC1 individuals grown in York

F2s and BC1s were screened for the Q254* (339 bp), R158K (345 bp), W261* (327 bp), Q141* (492 bp), W145* (325 bp), A146T (368 bp) and Tint (164 bp) SNPs where appropriate. CODMb heterozygotes indicated by predicted fragment sizes after treatment with cleavage enzyme i.e. R158K (919 bp+203 bp), F271L (843 bp+278 bp) and Q254* (789 bp+332 bp). Likewise, T6ODMa heterozygotes were indicated Tint (359 bp+267 bp), Q141* (450 bp+161 bp) and W145* (468 bp+151 bp).

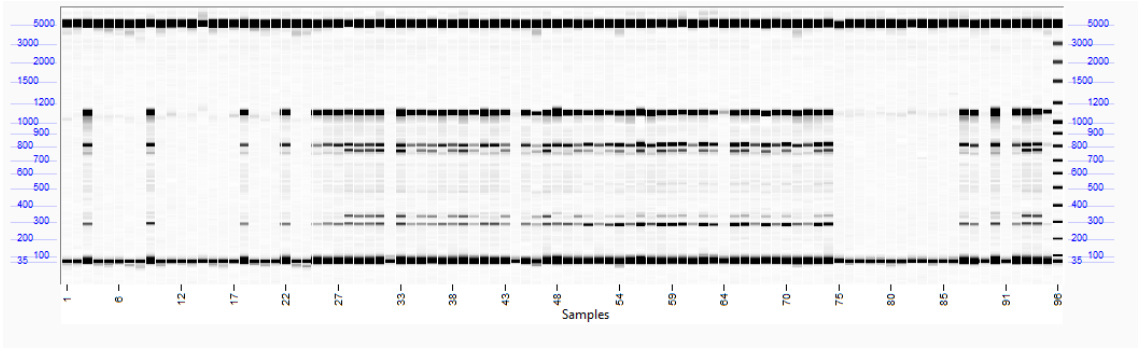
D-10858

X009 F2	S-122981	A1-A4
X008 F2	S-122988	B4-B7
X015 F2	S-122991	C7-B10
X041 F2	S-122994	C10-F11
X008 F2	S-122988	G11-G12

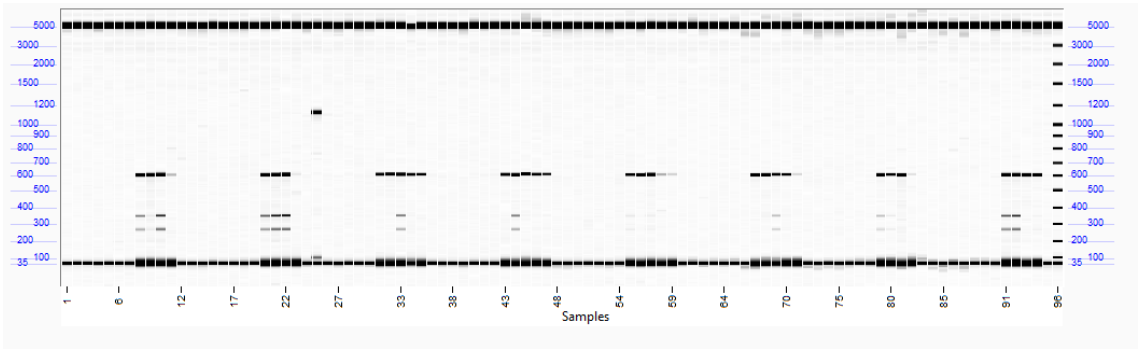
D-10858 Mutation screen



D-10858 CODMb TILL



D-10858 T6ODMa TILL



D-10858

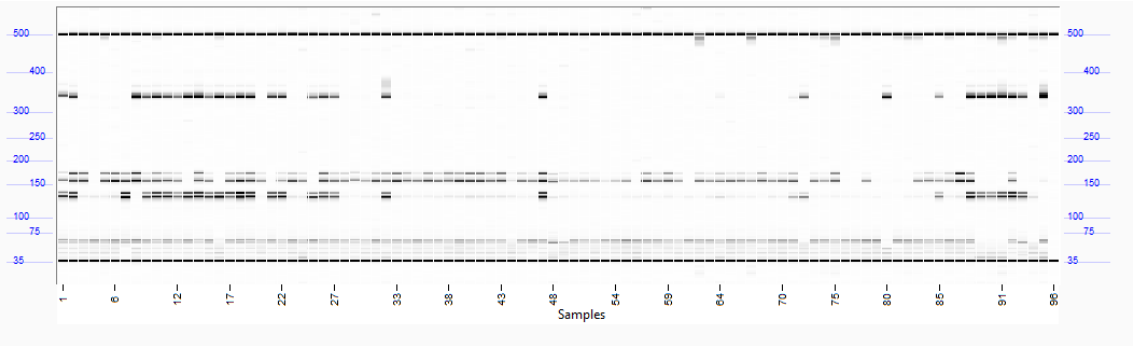
	1	2	3	4	5	6	7	8	9	10	11	12
A	E193K **R158K homo** **Q254* wt**	**E193K wt** **Q254* homo** F271L het	E193K R158K Q254*	E193K Q254* homo F271L het	Q254* het F271L het	Q254* homo F271L het	Q254* homo F271L het	**Q254* homo** F271L het **Tint het**	**Q254* het** F271L het **Tint homo**	Q254* het F271L het Tint het	Tint homo	**Q254* wt**
B	E193K **R158K het** **Q254* het**	E193K R158K Q254*	**E193K wt** **Q254* homo** F271L het	**Q254* homo** F271L het	Q254* homo F271L het	Q254* homo F271L het	Q254* het F271L het	Q254* het F271L het Tint het	Q254* het F271L het Tint het	Q254* het F271L het Tint het		F271L het
C	**E193K wt** **Q254* homo** F271L het	E193K R158K Q254*	E193K R158K Q254*	**Q254* homo** F271L het	Q254* het F271L het	Q254* het F271L het	F271L het Q254* het Tint homo	F271L het Q254* het Tint homo	**Q254* homo** F271L het Tint het	**Tint homo**	**Tint homo**	
D	E193K R158K Q254*	E193K R158K Q254*	E193K R158K Q254*	Q254* het F271L het	Q254* het F271L het	Q254* rxn fail	F271L het Tint homo	Q254* het F271L het Tint het	**Q254* het** F271L het Tint homo	Tint homo	Tint homo	F271L het
E	E193K R158K Q254*	E193K R158K Q254*	E193K R158K Q254*	**Q254* het** F271L het	Q254* homo F271L het	Q254* homo F271L het	**Q254* het** F271L het **Tint het**	F271L het Tint homo	F271L het Tint homo	Tint homo	Tint homo	Q254* het F271L het
F	E193K R158K Q254*	E193K R158K Q254*	**E193K wt** **Q254* homo** F271L het	Q254* het F271L het	Q254* het F271L het	Q254* het F271L het	**Q254* het** F271L het Tint homo	Q254* het F271L het Tint homo	Q254* het F271L het Tint het	Tint homo	Tint homo	Q254* het F271L het
G	E193K R158K Q254*	E193K R158K Q254*	E193K R158K Q254*	**Q254* het** F271L het	Q254* het F271L het	Q254* het F271L het	F271L het Tint het	Q254* het F271L het Tint het/homo?	**Q254* homo** F271L het **Tint WT**	Tint homo	Q254* homo F271L het	Q254* homo
H		E193K R158K Q254*	E193K R158K Q254*		Q254* het F271L het	Q254* homo F271L het	Q254* het F271L het Tint het	Tint het	F271L het Tint homo	Tint homo	**Q254* wt** F271L het	

Those genotypes in bold have been confirmed by sequencing.

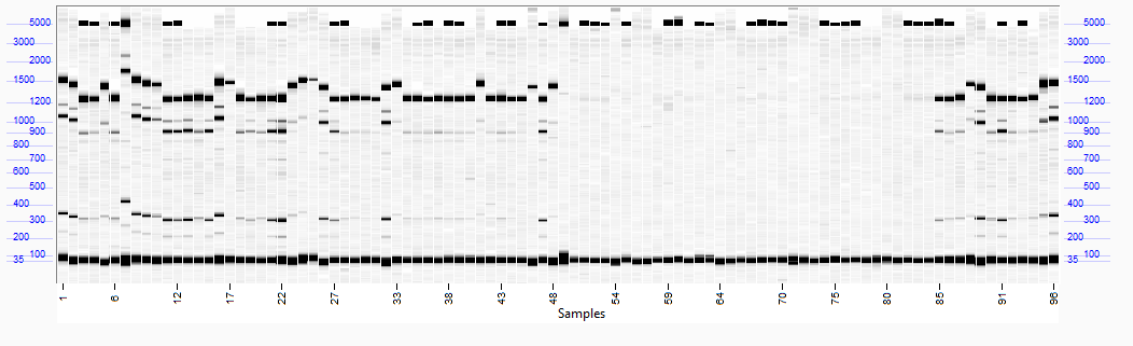
D-10859

X051 F2	S-122983	A1-A4
X051 BC1	S-122973	B4-A7
X042 F2	S-122989	B7-D8
X014 F2	S-122990	E8-D11
X051 F2	S-122983	E11-H12

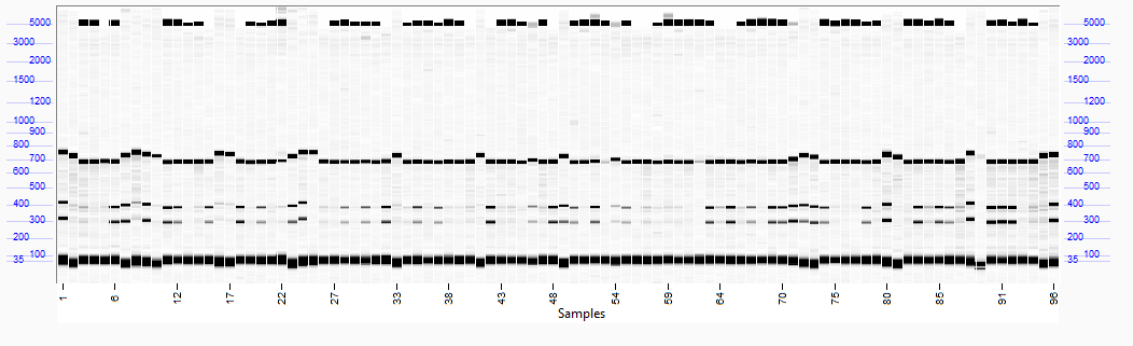
D-10859 Mutation screen



D-10859 CODMb TILL



D-10859 T6ODMa TILL



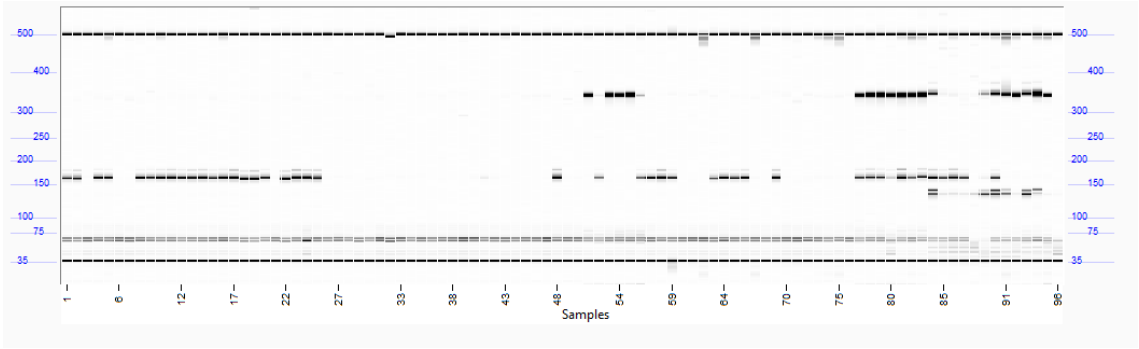
D-10859

	1	2	3	4	5	6	7	8	9	10	11	12
A	E193K R158K het F271L het Tint het	E193K R158K het F271L het Tint het	**E193K homo** **R158K homo** **Tint homo**	E193K R158K homo	F271L het Tint homo	F271L het Tint homo	Tint het	Tint het	Tint het	Tint het		E193K R158K het F271L het Tint ? rxn failed
B	E193K R158K het F271L het Tint homo	E193K R158K het F271L het	E193K R158K homo? F271L het Tint het	**E193K het** **R158K het** F271L het **Tint homo**	F271L het Tint homo	F271L het Tint het	Tint het	Tint homo	**Tint homo**	Tint het		E193K **R158K homo** F271L het? **Tint het**
C	F271L het Tint homo	E193K R158K het F271L het Tint het	E193K R158K het F271L het Tint homo	**E193K het** **R158K het** F271L het **Tint homo**	F271L het Tint het	F271L het Tint homo	Tint homo	**Tint homo**	Tint het	**Tint homo**	Tint het	E193K R158K het F271L het Tint het
D	F271L het	E193K R158K het F271L het Tint het	F271L het Tint het	F271L het Tint het	F271L het Tint homo	F271L het Tint homo	Tint het	Tint homo	**Tint homo**		Tint het	E193K R158K F271L het Tint het
E	F271L het Tint homo	E193K R158K het F271L het	E193K R158K het F271L het	F271L het	F271L het Tint het	F271L het Tint homo	Tint homo		Tint het		E193K R158K het? F271L het Tint het	E193K R158K F271L het
F	F271L het Tint het	E193K R158K homo? F271L het Tint homo	E193K R158K het F271L het Tint homo	F271L het Tint het	F271L het Tint homo	Tint het	Tint het	Tint homo	Tint het	Tint het	F271L het Tint het	F271L het
G	E193K F271L het Tint het	E193K R158K het F271L het Tint het	 F271L het Tint het	 F271L het	 F271L het Tint homo	**E193K het** **R158K het** F271L het **Tint homo**	 Tint het	 Tint het	 Tint het	 	 F271L het Tint het	 R158K homo? F271L het
H	R158K het F271L het Tint het	E193K R158K het F271L het	F271L het Tint het	E193K R158K het F271L het Tint het	F271L het Tint homo	R158K het F271L het Tint het	Tint homo	Tint het	E193K? R158K? Tint het	R158K? Tint het	E193K R158K homo? F271L het Tint het	R158K het F271L het Tint het

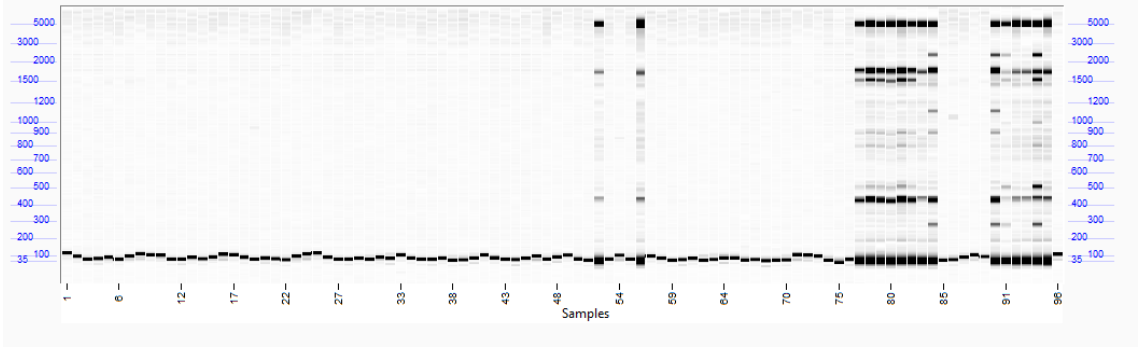
D-10860

X016 F2	S-122995	A1-A4
X016 BC1	S-122972	B4-B7
X015 F2	S-122991	C7-H7
X016 F2	S-122995	A8-C9
X016 BC1	S-122972	D9-D10
Q254* x HT1 F2	S-124776	E10-C11
X051 BC1	S-122973	D11-B12
X009 F2	S-122981	C12-H12

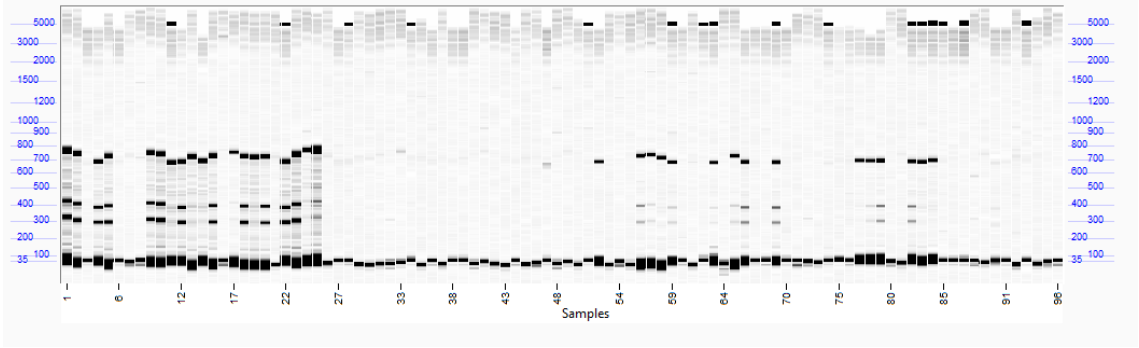
D-10860 Mutation screen



D-10860 CODMb TILL



D-10860 T6ODMa TILL



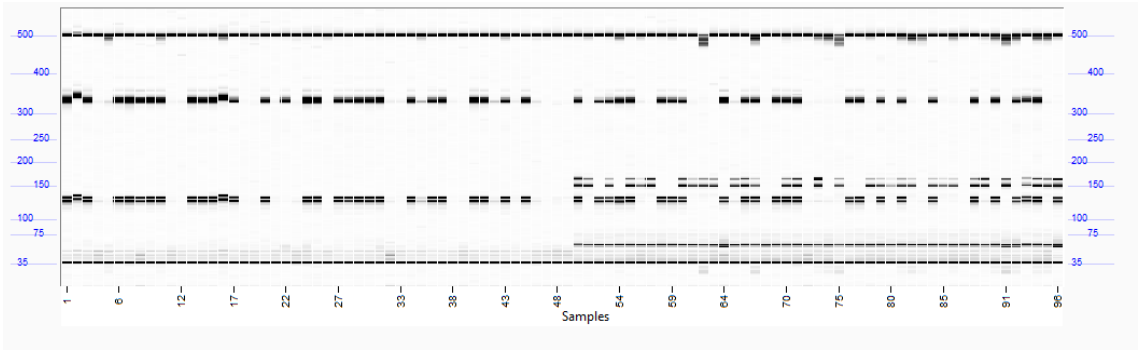
D-10860

	1	2	3	4	5	6	7	8	9	10	11	12
A	Tint het	Tint het	Tint homo	Tint homo				Tint het	Tint het		Q254* het F271L het Tint TILL rxn fail	
B	Tint het	Tint het	Tint het					Tint homo	Tint het		Q254* het F271L het Tint het	E193K R158K het F271L het Tint TILL rxn fail
C		Tint homo	Tint homo					Tint het			**Q254* homo** **F271L het** **Tint homo**	**E193K het** **Q254* het** **R158K het** **F271L no het signal**
D	Tint het	Tint het	Tint het				Q254*? F271L het Tint homo				**E193K het** **R158K het** F271L het **Tint homo**	**E193K wt** **Q254* homo** F271L het
E	Tint het	Tint homo							**Tint het**	**Q254* het** F271L het Tint homo		R158K homo? F271L het Q254* homo?
F		Tint homo	Tint het							Q254* het F271L het Tint het		E193K R158K het Q254* het F271L het
G		Tint het	Tint het					Tint het		**Q254* het** F271L het Tint het		Q254* homo F271L het
H	Tint TILL rxn fail	Tint TILL rxn fail	Tint homo			**Tint het**	Q254*? F271L het Tint het	Tint TILL rxn fail		Q254* het F271L het		

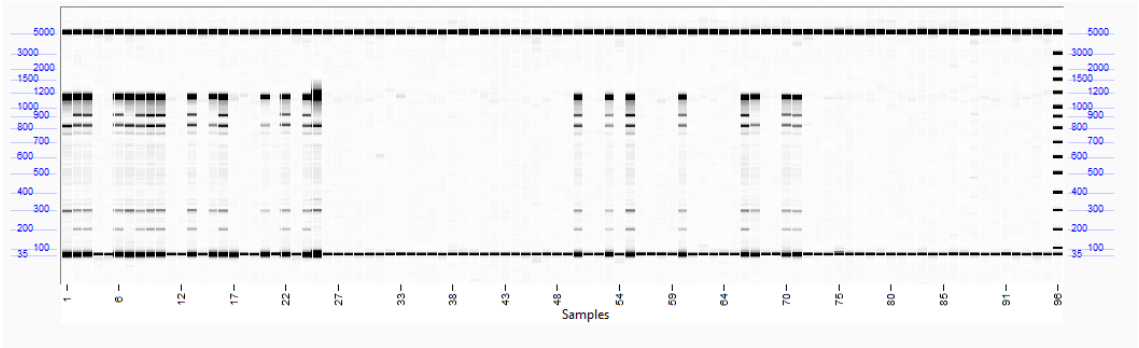
D-10861

X027 F2	S-122982	A1-A4
X027 BC1	S-122978	B4-A7
X030 F2	S-122987	B7-B10
X030 BC1	S-122977	C10-H12

D-10861 Mutation Screen



D-10861 CODMb TILL



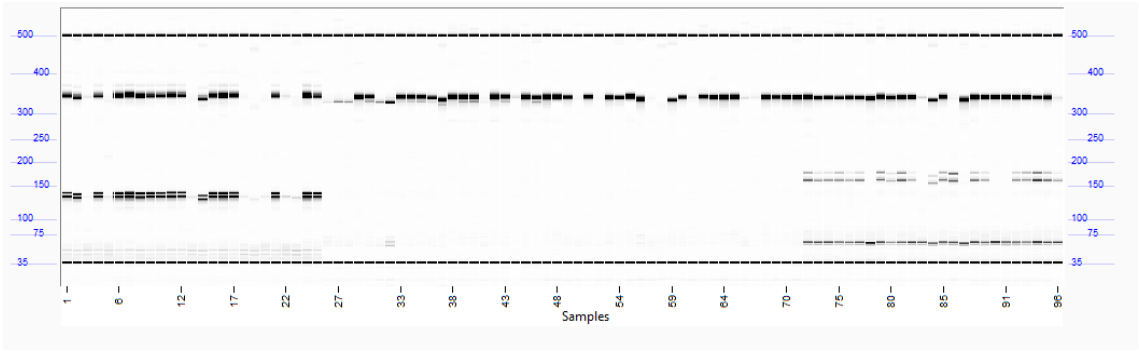
D-10861

	1	2	3	4	5	6	7	8	9	10	11	12
A	**E193K homo** R158K homo F271L het	E193K R158K het F271L het	E193K R158K rxn fail	**E193K homo** **R158K homo** F271L het		E193K R158K		Tint	Tint	Tint	E193K R158K Tint	Tint
B	E193K R158K het F271L het	E193K R158K het F271L het			E193K R158K		E193K R158K het F271L het Tint	E193K R158K	E193K R158K het F271L het Tint		Tint	E193K R158K
C	E193K R158K het F271L het			E193K R158K	E193K R158K	E193K R158K	Tint	E193K R158K	E193K **R158K homo** F271L het **Tint het**	Tint		Tint
D			**E193K homo** **R158K homo** F271L het	E193K R158K	E193K R158K		E193K R158K	E193K R158K het F271L het Tint		E193K R158K	E193K R158K Tint	E193K R158K
E		E193K R158K het F271L het		E193K R158K	E193K R158K	E193K R158K	E193K R158K het F271L het Tint	Tint	E193K R158K	E193K R158K	Tint	E193K R158K Tint
F	E193K R158K het F271L het	E193K R158K rxn fail	R158K het F271L het	E193K R158K			E193K R158K	Tint	E193K R158K het F271L het Tint	Tint	Tint	E193K R158K Tint
G	**E193K homo** **R158K homo** F271L het	E193K **R158K homo** F271L het		E193K R158K			E193K R158K het F271L het Tint	Tint	E193K R158K het F271L het Tint	E193K R158K Tint		Tint
H	E193K R158K het F271L het	E193K R158K het F271L het	E193K R158K het F271L het		E193K R158K		Tint	E193K R158K		Tint	E193K R158K Tint	E193K Tint

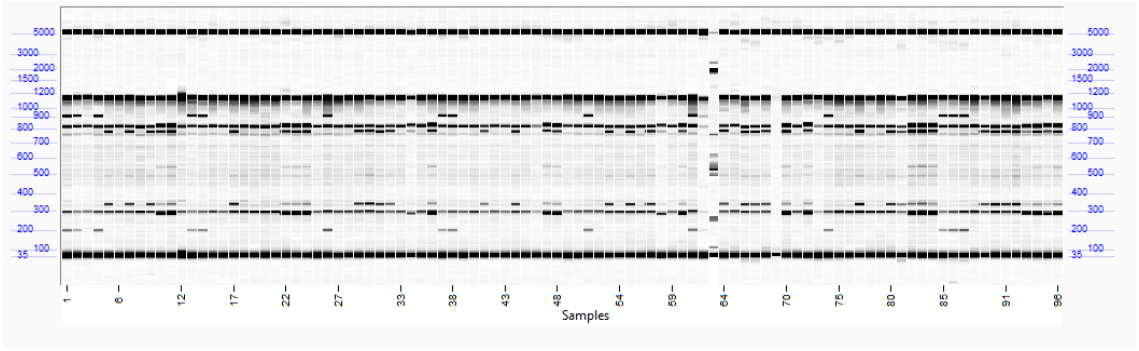
D-10862

X001 F2	S-122996	A1-A4
Q254* x W261* F3	S-124758	B4-F6
Q254* x HN1 F2	S-124770	G6-G9
Q254* x HT1 F2	S-124776	H9-H12

D-10862 Mutation screen



D-10862 CODMb TILL



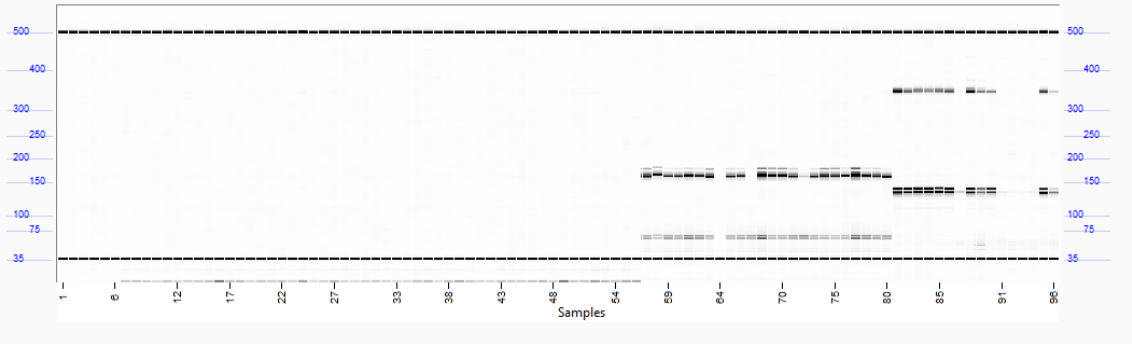
D-10862

	1	2	3	4	5	6	7	8	9	10	11	12
A	E193K R158K het F271L het	E193K R158K het F271L het	**E193K homo** **R158K homo** F271L het	E193K R158K het F271L het	W261*? Q254* het F271L het	F271L het	Q254* het F271L het	F271L het	Q254* het F271L het	Q254* het F271L het Tint	Q254* het F271L het Tint	**Q254* homo** F271L het **Tint het**
B	E193K R158K het F271L het	E193K R158K het F271L het	F271L het	W261* F271L het	W261*? Q254* het F271L het	Q254* homo F271L het	F271L het	F271L het	F271L het	Q254* het F271L het Tint	Q254* het F271L het Tint	Q254* het F271L het
C	F271L het	E193K R158K het F271L het	F271L het	W261* F271L het	W261*? Q254* het F271L het	W261*? Q254* het F271L het	Q254* het F271L het	Q254* het F271L het	F271L het	**Q254* het** F271L het **Tint homo**	F271L het	Q254* het F271L het
D	E193K R158K het F271L het	E193K R158K het F271L het	F271L het	W261* F271L het	Q254* het F271L het	F271L het	F271L het	Q254* het F271L het	Q254* homo F271L het	**Q254* homo** F271L het **Tint het**	Q254* het F271L het Tint	**Q254* het** F271L het **Tint het**
E	F271L het	F271L het	E193K R158K het F271L het	Q254* homo F271L het	**W261* het** **Q254* het** F271L het	Q254* homo F271L het	Q254* het F271L het	F271L het	Q254* het F271L het	**Q254* het** F271L het **Tint het**	**Q254* homo** F271L het **Tint homo**	Q254* het F271L het Tint
F	E193K R158K het F271L het	E193K R158K het F271L het	Rxn failure	**W261* het** **Q254* het** F271L het	Q254* homo F271L het	**W261* het** **Q254* het** F271L het	Q254* het F271L het	Q254* het F271L het	Rxn failure	Q254* het F271L het	F271L het	**Q254* het** F271L het Tint
G	**E193K homo** **R158K homo** F271L het	E193K R158K het F271L het	F271L het	W261* F271L het	W261* Q254* het F271L het	**Q254* het** F271L het	**Q254* homo** F271L het	Q254* het F271L het	Q254* het F271L het	Q254* het F271L het Tint	Q254* het F271L het	**Q254* het** F271L het Tint
H	E193K R158K het F271L het	E193K R158K het F271L het	E193K R158K het F271L het	W261* F271L het	W261* Q254* het F271L het	Q254* het F271L het	Q254* het F271L het	Q254* het F271L het	Q254* het F271L het Tint	Q254* het F271L het Tint	Q254* het F271L het Tint	Q254* het F271L het Tint

D-10865

Q141* M4	S-122951	A1-F1
Q254* x W261* F3	S-124758	G1-H7
X042 F2	S-122998	A8-H10
X001 F2	S-122996	A11-H12

D-10865 Mutation screen



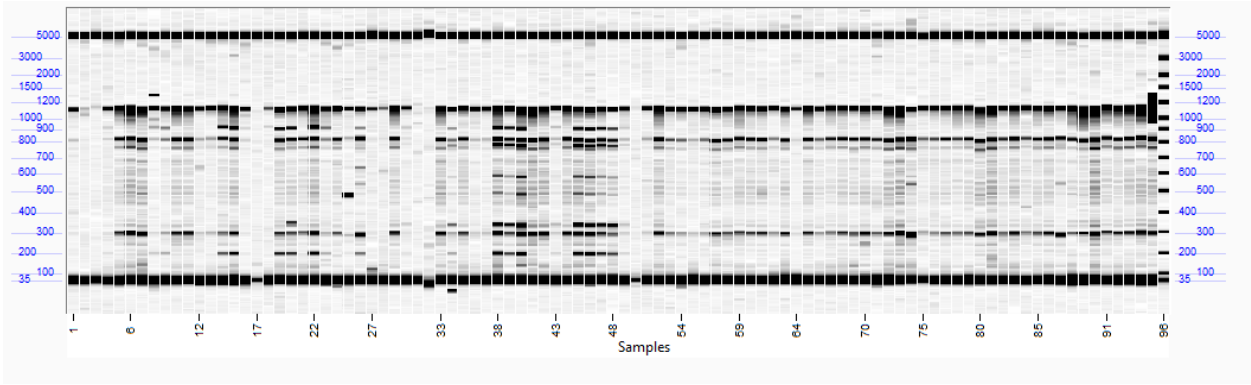
D-10865

	1	2	3	4	5	6	7	8	9	10	11	12
A								Tint	Tint	Tint	E193K R158K	E193K **R158K het**
B								Tint	Tint	Tint	E193K R158K	E193K **R158K het**
C								Tint	**Tint wt**	**Tint homo**	E193K R158K	
D								Tint	Tint	Tint	E193K **R158K homo**	
E								**Tint het**	Tint	Tint	E193K **R158K homo**	
F								Tint	Tint	Tint	E193K **R158K het**	
G								Tint	**Tint het**	Tint		E193K **R158K het**
H								**Tint wt**		Tint	E193K R158K	E193K R158K

D-10868

X027 BC1self	S-186656	A1-G3
X051 BC1self	S-186602	H3-B4
X009 F3	S-186609	C4-C7
X059 F2	S-186709	D7
X071 F2	S-186702	E7-B8
HM6		C8-A10
HT5		B10-H11
X031 F2	S-186693	A12-H12

D-10868 CODMb TILL



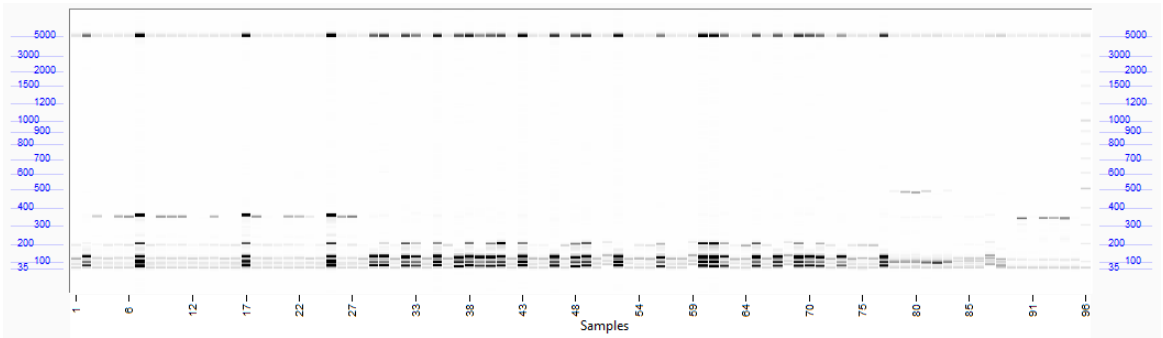
D-10868

	1	2	3	4	5	6	7	8	9	10	11	12
A	**R158K wt**	R158K het	PCR fail	**R158K het**	Q254* wt or homo R158K wt or homo F271L	Q254* wt or homo R158K wt or homo F271L	Q254* wt or homo R158K wt or homo	**W145* WT**				
B	PCR fail	F271L **R158K homo**	R158K pos het seq rxn failed	R158K het	Q254* het R158K het	Q254* wt or homo R158K wt or homo F271L	PCR fail	**W145* het**				
C	PCR fail	F271L **R158K homo**	R158K het	Q254* wt or homo R158K wt or homo	Q254* wt or homo R158K wt or homo F271L	Q254* wt or homo R158K wt or homo	Q254* wt or homo R158K wt or homo					
D	**R158K het**	**R158K homo**	R158K het	Q254* wt or homo R158K wt or homo	Q254* wt or homo R158K wt or homo	Q254* wt or homo R158K wt or homo F271L	**Q254* homo** **Q141* WT**					
E	F271L R158K wt or homo	**R158K wt**	R158K wt or homo	Q254* wt or homo R158K wt or homo F271L	Q254* wt or homo R158K wt or homo	Q254* het R158K het F271L	**W145* homo**					
F	F271L R158K wt or homo	R158K het	R158K het	Q254* wt or homo R158K wt or homo	Q254* het R158K het F271L	Q254* het R158K het F271L	**W145* WT**					
G	F271L R158K wt or homo	R158K het	R158K het	PCR fail	Q254* het R158K het F271L	Q254* het R158K het F271L	**W145* het**					
H	**R158K het**	**R158K het**	**R158K wt**	PCR fail	Q254* het R158K het F271L	Q254* het R158K het F271L	**W145* homo**					

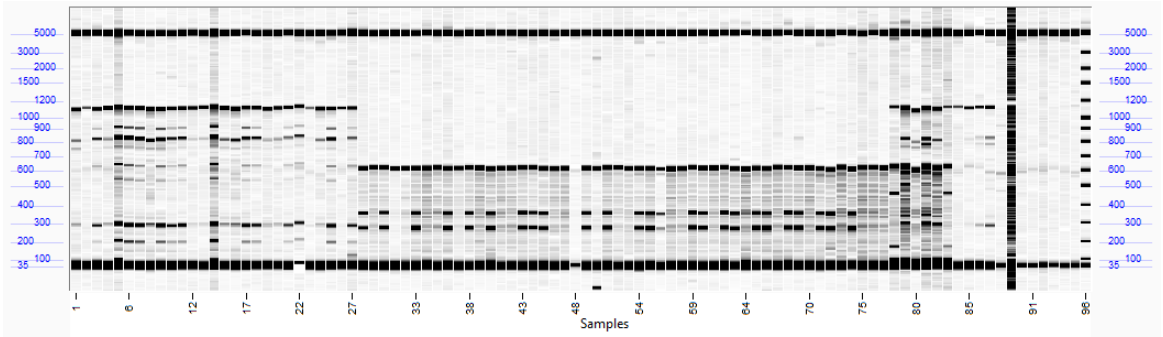
D-10869

X041 F3	S-186742	A1-B1
X051 BC1 self	S-186603	C1-C4
X016 BC1 self	S-186625	D4-D7
X016 BC1 self	S-186626	E7-E10
X059 F2	S-186710	F10-C11
Q254* x HT1 F3	S-186633	D11-H11
X026 F2	S-186689	A12-G12

D-10869 Mutation_screen



D-10869 CODMb and T6ODMa TILL



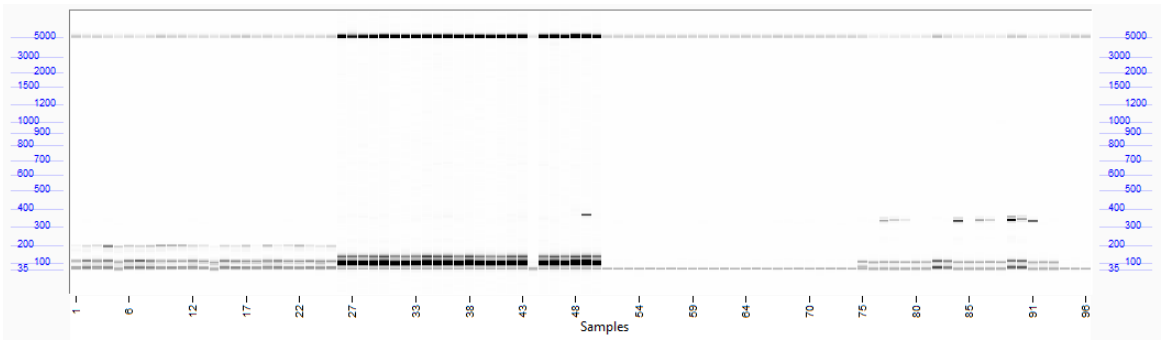
D-10869

	1	2	3	4	5	6	7	8	9	10	11	12
A	F271L het Tint	R158K het Tint	R158K het Tint	R158K het Tint	Tint het	Tint homo	Tint het	wt	Tint het	wt	**Q254* het** **Q141* homo**	
B	F271L TBD Tint	R158K het Tint	R158K het Tint	**R158K homo** Tint	wt	wt	wt	wt	Tint het	Tint het	Q254* het wt	**W261* het**
C	R158K het Tint	R158K het Tint	Tint	R158K het Tint	Tint homo	Tint het	Tint het	Tint het	wt	Tint homo	wt Q141* het	
D	Tint	Tint	Tint	Tint het	Tint het	Tint het	wt	Tint het	Tint het	Tint homo	wt	**W261* het**
E	R158K het Tint	Tint	R158K het Tint	wt	wt	Tint het	wt	Tint homo	Tint het	wt	**Q254* homo** Tint	**W261* het**
F	R158K het Tint	R158K het Tint	R158K het Tint	Tint het	Tint het	wt	Tint het	Tint het	wt	wt Q141* het	Q254* het Tint	**W261* het**
G	R158K het Tint	Tint	**R158K het**	Tint homo	wt	Tint homo	Tint het	Tint het	Tint het	**Q254* homo** **Q141* homo**	Q254* het Tint	
H	Tint	Tint	Tint	Tint homo	Tint het	Tint TBD	Tint het	Tint pos homo	Tint het	**Q254* het** **Q141* homo**	wt Tint	

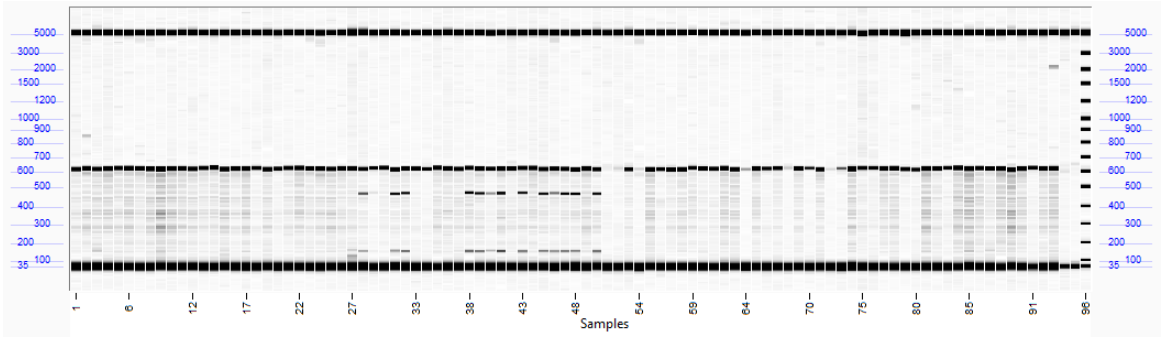
D-10870

Q254* x HT1 F3	S-186680	A1-A4
X071 F3	S-186703	B4-B7
X029 F2	S-186686	C7-B10
Q254* x HN1 F3	S-186635	C10-E12

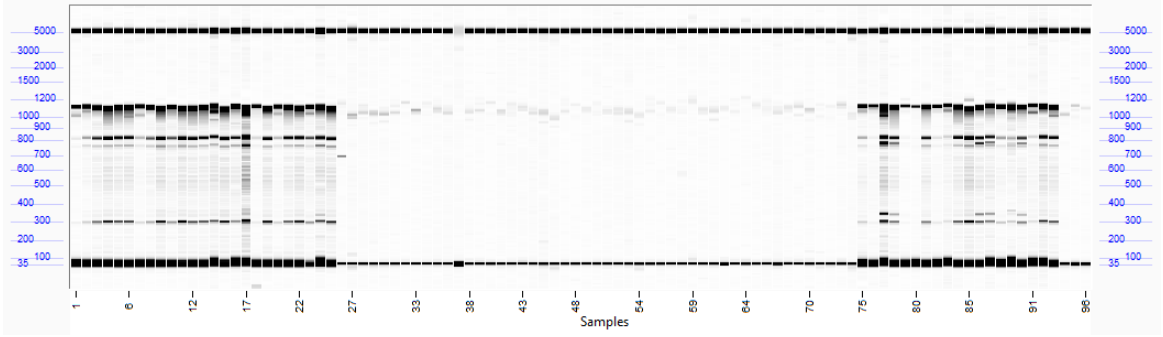
D-10870 Mutation screen



D-10870 T6ODMa TILL



D-10870 CODMb TILL_FM24OCT14

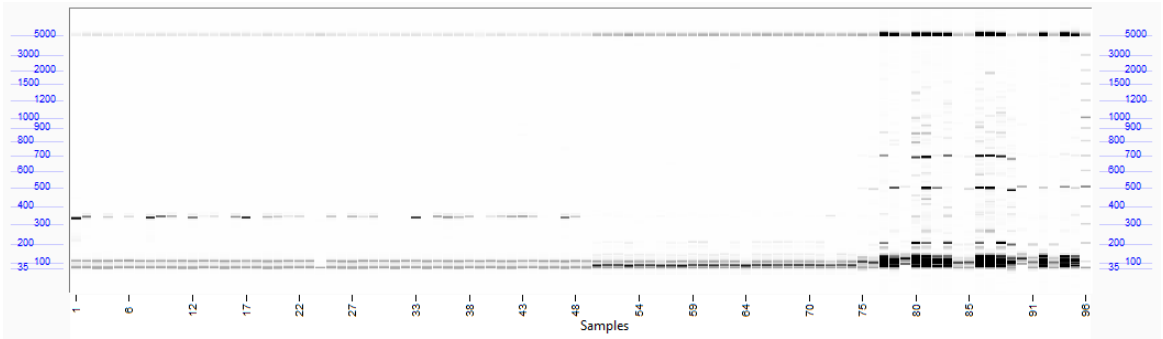


D-10870

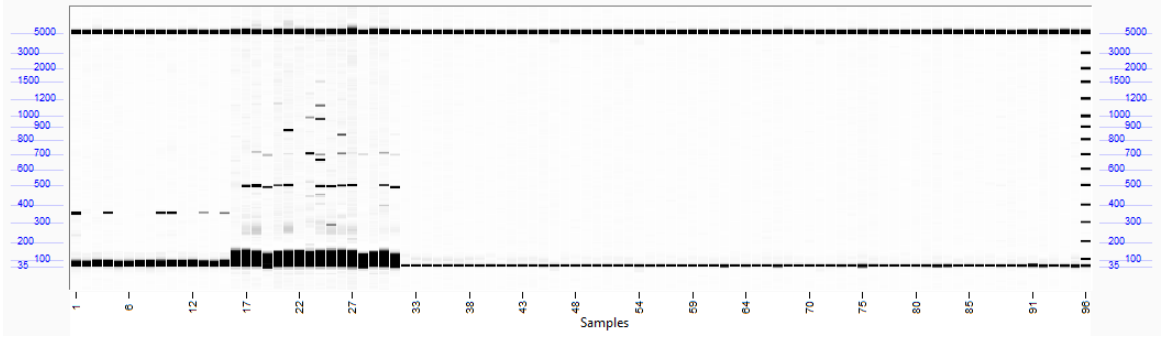
	1	2	3	4	5	6	7	8	9	10	11	12
A		Tint	Tint	Tint	wt	W145* het	**W145* homo**				wt	**Q254* het**
B	Tint	Tint		wt	wt	wt	W145* het				wt	Q254* het
C	Tint	Tint	Tint	wt	wt	W145* het				wt	wt	**Q254* homo**
D	Tint	Tint	Tint	W145* het	wt	wt				wt	**Q254* homo**	wt
E	Tint	Tint	Tint	wt	wt	W145* het				Q254* het	wt	wt
F	Tint		Tint	wt	W145* het	W145* het				Q254* het	Q254* het	
G	Tint	Tint	Tint	W145* het	W145* het	W145* het				**Q254* het**	Q254* het	
H	Tint	Tint	Tint	W145* het	W145* het	W145* het				wt	wt	

D-10871	***&***		***D-10872***		
Q254* x HN1 F3	S-186634	A1-A4	X001 F3	S-186724	A1
X008 F3	S-186676	B4-A7	X001 F3	S-186725	B1-G2
X065 F2	S-186696	B7-B10	X074 F2	S-186704	H2-G4
X069 F2	S-186698	C10-G12			

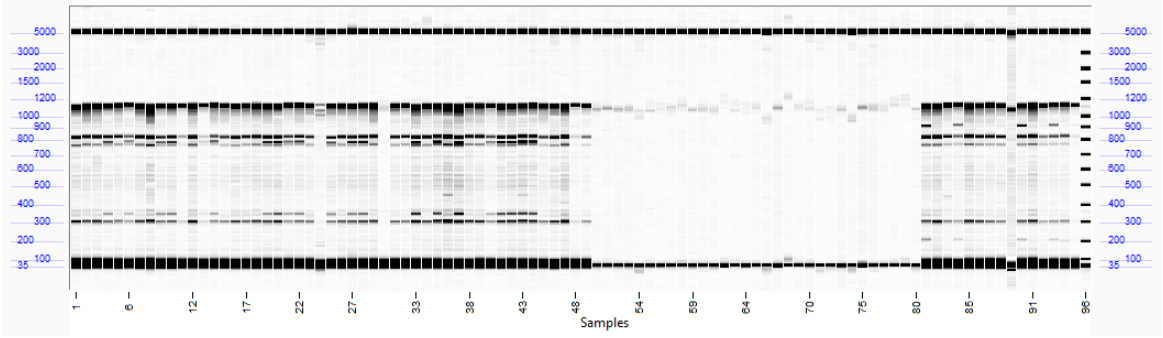
D-10871 Mutation screen



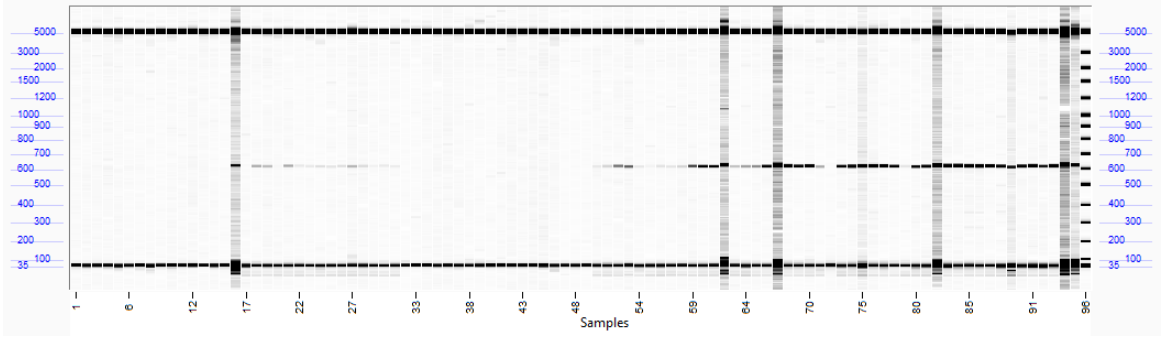
D-10872 Mutation screen



D-10871(Cols1-7)+D-10872 (Cols11-12) CODMb TILL



D-10871 (Cols7-12)+D-10872 (Cols2-4) T6ODMa TILL



D-10871

	1	2	3	4	5	6	7	8	9	10	11	12
A	**Q254* homo**	Q254* het	Q254* homo	Q254* het	Q254* het	Q254* het	wt	Tint	Tint		Q141* Tint	Q141* Tint
B	**Q254* homo**	Q254* het	wt	wt	wt	Q254* het		Tint	Tint		**Q141* homo**	**Q141* homo**
C	wt	wt	Q254* het	Q254* het	Q254* het	Q254* het	Tint	Tint	Tint	pos Q141*	Tint	Tint
D	Q254* het	Q254* homo	Q254* het	Q254* het	Q254* pos het	Q254* het	Tint	Tint	Tint	**Q141* het**		**Q141* homo**
E	wt	wt	Q254* het	**Q254* homo**	Q254* het	wt	Tint		Tint	Tint	**Q141* het**	Tint
F	Q254* het	Q254* het	Q254* het	wt	**Q254* homo**	wt			Tint	**Q141* homo**	Q141* Tint	Q141*
G	**Q254* WT**	wt	wt	wt	wt	**Q254* homo**		Tint	Tint	pos Q141*	Q141*	Q141*
H	**Q254* homo**	Q254* homo	wt	wt	Q254* het	**Q254* homo**	**W145* het**		**W145* het**	Tint	Tint	

D-10872

	1	2	3	4	5	6	7	8	9	10	11	12
A	R158K het	**R158K het**	Q141* het	**Q141* het**								
B	wt	R158K het	**Q141* homo**	**Q141* het**								
C	wt	wt	**Q141* homo**	**Q141* het**								
D	R158K het	wt	**Q141* homo**									
E	wt	R158K het	Q141*									
F	wt	wt		Q141*								
G	wt	R158K TBD		Q141*								
H	wt		Q141*									

Appendix H- Genotyping of material grown in Tasmania

Field plots each seed batch was sown in

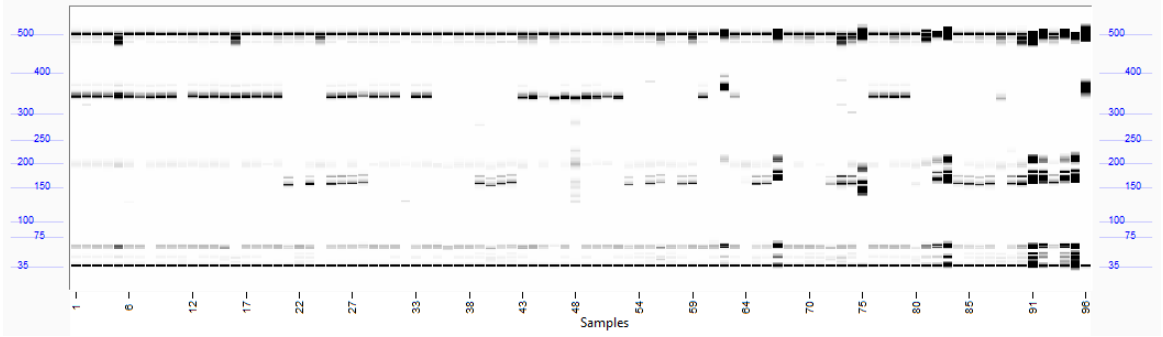
Seed batch id	Plot
S-186660	MD 801
HT6	MD 802
S-186608	MD 803
S-186658	MD 804
HM6	MD 805
S-186610	MD 806
S-186609	MD 807
S-186720	MD 808
S-186729	MD 809
S-186620	MD 810
S-186622	MD 811
HT5	MD 812
S-186728	MD 813
S-186667	MD 814
S-186716	MD 815
S-186742	MD 816
S-186661	MD 817
S-186609	MD 818
HM7	MD 819
S-186617	MD 820
S-186611	MD 821
S-186601	MD 822
S-186660	MD 823
S-186613	MD 824
HM6	MD 825
S-186671	MD 826
S-186608	MD 827
S-186728	MD 828
S-186619	MD 829
S-186622	MD 830
HT6	MD 831
S-186611	MD 832
S-186609	MD 833
S-186659	MD 834
S-186610	MD 835
S-186613	MD 836
S-186640	MD 837
S-186719	MD 838
HM6	MD 839
S-186608	MD 840
S-186617	MD 841
S-186667	MD 842
S-186671	MD 843
S-186620	MD 844
HT5	MD 845
S-186716	MD 846
S-186605	MD 847
S-186729	MD 848
S-186742	MD 849
S-186601	MD 850
S-186608	MD 851
S-186659	MD 852
HM7	MD 853
S-186662	MD 854
S-186637	MD 855
S-186661	MD 856
S-186609	MD 857
S-186619	MD 858
S-186678	MD 859
S-186621	MD 860
S-186605	MD 861
S-186658	MD 862
S-186716	MD 863
S-186624	MD 864
S-186637	MD 865
S-186740	MD 866
S-186613	MD 867
S-186678	MD 868
T 32	MD 869
S-186607	MD 870

Seed batch id	Plot
S-186682	MD 871
S-186685	MD 872
S-186716	MD 873
S-186628	MD 874
S-186720	MD 875
S-186632	MD 876
S-186717	MD 877
S-186640	MD 878
S-186613	MD 879
HN1	MD 880
S-186628	MD 881
S-186607	MD 882
S-186624	MD 883
S-186719	MD 884
S-186685	MD 885
S-186665	MD 886
S-186682	MD 887
S-186717	MD 888
HN1	MD 889
S-186662	MD 890
S-186632	MD 891
S-186664	MD 892

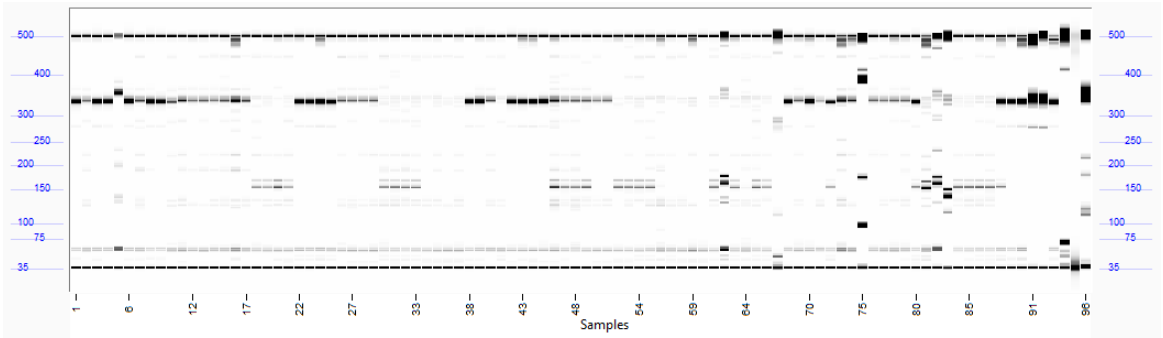
Genotyping of six DNA plates representing field grown material

Plants were screened for the Q254* (339 bp), R158K (345 bp), and Tint (164 bp) SNPs where appropriate. CODMb heterozygotes indicated by predicted fragment sizes after treatment with cleavage enzyme i.e. R158K (919 bp+203 bp), F271L (843 bp+278 bp) and Q254* (789 bp+332 bp). Likewise, T6ODMa heterozygotes were indicated Tint (359 bp+267 bp).

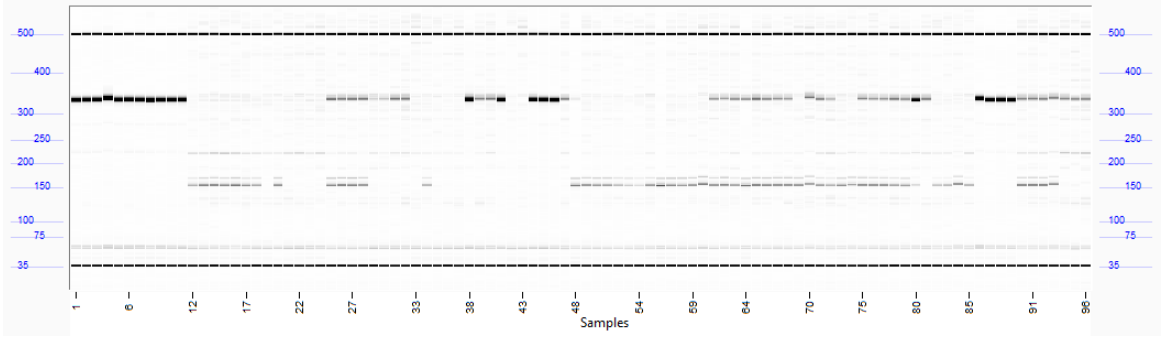
D-11113 Mutation Screen



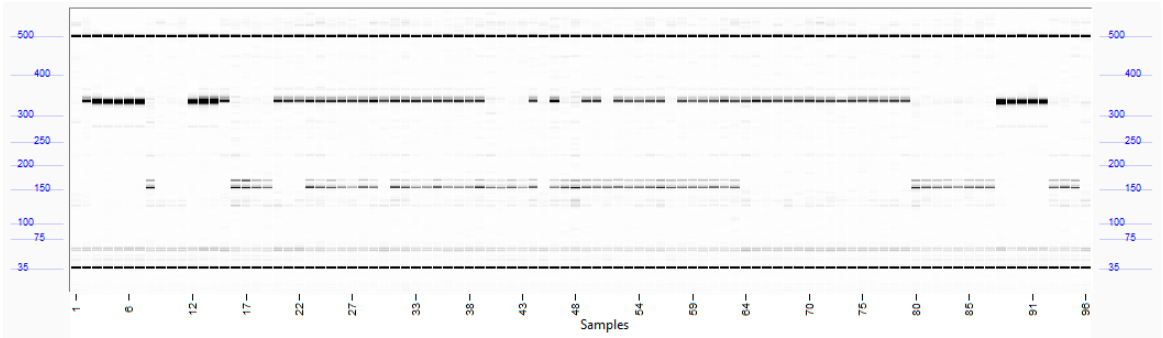
D-11114 Mutation Screen



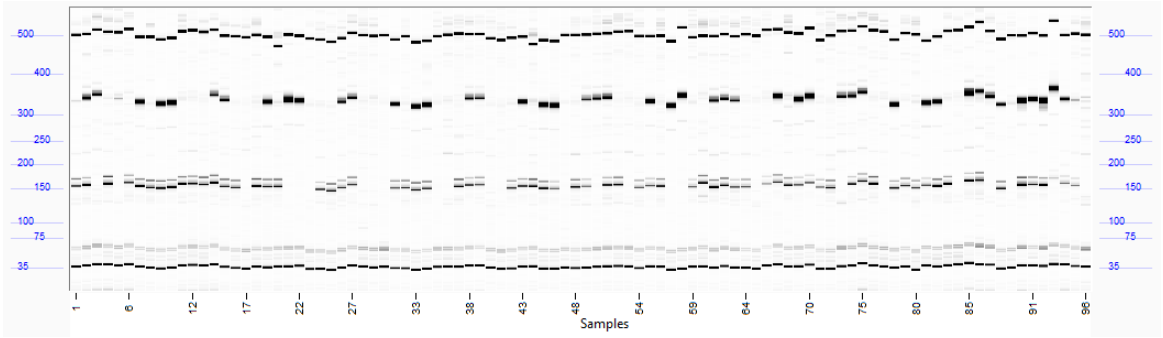
D-11115 Mutation Screen



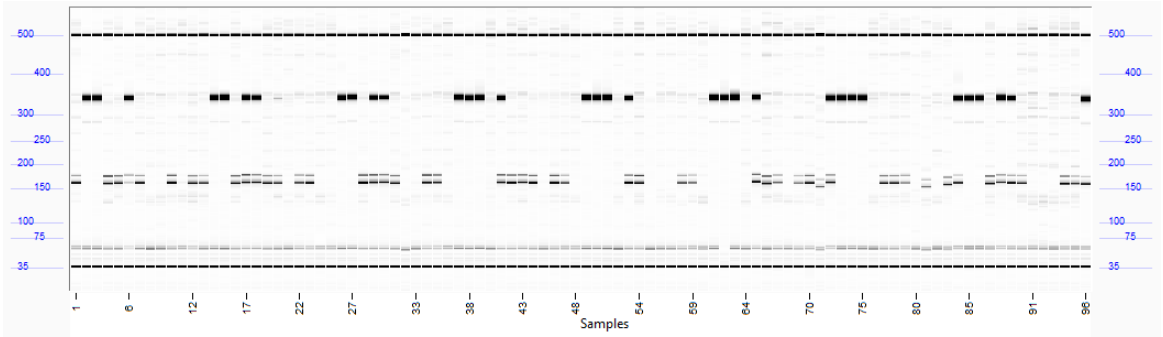
D-11116 Mutation Screen



D-11117 Mutation Screen



D-11118 Mutation Screen



TILLING plate layouts (to determine zygosity in individuals where SNPs were segregating)

T6ODMa TILL Plate 1

	1	2	3	4	5	6	7	8	9	10	11	12
A					Tint		Tint		R158K Tint	R158K Tint		
B		Tint		Tint	Tint		Tint		R158K Tint	R158K Tint		
C		Tint		Tint	Tint				R158K Tint	R158K Tint		
D				Tint	Tint	Tint	Tint	R158K	R158K Tint	R158K Tint		
E	Tint			Tint	Tint	Tint		R158K	R158K Tint	R158K Tint	Tint	
F				Tint	Tint	Tint		R158K	R158K	R158K Tint	Tint	
G	Tint			Tint	Tint	Tint		R158K Tint	R158K Tint	R158K Tint	Tint	
H	Tint					Tint		R158K Tint	R158K Tint			
	7	8	10	11	12	2	3	3	4	5	12	
	D-11113					D-11115		D-11116				

T6ODMa TILL Plate 2

	1	2	3	4	5	6	7	8	9	10	11	12
A		R158K Tint	R158K		Tint	Q254* Tint	Q254* Tint	Tint	Tint	Tint		
B		R158K Tint	R158K Tint		Tint	Q254*	Q254*		Tint	Tint		Tint
C		R158K Tint	R158K Tint		Tint	Q254* Tint	Q254* Tint					
D		R158K Tint	R158K Tint		Tint	Q254* Tint	Q254* Tint		Tint		Tint	Tint
E	R158K Tint		R158K Tint		Tint	Q254*	Q254*	Tint	Tint			
F	R158K Tint	R158K Tint	R158K Tint		Tint	Q254* Tint	Q254* Tint	Tint	Tint		Tint	
G	R158K	R158K Tint	R158K Tint	Tint	Tint	Q254* Tint	Q254* Tint	Tint			Tint	
H	R158K Tint	R158K Tint	R158K	Q254*? Tint	Q254* Tint	Q254* Tint	Q254? Tint	Tint				
	1	2	3	7	8	9	10	11	12	1	7	8
	D-11117									D-11118		

T6ODMa TILL Plate 3

	1	2	3	4	5	6	7	8	9	10	11	12
A		Tint		Tint	Q254* Tint	Q254* Tint	Q254*					
B		Tint	Tint		Q254* Tint	Q254* Tint						
C		Tint	Tint		Q254* Tint	Q254* Tint						
D		Tint	Tint		Q254* Tint	Q254* Tint						
E		Tint	Tint		Q254* Tint	Q254* Tint						
F	Tint	Tint	Tint	Q254* Tint	Q254*	Q254* Tint						
G	Tint		Tint	Q254* Tint	Q254*	Q254* Tint						
H		Tint	Tint	Q254* Tint								

9	10	11	12	1	2	3
D-11118				D-11119		

CODMb TILL Plate 1

	1	2	3	4	5	6	7	8	9	10	11	12
A		R158K		Q254*	Q254*			Q254*	Q254*	Q254*		Q254*
B		R158K		R158K		Q254*		Q254*	Q254*	Q254*		
C	R158K	R158K		Q254*		Q254*		Q254*	Q254*	Q254*		
D	R158K	R158K	R158K	Q254*		Q254*	Q254*	Q254*	Q254*			Q254*
E	R158K			Q254*		Q254*	R158K	Q254*	Q254*			Q254*
F	R158K		R158K	Q254*	Q254*		Q254*	Q254*	Q254*		Q254*	Q254*
G	R158K		R158K	R158K	Q254*		R158K	Q254*	Q254*		R158K	R158K
H	R158K			Q254*	R158K		Q254*	Q254*	Q254*		R158K	

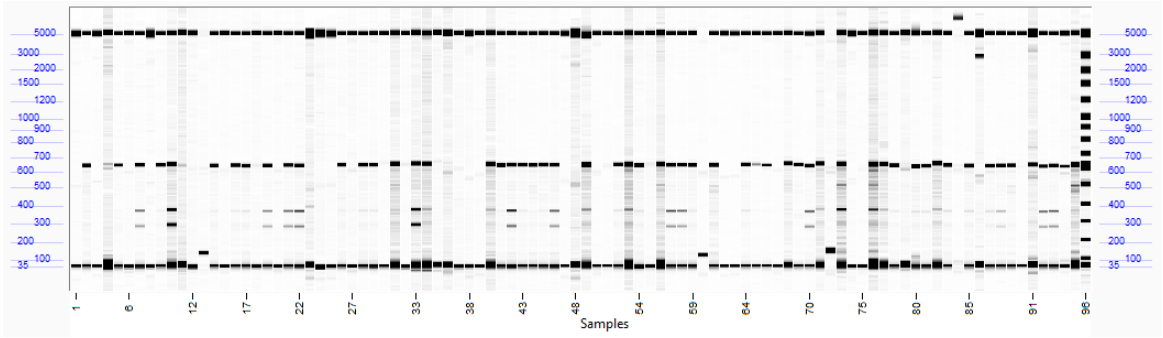
6	7	8	1	2+5	6	9+11	12	1	2	5	6
D-11113			D-11114					D-11115			

CODMb TILL Plate 2

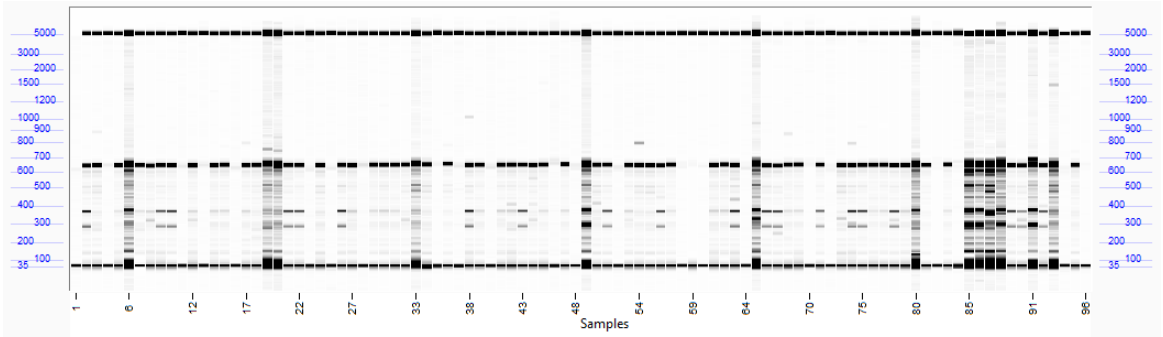
	1	2	3	4	5	6	7	8	9	10	11	12
A		R158K Tint	Tint			R158K Tint	Tint		Q254* Tint		Tint	Q254* Tint
B		R158K Tint	Tint			R158K Tint	R158K Tint		Q254* Tint		Q254* Tint	Q254* Tint
C		R158K Tint	R158K Tint			Tint	R158K Tint		Q254* Tint		Q254* Tint	Q254* Tint
D	Tint	R158K Tint	R158K Tint	Q254*	R158K Tint	R158K Tint	R158K Tint		Q254* Tint		Q254* Tint	
E	R158K Tint	Tint	R158K Tint	Q254*		R158K Tint	R158K Tint	Tint	Q254* Tint		Q254* Tint	
F	R158K Tint	R158K Tint	R158K Tint	Q254*	R158K Tint	R158K Tint	R158K Tint	Tint	Q254* Tint		Q254* Tint	
G	R158K Tint	R158K Tint	R158K Tint	R158K	Tint	R158K Tint	R158K Tint	Q254* Tint			Tint	
H	R158K Tint	R158K Tint			Tint	R158K Tint		Q254* Tint		Q254* Tint	Q254* Tint	

8	9	10	2	6	7	8	6	7	4	5	6
D-11115			D-11116				D-11117		D-11118		

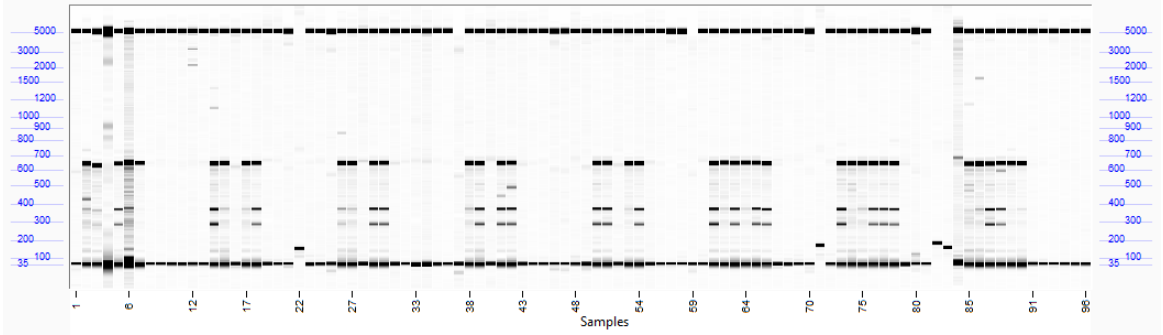
T6ODMa TILL Plate1



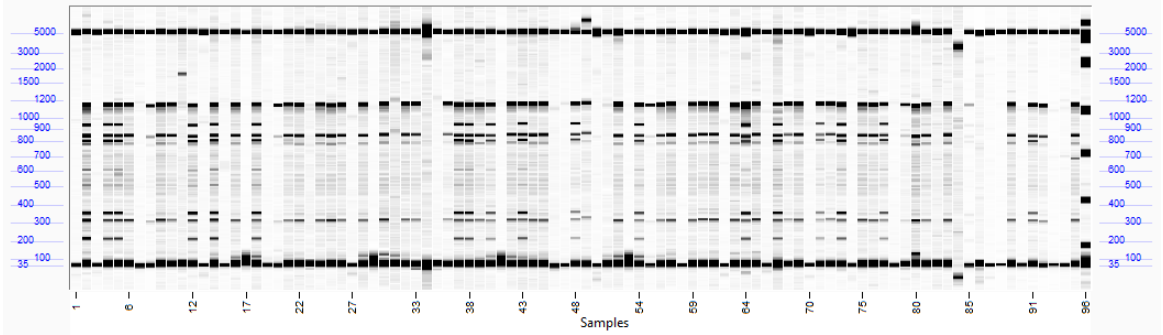
T6ODMa TILL Plate2



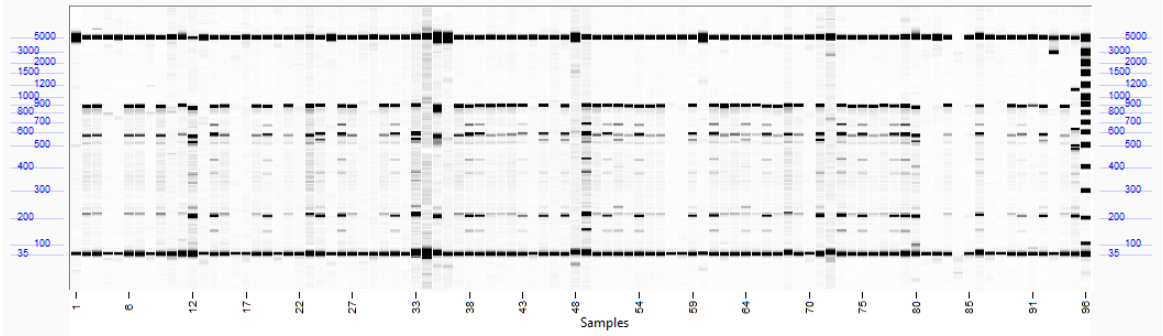
T6ODMa TILL Plate3



CODMb TILL Plate1



CODMb TILL Plate2



Mutations identified with AS-PCR

D-11113	1	2	3	4	5	6	7	8	9	10	11	12
A	R158K	R158K	R158K	R158K Tint	R158K	Tint	Q254* het R158K het	WT T6ODMa	Tint	Tint	WT	Tint homo
B	R158K	R158K	R158K	R158K Tint	R158K	Tint	Q254* het R158K het	Tint homo	Tint	Tint	Tint homo	Tint homo
C	R158K		R158K	R158K Tint		R158K homo	R158K homo	Tint homo	Tint	Tint	Tint homo	Tint homo
D	R158K	R158K	R158K	R158K Tint		Q254* het R158K het	Q254* het R158K het	R158K homo		R158K	Tint homo	Tint homo
E	R158K	R158K	Tint	R158K		R158K homo	Tint homo			R158K	Tint homo	Tint homo
F	R158K	R158K		R158K		R158K homo	WT T6ODMa	R158K homo		R158K	Tint homo	Tint homo
G	R158K	R158K	Tint	R158K	Tint	Q254* het R158K het	Tint homo	R158K homo		R158K	Tint homo	Tint homo
H	R158K	R158K			Tint	R158K zy_nd	Tint homo		Tint	Tint homo	Tint homo	

S-186660 X027 F3	pred all E193K/R158K homo
HT6 control	
S-186608 X051 F3	pred all E193/R158K/Tint homo
S-186658 X027 F3	pred all E193K/R158K homo
HM6 control	
S-186610 X042 F3	pred all Tint homo
S-186609 X009 F3	Q254* and E193K/R158K mutations to segregate
S-186720 X042 F3	Tint segregating
S-186729 X001 F3	pred all E193K/R158K homo
S-186620 X016 F3	pred all Tint homo
S-186622 X016 F3	pred all Tint homo
HT5 control	
S-186728 X001 F3	pred all E193K/R158K homo
S-186667 X041 F3	pred all Tint homo
S-186716 X042 F3	pred all Tint homo
S-186742 X041 F3	pred all Tint homo
S-186661 X041 F3	pred all Tint homo
S-186609 X009 F3	Q254* and E193K/R158K mutations to segregate

Mutations identified with AS-PCR

D-11114	1	2	3	4	5	6	7	8	9	10	11	12
A	R158K het Q254* het	R158K het Q254* het	R158K	Q254*	Tint		R158K Tint		Tint	R158K	Tint	Q254* homo
B	R158K homo	R158K	Tint	R158K		R158K het Q254* het	R158K		Tint	R158K	Tint	Q254* homo
C	Q254* homo	R158K	Tint	R158K		Q254* homo	R158K		Tint	R158K	Tint	Q254* homo
D	R158K het Q254* het	R158K	Tint	R158K		Q254* homo	Tint		R158K het Q254* het	R158K	Tint	Q254* homo
E	Q254* homo	R158K	Tint	R158K		R158K het Q254* het	Tint	Tint	R158K homo	R158K	Tint	Q254* homo
F	R158K het Q254* het	R158K	Q254*	Tint	Q254* homo	R158K Tint	Tint	Tint	R158K het Q254* het	R158K	Tint	Q254* homo
G	R158K homo	R158K	Q254*	Tint	R158K het Q254* het	R158K Tint	Tint	Tint	R158K homo	R158K	Tint	Q254* homo
H	Q254* zy_nd	R158K	Q254*	Tint	R158K homo	R158K Tint			R158K	Q254*	Q254 het	Q254* homo

S-186609 X009 F3	Q254* and E193K/R158K mutations to segregate
HM7 control	
S-186617 X001 F3	pred E193K/R158K homo
S-186611 X042 F3	pred Tint homo
S-186601 X008 F3	pred Q254* homo
S-186660 X027 F3	pred E193K/R158K homo
S-186613 X042 F3	pred Tint homo
HM6 control	
S-186671 X009 F3	Q254* and E193K/R158K mutations to segregate
S-186608 X051 F3	pred E193K/R158K and Tint homo
S-186728 X001 F3	pred E193K/R158K homo
S-186619 X016 F2	pred Tint homo
S-186622 X016 F2	pred Tint homo
HT6 control	
S-186611 X042 F3	pred Tint homo
S-186609 X009 F3	Q254* and E193K/R158K mutations to segregate
S-186659 X027 F3	pred E193K/R158K homo
S-186610 X042 F3	pred Tint homo
S-186613 X042 F3	pred Tint homo
S-186640 Q254* x HN1 F3	pred Q254* mutation segregating

Mutations identified by AS-PCR

D-11115	1	2	3	4	5	6	7	8	9	10	11	12
A	Q254* homo	Q254* homo	Tint het	R158K Tint		R158K het Q254* het	Tint	Tint	**R158K homo Tint**	WT CODMb Tint	R158K	Q254*
B	Q254* homo	Q254* homo	Tint het	R158K Tint	Tint		Tint	Tint	R158K het Tint	WT CODMb Tint	Tint	R158K Tint
C	Q254* homo	Q254* zy_nd	WT T6ODMa	R158K Tint			Tint	Tint	R158K het Tint	**R158K homo Tint**	Tint	R158K Tint
D	Q254* homo	Tint het	Tint homo	R158K Tint		R158K het Q254* het	Tint	WT CODMb Tint	R158K het Tint	R158K het Tint	Tint	R158K Tint
E	Q254* homo	Tint homo		R158K		Q254* homo	Tint	R158K het Tint	WT CODMb Tint	R158K het Tint	Tint	R158K Tint
F	Q254* homo	Tint homo		R158K	R158K het Q254* het	Q254* homo	Tint	R158K het Tint	**R158K homo Tint**	**R158K homo Tint**	Q254*	R158K
G	Q254* homo	Tint homo		R158K	R158K homo	R158K zy_nd	Tint	R158K het Tint	**R158K homo Tint**	R158K het Tint	Q254*	R158K
H	Q254* zy_nd	Tint zy_dn		R158K	R158K homo	Tint	Tint	R158K zy_nd Tint	R158K het Tint	R158K	Q254*	R158K

S-186640 Q254* x HN1 F3	pred Q254* mutation segregating
S-186719 X042 F3	pred Tint segregating
HM6 control	
S-186608 X051 F3	pred homo for E193K/R158K and Tint
S-186617 X001 F3	pred E193K/R158K homo
S-186667 X041 F3	pred Tint homo
S-186671 X009 F3	Q254* and E193K/R158K mutations to segregate
S-186620 X016 F3	pred Tint homo
HT5 control	
S-186716 X042 F3	pred Tint homo
S-186605 X051 F3	pred E193K/R158K segregating, Tint homo
S-186729 X001 F3	pred E193K/R158K homo
S-186742 X041 F3	pred all Tint homo_check seq of F2 plant
S-186601 X008 F3	pred all Q254* homo
S-186608 X051 F3	pred homo for E193K/R158K and Tint
S-186659 X027 F3	pred E193K/R158K homo
HM7 control	

Mutations detected with AS-PCR

D-11116	1	2	3	4	5	6	7	8	9	10	11	12
A			Tint	**R158K Tint homo**	R158K Tint het	Tint	**R158K homo Tint**	WT CODMb Tint	R158K	R158K	Tint	Q254*
B	R158K		Tint	R158K Tint het	R158K Tint het	Tint	**R158K homo Tint**	R158K het Tint	R158K	R158K	Tint	Q254*
C	Q254*		Tint	R158K Tint het	**R158K Tint homo**	Tint	WT CODMb Tint	R158K het Tint	R158K	R158K	Tint	Q254*
D	Q254*	Q254* homo	R158K WT T6ODMa	**R158K Tint homo**	R158K Tint het	**R158K homo Tint**	**R158K homo Tint**	R158K het Tint	R158K	R158K	Tint	Q254*
E	Q254*	R158K het Q254* het	R158K WT T6ODMa	R158K Tint het	R158K Tint het		R158K het Tint	**R158K homo Tint**	R158K	R158K	Tint	Tint homo
F	Q254*	R158K het Q254* het	R158K WT T6ODMa	R158K WT T6ODMa	R158K Tint het	R158K homo Tint	R158K het Tint	**R158K homo Tint**	R158K	R158K	Tint	Tint homo
G	Q254*	R158K homo	**R158K Tint homo**	**R158K Tint homo**	**R158K Tint homo**	WT CODMb Tint	R158K het Tint	R158K het Tint	R158K	R158K	Tint	Tint homo
H	Tint	Tint	R158K Tint het	R158K Tint het	Tint	WT CODMb Tint	R158K het Tint	R158K	R158K	Tint	Q254*	

HM7 control	
S-186637 Q254* x HN1	pred Q254* homo
S-186661 X041 F3	pred Tint homo_check F2 seq
S-186609 X009 F3	Q254* and E193K/R158K mutations to segregate
S-186619 X016 F3	pred Tint homo
S-186678 X051 F3	pred Tint segregating, E193K/R158K homo
S-186621 X016 F3	pred Tint homo
S-186605 X051 F3	pred E193K/R158K segregating, Tint homo
S-186658 X027 F3	pred E193K/R158K homo
S-186716 X042 F3	pred Tint homo
S-186624 X016 F3	pred Tint homo
S-186637 Q254* x HN1	pred Q254* homo
S-186740 X041 F3	pred Tint homo

Mutations detected with AS-PCR

D-11117	1	2	3	4	5	6	7	8	9	10	11	12
A	Tint	R158K Tint het	R158K WT T6ODMa	Tint		Tint	**Q254* homo Tint**	Tint homo	Q254* Tint pos het	**Q254* Tint homo**	Tint homo	Tint het
B	Tint	**R158K Tint homo**	**R158K Tint homo**	Tint		Tint	**Q254* homo Tint**	Tint homo	Q254* WT T6ODMa	Q254* WT T6ODMa	WT T6ODMa	Tint het
C	Tint	R158K Tint het	**R158K Tint homo**				Q254* het Tint	Tint homo	**Q254* Tint homo**	**Q254* Tint homo**	WT T6ODMa	WT T6ODMa
D	Tint	R158K Tint het	**R158K Tint homo**			Tint	Q254* het Tint	Tint homo	**Q254* Tint homo**	Q254* Tint het	WT T6ODMa	Tint homo
E	R158K Tint het	R158K WT T6ODMa	R158K Tint het	Tint		WT CODMb Tint	Q254* zy_nd Tint	Tint homo	Q254* WT T6ODMa	Q254* WT T6ODMa	Tint het	Tint homo
F	**R158K Tint homo**	**R158K Tint homo**	R158K Tint het	Tint		WT CODMb Tint	**Q254* homo Tint**	Tint homo	Q254* Tint het	Q254* Tint het	Tint homo	Tint homo
G	R158K WT T6ODMa	R158K Tint het	R158K Tint het	Tint		Q254* het Tint	Tint homo	Tint homo	Q254* Tint het	**Q254* Tint homo**	Tint homo	WT T6ODMa
H	R158K Tint het	R158K Tint het	R158K unclear	R158K? Tint		Q254* het Tint	Q254*? Tint het	Q254* Tint het	Q254* Tint het	Q254? Tint het	Tint het	WT T6ODMa

S-186613 X042 F3	pred all Tint homo
S-186678 X051 F3	pred Tint segregating, E193K/R158K homo
HT5 control	
S-186607 X014 F3	pred all Tint homo
S-186682 X023 F3	no markers
S-186685 X029 F3	no markers
S-186716 X042 F3	pred all Tint homo
S-186628 Q254* x HT1 F3	pred Q254* segregating, Tint homo
S-186720 X042 F3	pred Tint segregating
S-186632 Q254* x HT1 F3	pred Q254* homo, Tint segregating
S-186717 X042 F3	pred Tint segregating

Mutations detected with AS-PCR

D-11118	1	2	3	4	5	6	7	8	9	10	11	12
A	Tint het	Q254*	Q254*	Tint	WT CODMb Tint	Q254* het Tint	Tint	WT T6ODMa		Tint homo	WT T6ODMa	Tint homo
B	Tint het	Q254*	Q254*	Tint	Q254* het Tint	Q254* het Tint	Tint	Tint homo		Tint het	Tint homo	
C	WT T6ODMa	Q254*	Q254*	Tint	Q254* het Tint	Q254* zy_nd Tint	Tint	WT T6ODMa		Tint homo	Tint homo	
D	Q254*	Q254*	Q254*		Q254* het Tint	Tint	Tint homo	TILL rxn fail		Tint homo	Tint het	
E	Q254*	Q254*	Q254*		Q254* het Tint	Tint	TILL rxn fail			Tint het	Tint het	
F	Q254*	Q254*	Q254*		Q254* het Tint	Tint	Tint het		Tint het	Tint homo	Tint het	**Q254* Tint homo**
G	Q254*	Q254*	Q254*		WT CODMb Tint	Tint	Tint homo		Tint het	WT T6ODMa	Tint homo	Q254* Tint het
H	Q254*	Q254*	Tint	**Q254* homo Tint**	Q254* het Tint	Tint	WT T6ODMa		WT T6ODMa	Tint homo	Tint het	Q254* Tint het

S-186717 X042 F3	pred Tint segregating
S-186640 Q254* x HN1 F3	pred Q254* segregating
S-186613 X042 F3	pred Tint homo
HN1 control	
S-186628 Q254* x HT1 F3	pred Q254* segregating, Tint homo
S-186607 X014 F3	pred Tint homo
S-186624 X016 F3	pred Tint homo
S-186719 X042 F3	pred Tint segregating
S-186685 X029 F3	no markers
S-186665 X041 F3	pred Tint homo
S-186682 X023 F3	no markers
S-186717 X042 F3	pred Tint segregating
HN1 control	
S-186632 Q254* x HT1 F3	pred Q254* homo, Tint segregating

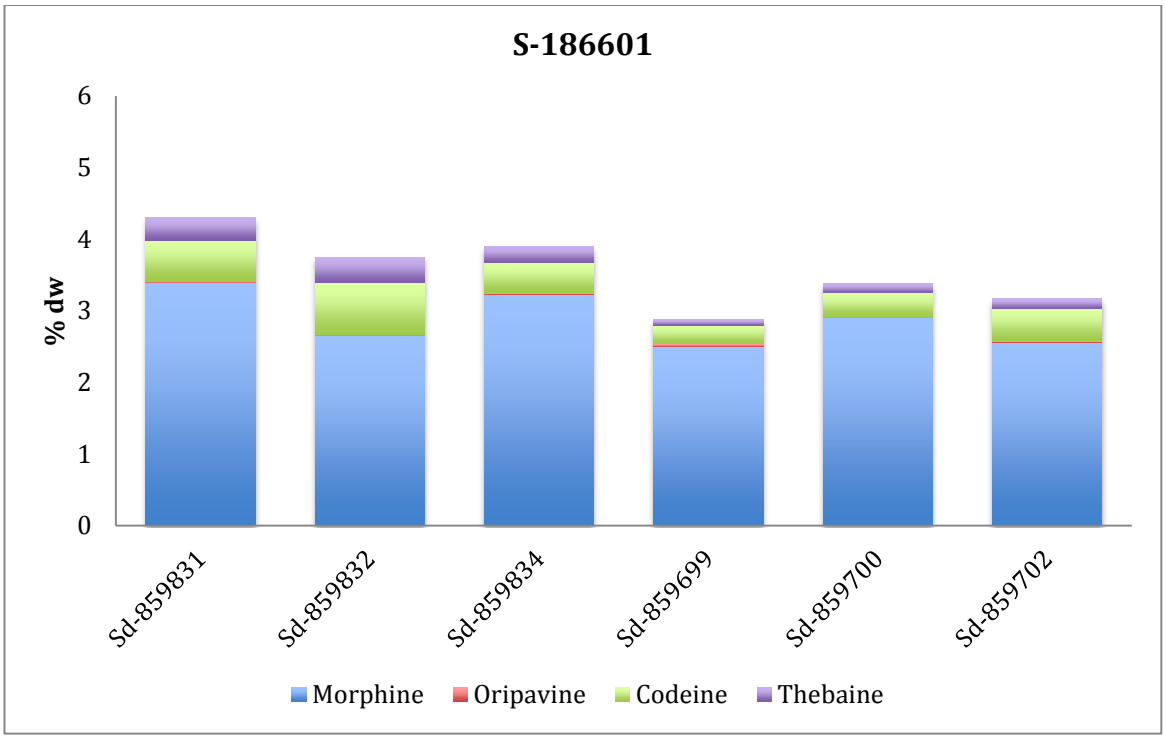
Mutations detected with AS-PCR

D-11119	1	2	3	4	5	6	7	8	9	10	11	12
A	Q254* Tint het	Q254* pos Tint homo	Q254* WT T6ODMa									
B	**Q254*** **Tint homo**	Q254* Tint het										
C	Q254* Tint het	Q254* Tint het										
D	Q254* Tint het	Q254* Tint het										
E	**Q254*** **Tint homo**	Q254* Tint het										
F	Q254* Tint het	Q254* Tint het										
G	Q254* Tint het	Q254* Tint het										
H	WT T6ODMa	WT T6ODMa										

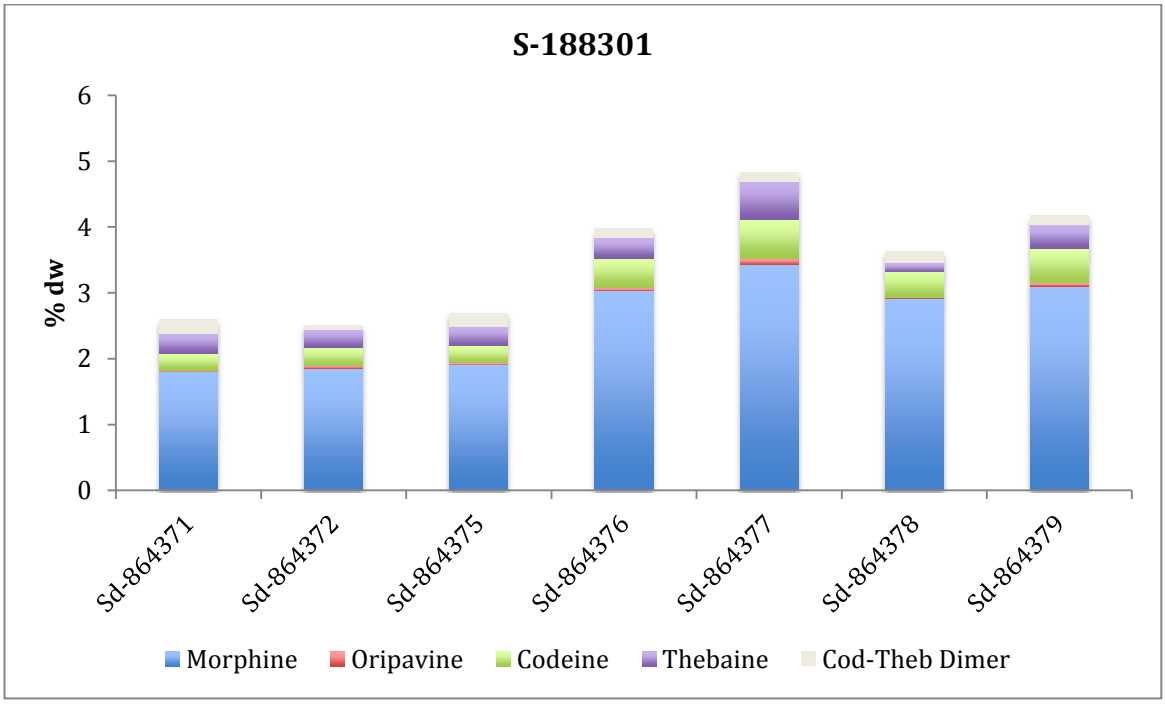
S-186632 Q254* x HT1 F3	pred Q254* homo, Tint segregating
S-186664 X041 F3	pred Tint homo

Appendix I- Alkaloid profiles of selected lines

X008

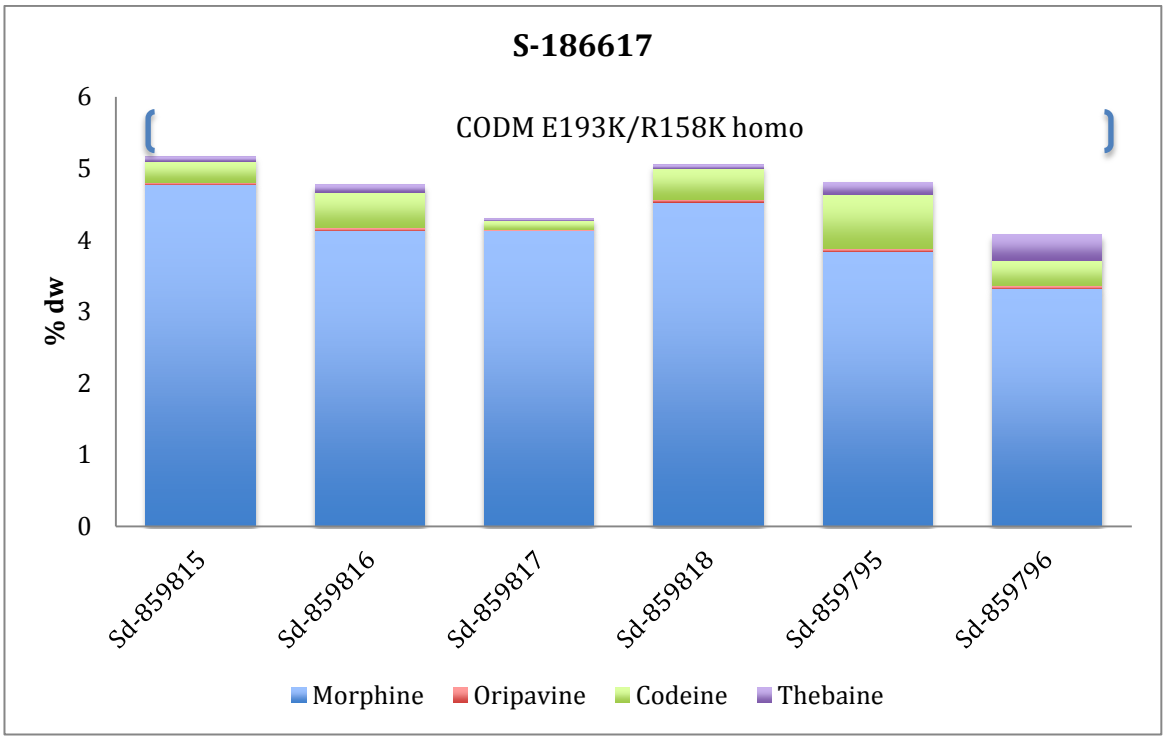


F3s sown in Tasmania. Average 0.46% codeine

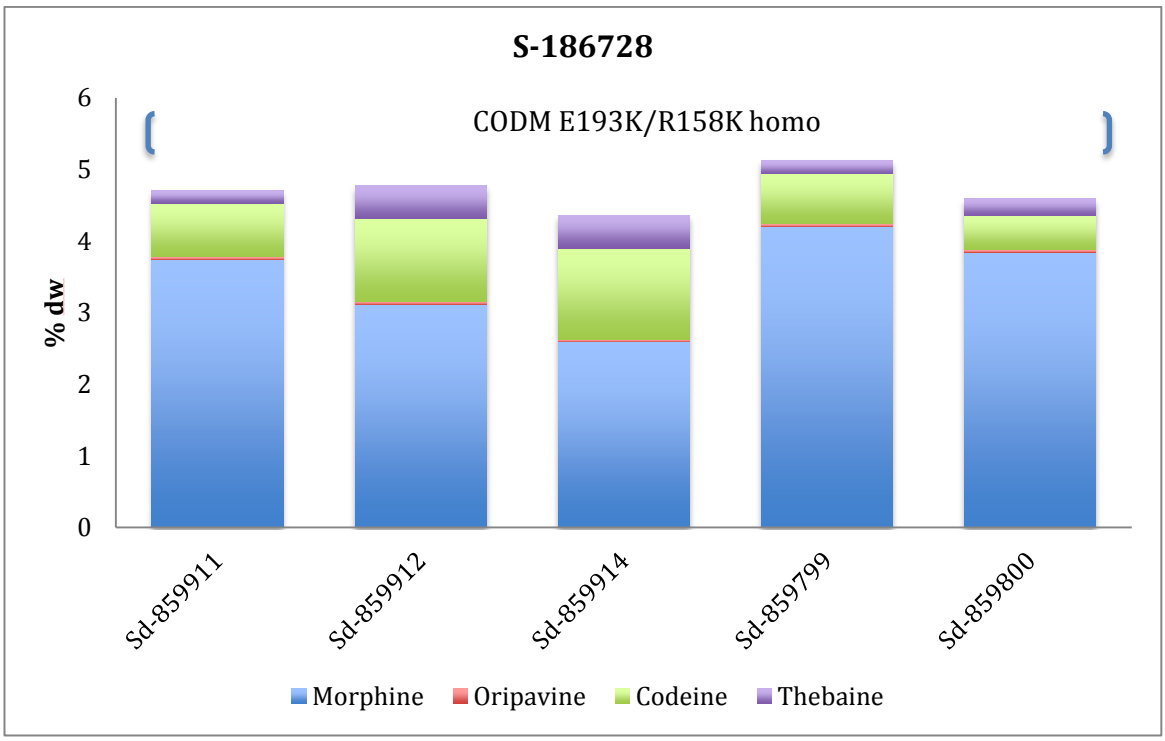


F4s sown in York. Average 0.39% codeine

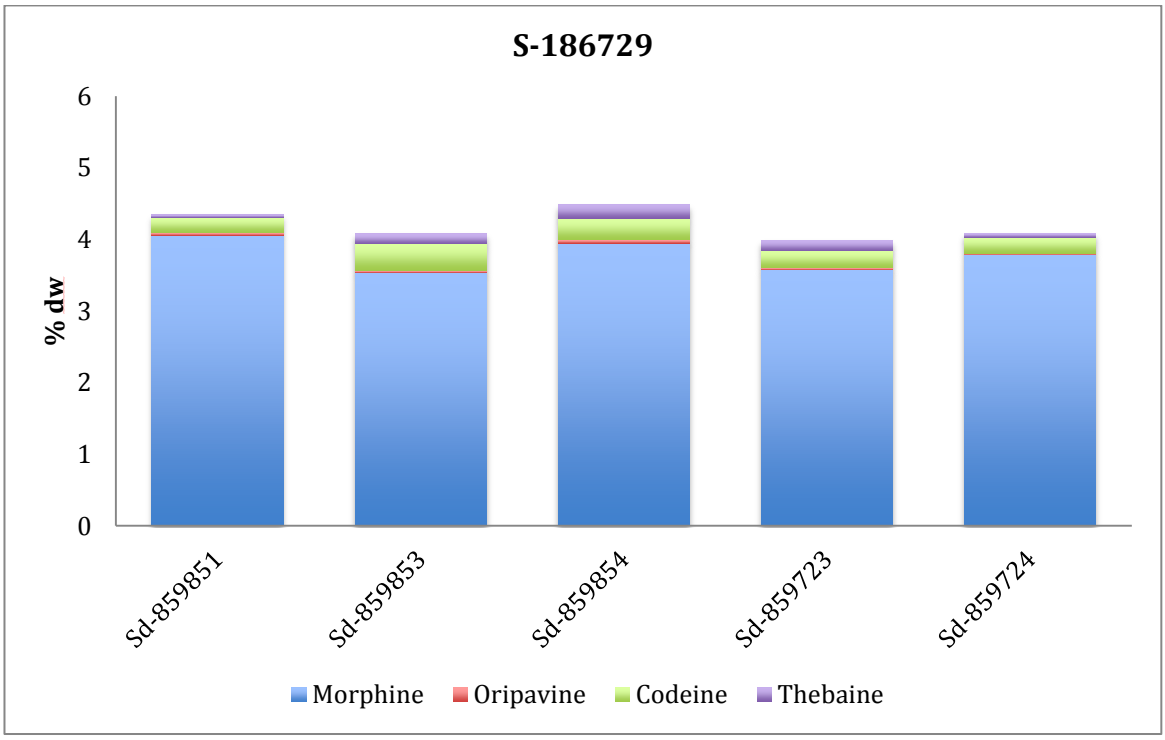
X001



All 8 genotyped had the R158K marker-assumed homozygotes-higher codeine than HM6

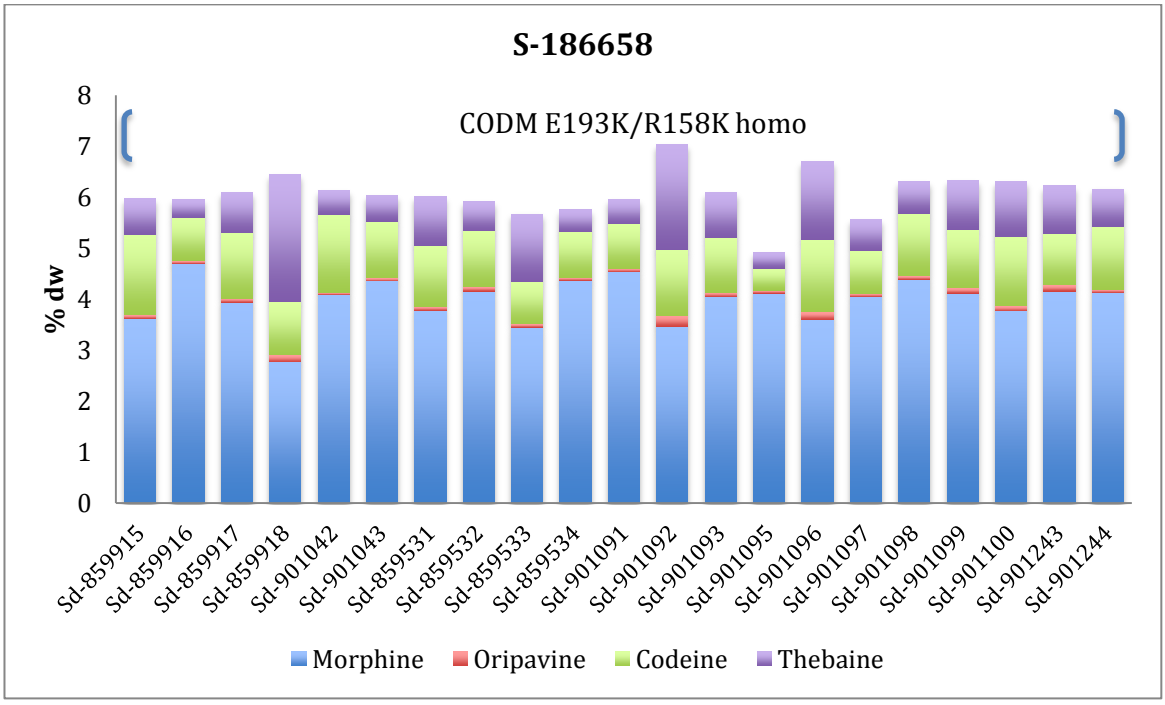


All 6 genotyped had the R158K marker-assumed homozygotes-higher codeine than HM6

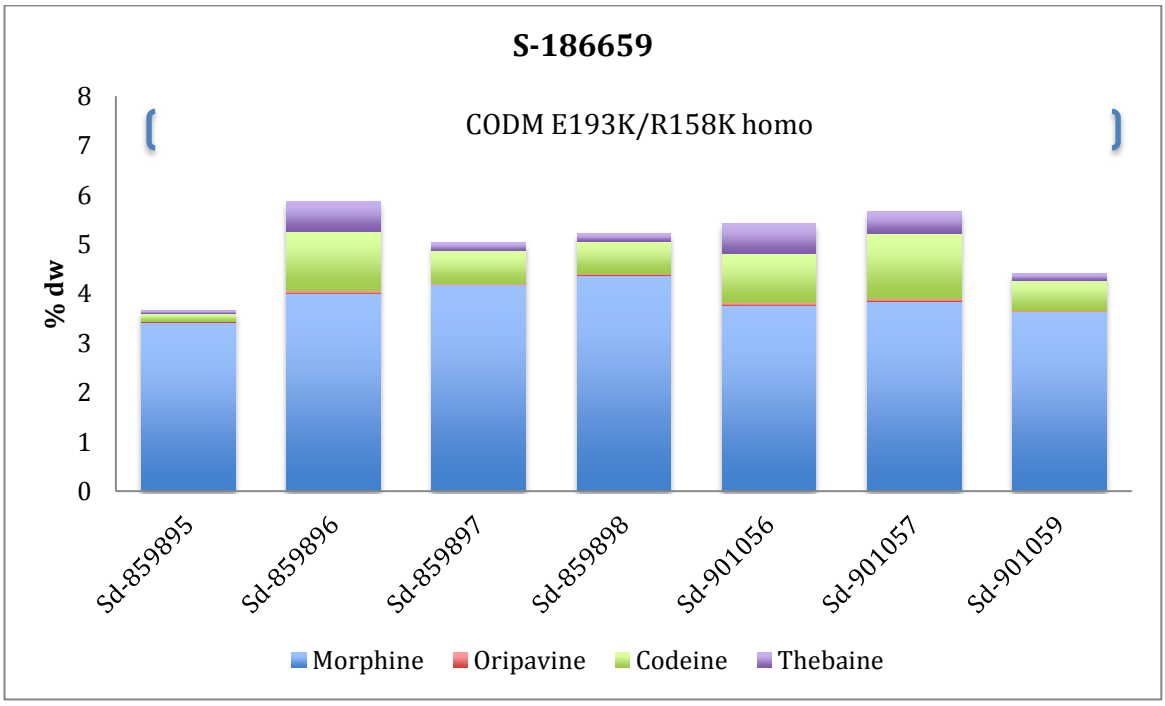


5 of the 6 genotyped had the R158K marker-must be segregating hence lower codeine than other F3 lines with these markers

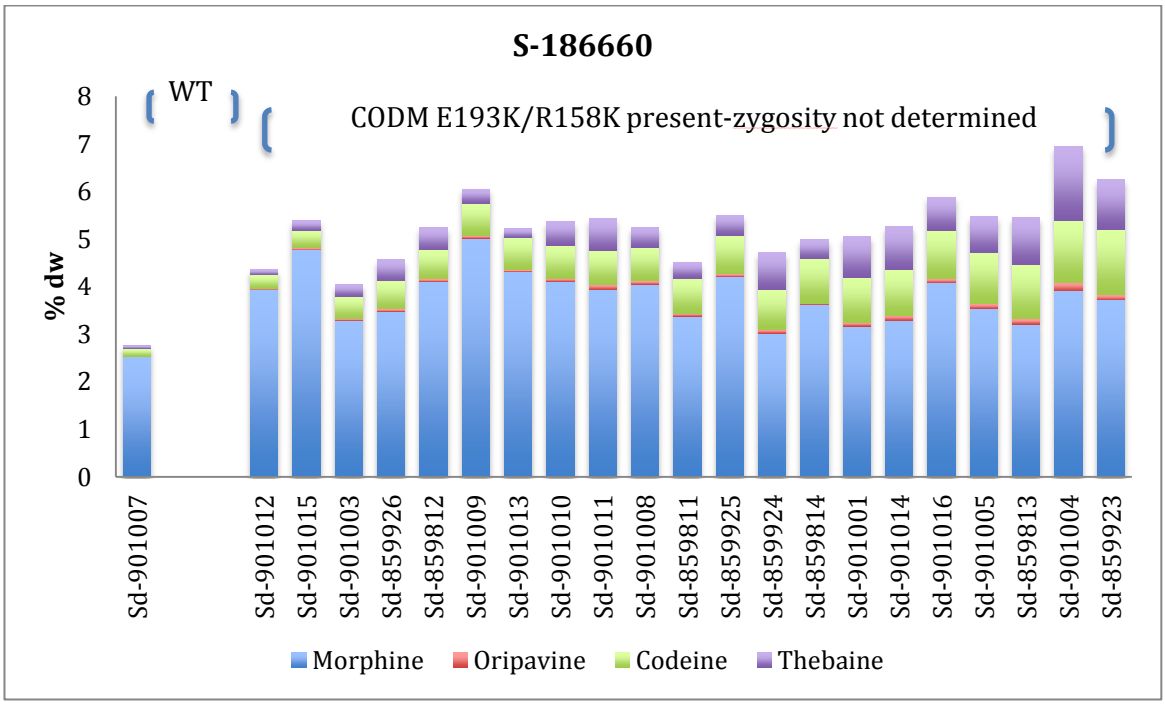
X027



All genotyped had the R158K marker-assumed homozygotes-high codeine (mean 1.11%)

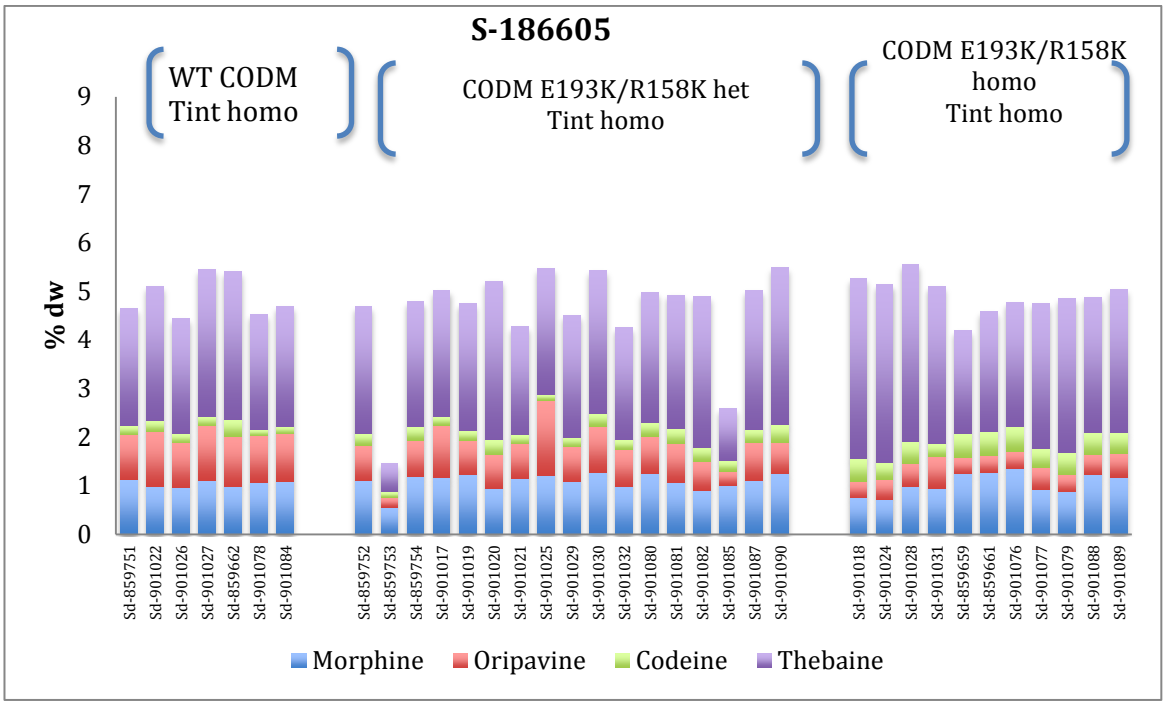


All genotyped had the R158K marker-assumed homozygotes-high codeine (mean 0.81%)

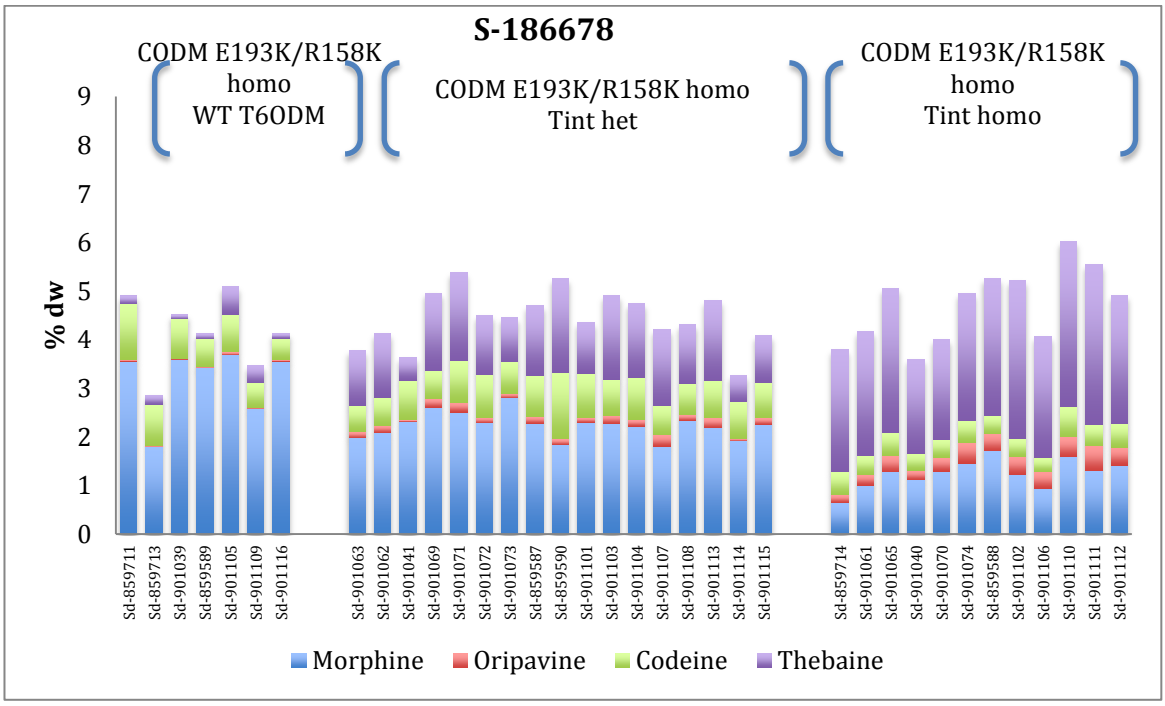


1 of the F3s genotyped did not possess the R158K marker (low codeine). Codeine varies from 0.29% to 1.37% among those that do have the R158K marker suggesting a segregating population.

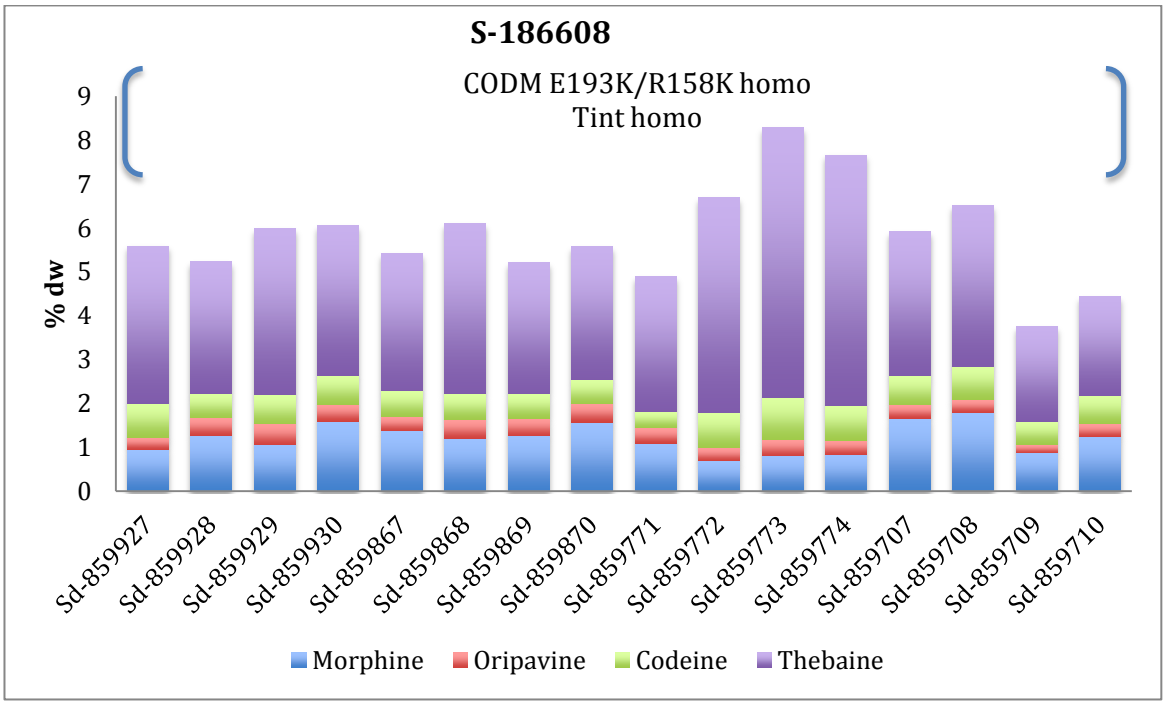
X051



Tint fixed, CODM E193K/R158K segregating. Codeine is increased in CODM E193K/R158K homozygotes at the expense of the oripavine present in high T forward screen lines.



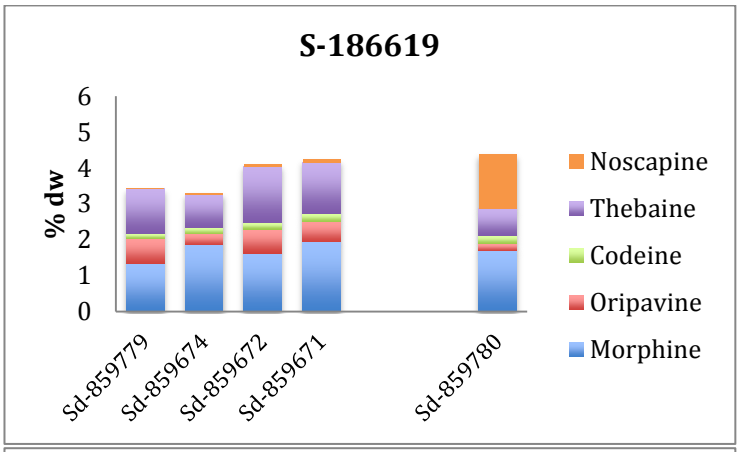
CODM E193K/R158K mutations fixed, Tint polymorphism segregating. The phenotypes demonstrate the usefulness of the Tint marker in predicting thebaine content. Those without the polymorphism had on average 0.23% thebaine. Thebaine is increased in both heterozygotes (mean 1.29%) and homozygotes (2.71%).

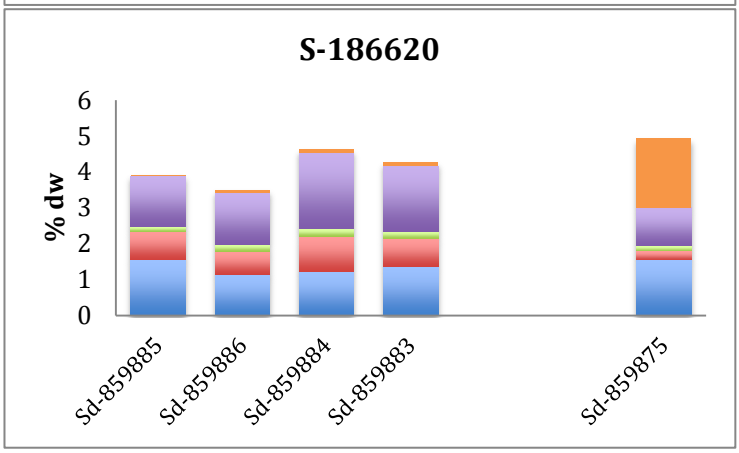


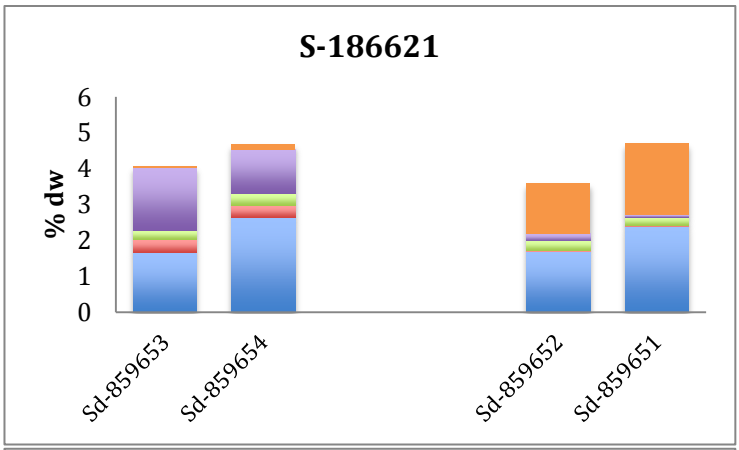
All three SNPs fixed. Phenotypes resemble the third groups of individuals derived from S-186605 and S-186678.

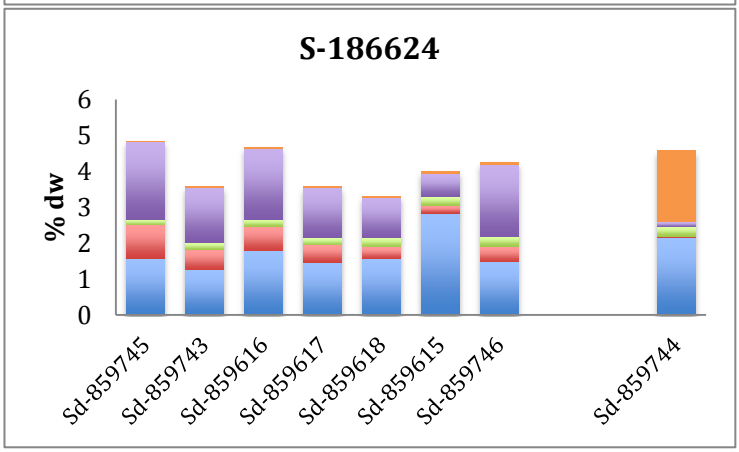
X016

In the F2s the noscapine gene cluster was in a hemizygous state and plants had small amounts of noscapine in their capsules. Segregation of the noscapine gene cluster in these F3s in Tasmania was expected to result in plants with circa 50% noscapine/50% morphinans, similar phenotypes to F2s, or plants with no noscapine in their capsules. For four of the five F3 seed batches this was the case (see below). Although not all the F3s grown were phenotyped it is clear in each case which individual plants (those on the right of each figure) have the noscapine gene cluster in the homozygous state i.e. noscapine comprises approximately 50% of total alkaloids in their capsules.


S-186619
% dw
6
5
4
3
2
1
0
Noscapine
Thebaine
Codeine
Oripavine
Morphine
Sd-859779
Sd-859674
Sd-859672
Sd-859671
Sd-859780


S-186620
% dw
6
5
4
3
2
1
0
Sd-859885
Sd-859886
Sd-859884
Sd-859883
Sd-859875


S-186621
% dw
6
5
4
3
2
1
0
Sd-859653
Sd-859654
Sd-859652
Sd-859651

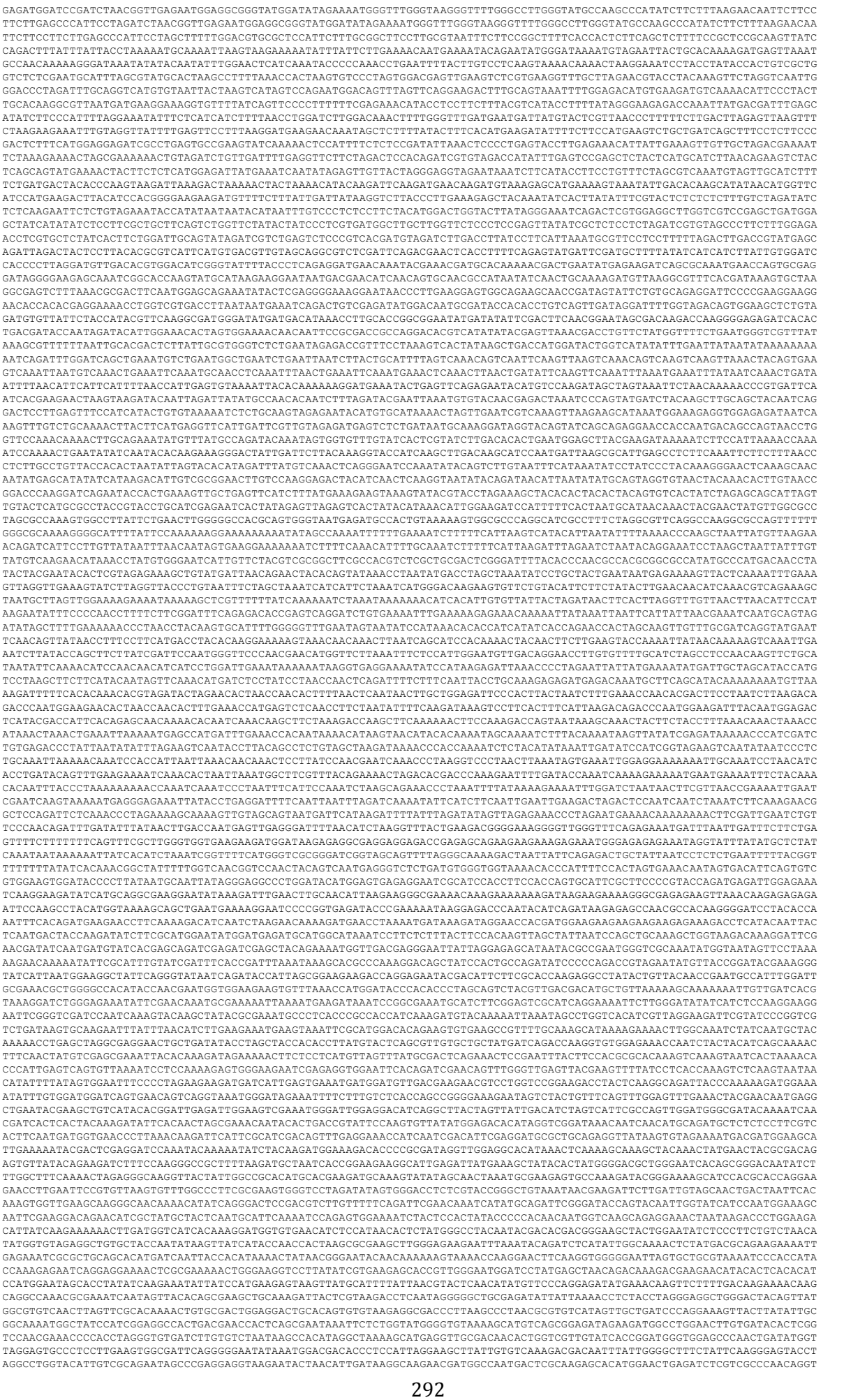

S-186624
% dw
6
5
4
3
2
1
0
Sd-859745
Sd-859743
Sd-859616
Sd-859617
Sd-859618
Sd-859615
Sd-859746
Sd-859744

Appendix J- Sequences of largest contigs containing CODM sequence

In each of the three contigs CODM coding sequence (including the three introns) is highlighted in green and the locations of the two polymorphic sites between different copies of CODM are highlighted in bold/underlined. CODM in each contig reads in an antisense manner.

>PS_89C05_000086 (41 kb)

AGACTGGGCTTGTGATGTGATTTGAGACTTGGGTTGAGCTGGGCTTTCACAGTGAACTGTGGCTGGGCCTTAACCTGAACTGTGGTCTTAGCTGAGCCAAGCTCAGTTGCAGCAGCAGCAG
TAACAGCAAGTAGTAGCAGCAGGGGGAGAACAAGGCAGCAGAAGCAAAGGAAGACAGGAAACAAAATCACACAGGGCCAAGGATGGCATTTACAATGACAATGAAGATGGTAGTGAAACAA
ACTAACAATGATCATGGGAAAAGAAGCAAAAATCAGTGGCAATGAGGCACAAAGGCAGTTAACCTAAGGCAAGGGAGCAGCAATGAAACAAATGGTAACAGTGGCAGTTAACAGGAAACAAA
ACCAAATAAACCAAGGCTGTTAGCCAAGGGAAGTGGGAGAACAAGGCACAAAACAGTTATAAAAAAAAAACAAATTGAGAGCCTAGGCATACAACTAGAGTGGGAAGAGAACCAGTTTGCTC
ACAGTCACTAGTGAGCACTAGTTTCTCCCCACTGCTCAATCAATCCAATGCTTCAAAGGCTTTAGCCTAACATTCACACAAAACATTAAACATCACAGATATAACATAGTAACATACACAA
ATCAAGAGCATTTACAGTGAAATTAAACAGTGAAACATTCTAATTAAACTATCCTAACATAGATTAAACAATGAAACATTAACAAAATTTACAAAATAAAATTAACAGTAACACAGACAAT
TAACAGCAAACACTAAGAATAAAATTGAAAATTGAAATTAACTGAATTTGAAATTAAAACTAACCTAAACATGGAATTAAACAATAAAGTATGGAATTGAAATTTGGACAGTGGCAGCAAA
TGAAGAACCCTAAATTAACAGAACCCTAACTGTTTATGAAATTGAAATTAGAATTTTACAAACTAAAATTAAACCTAATTGGTAACTTGGATAGTCCAAGAACACCCTTCTCTCTACACAT
CAGTTGCCTTTTATAGGTTTACAAGATATCCCCAAATTTCGAAACCCTAATTTTTGAATTGAAAAGTCCACTCACCTAATCCCTGAATTGATGTCGACCCATGCCTATTCTCTCTTCTCTG
CTCCTCTCCATGTCTTCAATTGCGTCTTATAGCCTCACCTAATTTATTGATTTATCACCTAGGGTTTCAGTGGTGAAAAATCAATGAGTTAGATGGCTATAGAGGTAGGGGAGGTAGCTAC
GGCGTGTAGGTGGTGTTTGGTGACAGTGGAGGAAGTGGCGATGGCAGGGTGGTGGTGCGGAAGGGTATGGTGGTGGTGGCAGTGAGAGCAGGCGGAGTGGTAGTTGCAGACAGATAGGTTG
AGAGGGAGGAGAAGACTCGAGTGTTTTGGGATGTAGGGAGTGTTTGGTTGGGTATAGGTGTTCGATACTTGGGTGTTAGGCGGGATATCAAGATTCGATGTTGGCGAATCTGAGTCACTGG
ATGCGAAGCTGGTAGGTGAATCGGATGGCTACCCCTGAAGAATCTGGTAGCGACCGTCGGATGTCTTGATACAACGAAGTTAACGGCGAAGATCGAGGTTAGGTGCTGTAGTGTTTAGCGG
AAGTATCGAGATTTGATGCGCAGGCAAAGAAGCGACCGTAGGATCTAAATGTGATCTAATCTGAAGGCTTGGAATTTAGGTACTATGGTGTTTTGCAGAGACTTCAGATTTTGATGCTCTA
CGAAGGATCGACCATTGGATGATGAGATGGATCCGATCTAACGGTTGAGAATGGAGGCGGGTATGGATATAGAAAATGGGTTTGGGTAAGGGTTTTGGGCCTTGGGTATGCCAAGCCCATA
TCTTCTTTAAGAACAATTCTTCCTTCTTGAGCCCATTCCTAGCTTTTTGGACGTGCGCTCCATTCTTTGCGGCTTCCTTGCGTAATTTCTTCCGGCTTTTCACCACTCTTCAGCTCTTTTC
CGCTCCGCAAGTTATCCAGACTTTATTTATTACCTAAAAATGCAAAATTAAGTAAGAAAAATATTTATTCTTGAAAACAATGAAAATACAGAATATGGGATAAAATGTAGAATTACTGCAC
AAAAGATGAGTTAAATGCCAACAAAAAGGGATAAATATATACAATATTTGGAACTCATCAAATACCCCCAAACCTGAATTTTACTTGTCCTCAAGTAAAACAAAACTAAGGAAATCCTACC
TATACCACTGTCGCTGGTCTCTCGAATGCATTTAGCGTATGCACTAAGCCTTTTAAACCACTAAGTGTCCCTAGTGGACGAGTTGAAGTCTCGTGAAGGTTTGCTTAGAACGTACCTACAA
AGTTCTAGGTCAAAATATAAGCTCAGATTCCATCAAATGTGACATGTGCAAGTCAGTTTAAGCTCACAGCAAAATGGAGAGTTTAGTCAATCTAGCTATCGAAGGTTTGCTTACTTAGAAC
GTACCTACAAAGTTCTAGGTCAAAATATAAGCTCAGATTCCATCAAATGTGACATGTGCAAAACAGTTTAAGCTCACAGCAAAATGGAGATGTCAATCTAGCTATCGAAGGCACAATCCTA
GCACTGATAACAAAAAAAGACATGTGATAAGAGTGTAAAGTGTATCTACACATGTGTAAAGAAAGATCTGAAGTCATGACTACTAGTCACCAAGAGATAGTTTCTCAGGCTAAGAACTGAG
GTCGAAATCTAGCTAGCTGTCCGGACTTTACGAGAATTGTGAATGAGTTGGAGGTATTTCACAATTACTCGCGTTGTACATCAATGGCATACACCCTCCTTGCTTACAACACAAAAACACA
AAAGACTATTTACATGACTCTTATTTACATTGACTACTCTCTTTTATTTTTGGAACAAGAGATGATGGAATTGATAAATACTTGACTTTTTTGTATTTTTCTGAAATTTTTTTTTTTTTTT
GAATAAGGAAACAGTTTTGATACATATACAAAAGGAAACAAAAATTACATGACACTTGCAAGAGGTAGCCCTTTTGATGCACCCAGTTAAATTCGATGGTTGTCTTTCTTAATGTAACC
TCCACCTTCTATCCCTACCAACCAAAGAACAAGCTAGTCAAGTTTCGTTCAGTATTCTAAAGTGATTGGCAATCGTAACTTCCTATCCAACACCTTGAAGATCGAGGCCATACATGTATTG
GTAGATCGTGCGCGTGCAAATTTCTTATCACTATGTGAATTGTGCTAGAATCAGGGTGCCTAAATATCTAGACTAAGACTCCTAATAATTACATATTTGCACAAGAGTCAACATTTCAAGG
TAAATGAGCTCCTTTTTTTATGATTTTTCATTTTTTAATTTTTTTGAATTTTTTCGATTTTTTCAAAAGAAGAAGGAGTTCGTTTTCAATTATGGCATATTATCAAAGTATCTACTTTTCAC
CCCCAAACCTAAACTAAACATTGTCCTCAATGTTTAAAAATATGAACAAAATTATAATACAACATATGAAGAGGATCATGTTGAGTAGAGAAAAAGGAAAGAGAATACCCGATTTCGGCGA
AAGCAAGATTAAACTTCGTTATTCAAGGAAAAAAAATCCAACATATTTCAGCCGAGATCATATTGGATTAGAAAAACATATACAAAATGAACAAAAGGGGTTTTAAGAAATTTTATCTACT
GGATTATATACAAAAATTCACCATACACTAACAATCGAAAGAGTTGAGGATCAACCCAAAAGACAAAGTGTAGAGGTTTCAACAGCTTCACACAATAATAATATGGTAGGCATGCAAGTGA
AGCTGTGAAACAAAATGAGCTACCCCCAAACCTGGATTTTACAGAAGATATAATTTTGAAAACAAAATCGCGCAGTTTCGGGGGTTCATCATGCACAAGGTCTAGCTCGAAATGAACTGTG
CTAGGGACGGGCAAACATGCTAATTCCAACTGTGGTTCCTCAAATGTATTATACTTAGAAGCAAAATAGTCCAAGAGGACTTGGGAAGCACACAGTCCAAACCTAGATTAGGGACTCTCAG
AAAAGTCGGTTTGACAATATGGGTACAAACCAAATCTAGGTTGGATGAAAGGCCATCAGATTGTGACTTATGTAGGACGATAGGCACATCATTATTAATCACATCAACATCTCCTAAGTCT
TCTACACGTGTCACAACATGCTCACAATCATCAAACAAATTGGCAAGATCACAATCAGAGTCATCAACATCATCAACAAATTTATTCTCATGCATCGGCAAATCAGGAGAATTATCACAAT
ATGACCTAGGTAGAGAATCAGACACATCAAATATATTTTCATGCATATTAGTGTCTACAGAAGATTCAACTATTCCTATGTCATGCCCATCTTCACAGAATAATTGTACGAATCCCATGTC
AATATCAAAATCATATGAATTACAATGAGTATTAGAAAAAACAACATGCATGAAGGTCGAGGGCGAGAAGCCATATGTTTTTGAACCAACAGGTTCCACAATATTTTCATGTTCTTCTAAC
ATATCATCATAATCATCATCATGGTAGCATGCATATTGGTCCTCATTAACAGTGGTGGTGTCATTAGTCGATTCATGTTCCTCTAAATTAGGTTCATATTCAACATTATTTACATGAATGG
GACTAGACACTTCATTAGGGACAACATACATTTCCTCATGAATTTCCTCCTTTTGTAGGTGAAGCAAAATCTGATCTAAATATGCATGAATTCGTGATGTAGATCTATCATAGTCTTGCTT
GCAGAGTATTAAAGCTTCTGTGGTGGAGTCTGAATTCATGGGAGTACCAAATTCTTCATGTTCAAATTGTGGTGAATGGTACATGTGTGCATGGTCATTATGGTTTACAAAATATTGATCG
CAACCCTCAAAGGATTGATTATGGTCCCAATGACTACCATTCTCACAATTTATGGGCATTTCATACCTTGGATTTTCTCTACATTGCCTAAAATTATGCAAAATATGACAATTCTCAACAG
GGTGGTCTAAACTACCACATCGCGGGACATGCATAGATTTCAGGTTGCCTAAGAAAATTAGATGTGACAGATTCCTCATGTGATTGCATTTCTAAAGCTGTGATTCGAGCTTCTATTTGATC
CATAAGTGATGATTGTGTGAGTGAGGCTCTAGCAAAATGTTCATTTTCACGTTGCGATGAATGATGCATGTATGAATAATCATTAGGGTTCATGTATTGCGGCTCGGGACTTTGAAAATGT
CCATAATAATTCCTATGACTATTGACATCATGGTCCGTGGAAGGTTCATAACGTGGCACATGCCCATAAGCAAGTTCTCTAAGAGTTTTATTTCCATCCCCAGGAGCAGACAAAAATCCAG
ACATGATTGGCTCAAAAGCTAAGTAACTACGTTACAAAGCCCAAATTTTGGTTTTTAATGGGTTTGGATTTTTGGGAAAAATTTGGTTTTGGTGGGAGACATTTTTTTTTTTTTTTTTTT
TTTAAATTTTGGGAGCAAGTTTTGGTTTTAATGGGTTAAAGAAGAAAAAATTTGGTTGTTAAAATGGGAGCAAGTAAAATTTGGTTTTTGTGGGAGACATTTTTTTTTTTTTTTTTTTT
AAATTTTGGGAGCAAGTTTTGGTTTTAATGGGTTAAAGAAGAAAAAATTTGGTTGTTAAAATGGGAGCAAGTAAAATTTGGTTTTTGAATGGGGGAAAATTTTTTTTTTTTTTTTTTT
TAAATTTTGGGAGCAAGTTTTGGTTTTAATGGGTTAAAGAAGAAAAAATTTGGTTGTTAAAATGGGAGCAAGTAAAATTTGGTTTTTGAATGGGAGAAAAGTTTTGGTTTTTGAAATTAGG
AGCAAAATTTTAGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAAAAAAAAAAAAAAAATTTGGGTTTTTGAATGGGAAAAACCCCCTTTTTTTTTTTTTTTATGGGGAAAAATTT
TTTTTTTTTTAAAATTTAGAAAAAAATTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAAAAGAAGAAAATAAAATTTGGTTTTTGAATGGGAGCAAGCCCACTGTTTGTTTTGT
GTTTGCTTTGGCTCAGCTGGTTTGAAAACGTTTGGTGAAACTGAGTCCAAAATCTCGGCCTAGAATTTGGGTACTCGGCCCACTAACGAGTTAACCTAGCGTGATCGGTTACAAGCTCAGC
TGGGTTTTAAAAACCCAGAATACAAAATACAAGTCCAAATTAAACAAGCTCACAAAAATTAAATACAAGCCCACAAATTAAACAAACAAGCCCACAAAAATTAATTACAAACCCAACAGAA
AATAAAAAGCCCAAAAATTGGCTTTAATATTACAAGCCCACAATTAAAAATTGGAAGCCCACAATTCGGGTTCTCTTAAATGGGTTACCTTTTAAGCACATCCCAGCTGCTCTGATATTGT
TGCAAAAACCCAGTTGGGCTTTGGTTCTTTTCCTTGGTGGCGTCCCAGCAGAGGAAAAACAGGTTCAGAATAGCAGGTCCAACAGCAAATGAAGTGAAGCAGATGCAGATGCAAATGCTAT
GCAGTGAAATGCAAAAATAGCTAATAAAACAACAAGAAAAACACAACCCGAGTCCCCGGCAGCGGCGCCAAAAACTTGGTGGGCCCGGAAAGGTGTGTAAAAATTGAAGTGAAATGAAACG
GGCCTACAGTAATACCGCAAGTGCACGGTTGTCAGTTATAGCTCGTGCAAGTACGGGTCGATCCACAGAGACTGGGTGTGTTTGGAGTTTCTAGCTATTTTGGGCTCTAAAATTGCTATTG
TTGGGCTTTGGGTACCAATGGATTTTGGTAACCAATGGGTTATGATTGTGATTTGGGCCTTTGAATACTCTTGGAATGAATTGGGGTGTTGTCTTAATGAAATGAACTGAGCTTGGGCTCA
GTTGGTTTTGTAGTGCACTAGGCTTTGGTTTTGAATTTGGGCTTGGGCTTATTGTAGTGAACTGAGCCTTTGGTGGACTGAATTAGACTGGGCTTGTGATGTGATTTGAGACTTGGGTTGA
GCTGGGCTTTCACAGTGAACTGTGGCTGGGCCTTAACCTGAACTGTGGTCTTAGCTGAGCCAAGCTCAGTTGCAGCAGCAGCAGTAACAGCAAGTAGTAGCAGCAGGGGGAGAACAAGGCA
GCAGAAGCAAAGGAAGACAGGAAACAAAATCACACAGGGCCAAGGATGGCATTTACAATGACAATGAAGATGGTAGTGAAACAAACTAACAATGATCATGGGAAAAGAAGCAAAAATCAGTG
GCAATGAGGCACAAAGGCAGTTAACCTAAGGCAAGGGAGCAGCAATGAAACAAATGGTAACAGTGGCAGTTAACAGGAAACAAAACCAAATAAACCAAGGCTGTTAGCCAAGGGAAGTGGG
AGAACAAGGCACAAAACAGTTATAAAAAAAAACAAATTGAGAGCCTAGGCATACAACTAGAGTGGGAAGAGAACCAGTTTGCTCACAGTCACTAGTGAGCACTAGTTTCTCCCCACTGCTCA
ATCAATCCAATGCTTCAAAGGCTTTAGCCTAACATTCACACAAAACATTAAACATCACAGATATAACATAGTAACATACACAAATCAAGAGCATTTACAGTGAAATTAAACAGTGAAACAT
TCTAATTAAACTATCCTGACATAGATTAAACAATGAAACATTCACAAAATTTACAAAATAAAATTAACAGTAACACAGACAATTAACAGCAAACACTAAGAATAAAATTGAAAATTGAAAT
TAACTGAATTTGAAATTAAAACTAACTTAAACATGGAACATTAACAAAATTTACAAAATAAAATTAACAGTAACACAGACAATTAACAGCAAACACTAAGAATAAAATTGAAAATTGAAAT
TAACTGAATTTGAAATTAAAACTAACTTAAACATGGAATTAAACAATAAAGTATGGAATTGAAATTTGGATAGTGGAATTAACAGTAACACAGACAATTAACAGCAAACACTAAGAATAAAA
ATTGAAAATTGAAATTAACTGAATTTGAAATTAAAACTAACCTAAACATGGAATTAAACAATAAAGTATGGAATTGAAATTTGGATAGTGGCAGCAAATGAAGAACCCTAAATTAACAGAA
CCCTAATTGTTTATGAAATTGAAATTAGAATTTTACAAACTAAAATTAAACCTAATTGGTAACTTGGCTAGTCCAAGAACACCCTTCTCTCTACACATCAGTTTCCTTTTATAGGTTTACA
AGATATCCCCAAATTTCGAAACCCTAATTTTTGAATTGAAAAGTCCACTCACCTAATCCCTGAATTGATGTCGACCCATGCCTATTCTCTCTTCTCTGCTCCTCTCCATGTCTTCAATTGC
GTCTTATAGCCTCACCTAATTTATTGATTTAACACCTAGGGTTTCAGTGGTGAAAAATCAATGAGTTAGATGGCTATAGAGGTAGGGGAGGTAGCTACGGCGTGTAGGTGGTGTTTGGTGA
CAGTGGAGGAAGTGGCGATGGCAGGGTGGTGGTGCGGCAGGGTATGGTGGTGGTGGCAGTGAGAGCAGGCGGAGTGGTAGTTGCAGACAGAGAGGTTGAGAGGGAGGAGAAGACTCGAGTG
TTTTGGCATGTAGGGAGTGTTTGGTTGGGTATAGGTGTTCGATACTTGGGTGTTAGGTGGGATATCAAGATTCGATGTTGGAGAATCTGAGTCACTGGATGCGAAGCTGGTAGGTGAATCG
GATGGCTATCCCTGAAGAATCTGGTAGCGACCGTCGGATGTCTTGATACAACGAAGTTAACGGCGAAGATCGAGGTTAGGTGCTGTAGTGTTTAGCGGAAGTATCGAGATTTGATGCGCAG
GCAAAGAAGCGACCGTAGGATCTAAATGTGATCTAATCTGAAGGCTTGGAATTTAGGTACTATGGTGTTTTGCAGAGACTTCAGATTTTGATGCTCTACGAAGGATCGACCATTGGATGAT



GGAAGACCCTACCCAAGGTGACCATAAGTCCTTTAAAGTTCTATGCGGGGAAGGCAGTAAGACCTTACATGGGAGGCGATTAAGACCTATAAATATCACACGAAGACAAGCAATCTAAGGA
GCGCAAGTGGACATTGAGGAACAAAAAGGATACACTGAGGAAGGAAAGAAACACACTCCTATCGAGCTAAGTATACCCAACTCTCAGATTTAAATATTATAATAACCAAAGCATGACTTAG
GAGCACAACCATAATTCACGCACACAAGGTTCGCATCATAAAAACTTATTTATACAGATGTCAATATACGAAGTCGCAGAGTCTATATCATCTTCGCATTCGCATACTCGCATATCTATG
CCTGACAATCTTAAAAGCATAAGCTGATAAAAAAAAATTCCATTTGCATACTAAATTTAAGAATCTTCCAAGGAAATAGAAAGCGATGGAAGATAGGTGAGAAATGCTTCCCACCTCCTTG
GAAGATACTCAAACCATATTTTCGCACAAGCACGAAACAAACATAGCTTAAAAATTCAATCAAAATATTAATTTATTATTTAGATTATTCAGATGCATGCTCGTTTATAAGATTAGAATAC
AGAAAGCGACTGAAACATATAGACAAGACACGTAAAGCACGCGATTAAGGTCTCCATCTGCGAAGTCATATCACCATCCACGCATCAAACCCGGGGAGTAACTGGGATCAATATCACTGGA
GTAATTAGGATCAGCATCGCGAGGCGACTCCAAACTAAGACGCGAGCAACAAAATAAATCCTCCCCATAGACCCGACGGTTCAGTTATTCCGAAGAGTGACTAGAGACTCCAAGCTCAAGG
CGAAAATCACAATGAGTTAGATCCTTCCAACAGAATCGGTGGATAGCTCTTTTTGAAGAGTGATCCTACCGACAGGCTCGGTGAACAACTCTTTCTGAATAGTGACCCGAGGATGAAATTG
GGGATTTCACAGCTGAGAGGAAATACCGTCTTGCTCAAAGAACAAAAACGCATCTGCAAAAGGCCCCCAATCATGATGAAAACAAGATCTAGACCTGCAAAAGAAGAAAATCACCAAATG
AGATAAGGAATTTGATAAAGCAAGAGAAAATTACAAAGCATGGAGTACTAAGGCAAAGATTTACCGAAAAAGTGATATACCACATTCAACCATCAAAAACATGAATTTTGAAGAACGCATA
AAACTCTAAAAGCTAAAGAGGAAGAAAAATATGAGAAAACGTTCCTCACCACTCCTGGTATTTATATAGGGGAAGCAATAAAGACCCACCGCCGGAAAATAAGGAAGTTGACAACATGTGT
AATCAATATGCGGAAACCACGGAGATACGTTGAATAAAGAAGATTTACCAAGAATTCTCGCATTATTCCACTTGCAGAAGAATAACGCGAGAAGAGGCAAAATGTAGGGGCGAAATATCGC
ACAGAGCTAACCCATAGTCTCGCCCTCACTAAAATAATTACTGGTATAAGCGAAGAGAGAAATTAGGAGAAGGACTACAACGCAATAAAATGGGACAAACACGAGAACACCCACACGACAA
GCACAACACGAAGAGCACGAAGCCGCAACGCGAGATATAGTGACATCTTCCCAATGCCTGGGGAATAAGACAGAGTGAGCAGGAATCAGCAAACAGAGATTTGCATGTGAACTCAATTGTC
TTCTACCATGATGACCGACTATATATGGGTTCAAATAATGAAGAGAGAGGCAGATTGAAAAGAATTCAGTCTAGAACAATAGTAGAAAATCACCACTAGAAAAAAGATAATAAACACTAGA
GTTATCTTATTATCATCATCTTCCTTATATGTATAAATCTTCATGATATCAATAAGAACATCACTTCACTCCATACAAAGCTTGTGTAAACATCATATCTAGTGTCAAGGGGTGAATAATA
TCATTAATGTTTGAGTTTACTTGCATGACTTAGATTTAGTGTCTTATCATTTCTTAAGGTGTAGTTGTGGGATTTTCCGCAACTACACTGGAGAATAGTATTTTTCTGAAAATGGGTAATA
TGATAAAAGACTTGATTGAATCAACTGATCTCATATCTGTTCCTGGCAAACAGTCAATAGCCGAGGATCTCACTATGTCAGTTGAAGAAGCTGTGGACCAAGTTACTCTACCCGTCAGTGA
TTTCTCGCTCCTCACTTCCGGTTCACCTAAACGGTCTAAGATCATCGAAGAGCTTCAAAATGCATGTCTGCATTGGGAATTCTTCATGGTAATAACTTATTTATGAATAGAATTTGAATTA
TAATGATTTTTGCACTAGTTACATGCATATTGAAGTAGTTTATTTATTATGAGTTTGCTCGGGTGACAATTATAGATGGTTAAGCATGGAGTTCTAAAGAGTTTGAAAGAGGAATTGATCA
ACTCCTGTATGAATTTTTTCGACTTAAGTCCGGAAGAAAAGTCGGTTTACAACGGAGAGAATGCTTCAGACCCATGCATTTTGGACCCAATTATATGTGGAACCAATTTCAACCCTACCTC
TGAGAGTGTCTTCTTTTGGAGGGATTATCTTAGGGTCTTAGTGCATCCAGTGTTAATTCCCCTGTTAAACCCTAGGGTATGAGGTAAACCTCTAATGTTCTTAAACGTGTATACTCAGTT
AGACTTGTCCGAGTTTGTAAATGCGCCTGGCTAAATTTTGTTTTGTTCATACAGTCAGGTTTTAAGTGAGTACAACAAAAGATCAAGAGAAATAATGAAAAAACTGCTCAGAGCAATTATG
GAGGGACTAGAATTAGAGAAGAGGACATCGAAAAGTCCCTTGAACTAAATTCTGTTGTTCAAAAGTTTTCTGCAAATATCTATCCTCCTTGCCCACAGCCTGATCTTGCAGTTGATTTGCT
CCCACATGCTGATCGTGGTCTCTCAACCCTACTCATACAAAATGACGTTGGTGGATTGCAAATTAAGAACATGGGGAAATAGGTTGCTTTGGATGCGGTTCCAACTCTATCATGGTTAACA
CTGGTGATCATATGGAGGTACATGTAAATTAGTTCTCTGTCTCGTTCTCGTAACCCTAGTCTTATTTGAGTAAATTAAATTGCCGCGATCGTGCATTTATTATTGATGTTGGAAAACGATT
AATAAAAAAATGATTTTTTAAAGTTTTAATAAAAAATTATAATTTATTGTTTTATTTTGTGAATGAAACTTTTTAGTCCCACATCGTGGAGTTTCCAATTTTTAGTAGTTTTAAGAAACTA
TATAAAGCTTTTAGTCCCACATCGGGGAGTTTTTCTTCTTAAGTTGTATTTGTCAATTATATAAAGAAATTCACTACTTTTGTAAAATCTATGGGAAAGGGGTTGCTCTATATTTTAGAGC
GACCCCTAAGGGAAAATATTTTATAGCGGTTTTTAAGCATTCGCGATTTTCCTTAACGGTTTTTCTCGGAGTTGCCAAGCTCAAGTTGAGCATCTACTACATATGCTAGTAGTAGGTGTATT
AGGGTGTTTTATCCTGGAGATATCCGTCCTGTGAGGGCTATAGCATCACTCTTGAGTGTAGCCGGGCGCTAATGTCTTAAGGGCAGCGTGTTGAACACGTGACTCACTCTATTTTCCAAAG
TTTTGCCTTGTTGCTGTTGCGGAGATACGGGGAGCTTGTTCGTTTCGTCAAAGCAATCACTTCCATTATAAAGGAGCTAAGTATCAATAACTTTTGTTTATTTGATTTTTCTTTGTTTTTG
ATTATTGCACCCAACAATCTTAAGACATTAGTATTTGTAATAATCGAAAACGATTGATTGGTTTGTGAATCATGGATGTTAATTGTGGAGTGAAAAACAAAAACGAATTTCTGGTCAGATG
TAGAGTTTCAATTTTATCTTTTAATATAGAAGGAATTTCGATGAACCCTTTTTACACAACGTAGTAGACATCCTGATAGTTACCCGCATAAAATTTCAGAATTTTTGGAGTTGTACAAGTA
TTTGTTTGATATTTTACAAAACAGAAAAACGTTTCTGAAAAATTCTGACGAGCAGAAAATTGTTGTTAACTAAATTAATTTTTGTGAGGGTAACCATGAGTTTTTTGAAACGCTGATTCAA
ACGAAGTTTTTATATATAGATGTTATCTTCAAAACTCATATTATACTTTCTGATTTGGAATTGTGATTTGTTAATAAAGGTGTCATCTCCGAGGGACATGGCTAATAGCCGCATGTGGTTG
GAAACAAATCTTCCAAAGATGGCAAAGGTGAGAACATCGAACGCAACTATAAGCATGGTCTTCATGATAAAGGTATATTTCGTAAACCTGAATCCAGAATTACTTTAGTTAAGGGTGAGTG
TTATGTTTGTAAAATTCCTGGCCACGCGGCAGTAAATTGTAGACAACATAAAAACCTTAATAAGAAGAAAGTTAATGCTAATTTAGTTGAAACAAACTAGAACGAGTTTGGTGGCGTGATG
TCGGAAGTTATTTTAATAACCAATGTGAGAGACCAGTAGGTGGACTCTGGAGCCACCAAGAATGTTTGTTGAAACAGAGACCTGTTCACCTCCTATCATAGGATAGGGGATGTCGAGAAAC
TCTTATTGAGTAACTTATATACAATAGAGGTTGCATAAAAGGAAAAGGTCGAGCAGAAGCTCATATCTGTAATATTCTCACATTGAATGAAGTTTTCATGTTCCGAGCATATGCGAAAATC
TTGTATCTTGTTCTGTTGTAGATGGAAAAATATTTAAGATCTTAATTGAATCTGGAAAACTTGTTGTAACTAGGCAGTGATTTTTTAAGCAAGAGTTATAGGACTTTGGGTCTATATAAGC
TTAACGGAAAAACTGATGATGTGAACGTAATTGATTCTTGTGATATCTTTGTGTGATTGATTGTTTTACGTGGTAGACTTGGAACCGTAAACTTATAAGTCAATGCTTAACTGGCTAGCAT
AGGCTTCGTACCCAAATTTAGTTTGGATTTTGAACACAAAAGTGAAATCTGTGAACACAAAATTTGCATTACGACCTGTATAATTACGTGTTGTTTGGTTAGTTTGATTGCGTGCTTTGTC
AATCGAGCGAAAAATAAATTCATGCTTTGTGCGTAATCTGTGGTTTTCTATTATGTAATGTCAAGTATGAGTATTATATATGTTTCATGCATCTGTGTTCTTCTTTTTTATGTGATTTATT
CATCTTTTCCTCTTACCTGATTTATTCTATTATTTCCCTGTGATTCCTCTTACTTATGAGGTCTTAAACATGTTTTCCTGCTCTGATACCAATTGAAAAAGCGGGGATCTAACAACCACAC
CAAATATTTCGCTTCGTCAATCTGAATGGACTAACTCCAATATACTTTCAAGAGAATCAACTAGACAGCCAGACTCAATCTTTGGAAAAGTATATCAAAGAGTTATATCTAAATTTCTCAA
ATCAATCTACAATCAAACAAATAGGAATTTGCGAGCCCGATTGAATAAGAGAAATAACTTGGACGGTATCACAAACCAATATCCAAGTGTCAATCAAGTCAATCAACAACCCAAAGGCCGG
ATTCAAGAAATTGATTGAACTTACGCACAACCTGTGATATTTCAATTATATAGAGAATATAATGCGGAAAAGAAATAATACAGACACCAGAATTTTGTTAAAGAGAAAACCGCAAATGCAG
AAAAACCCGGGACCTAGTCCAGCTTAGAACACCACATTGTATTAAGCCGCTACAGACACTAGCCTACTACCAATGAACTTCAGACTGGAATGTAGTTGATCCCTAACCAATCTCACACTG
ATCAAGGTACAGTTGCGCTCCTTACGTCTCTAAATCCCAGAAGGATTCTACGCACTTGATTCCCTTAGCTGATATCACCTACAACTAAGAGTTGCTACGACCCAAAGTCAAAGACTTGATA
AACAAATCTGTCTCCCACAGAAATACCTATGAGTTTTGTTTCGTCTTATGATAAATCAAGGTGCACATGAACCAATTGATAAACCAGACTTATATTCCCGAAAAACATCCAAGTATTATCA
ATCATCTCACAATAATGTTAATCGTATGGAAGAGAAACAAGATATTGTGGAATCATAAACGATGAGACGAAGATGTTCGTGACTTATTTAATCTTACCATTGGAGATAAATCTCGAGAA
AATCTTAGAGAAGATAGTACTCAATCACGATAGTATAAAGTAAGATTAGAACACGCAACTACAAAAAAAAAATAGCTGGGTTTGGCTTCACAATCCCAATGAAGTCTTTAAGTCGTTAACCA
ACAGGGTTTTGGAATAAACCTAAGGTTGAAGGAGAATCGACTCTAGTCGCAACTAGTATCACACAAGAGGTGTGTGGATTAGGTTTCCCAGTTGCTAGAGTTCTCCCTTATATAGTCTTCA
AATCAGGGTTTGATAGCTTTGGAACAAAGCAAACAATATTCACCGTTAGATGAACTCCTGATTAGATTCATGCTAATATCTTTCAACTTTTAGATCGAACTTAGCTTGTTACACACAAATA
AAATGTACCTTCAGTTAGATATGGGTAACCGTACCTAAACATGTAAACCAAGTTGGCTCAACAATAGTTAACTGAAGTTAGCCATATGAACACTTTCATATCAATCTAATTCATCTTAACC
AAAACTAGTTCATATGACTCAAATGAAACTAGTTATAGAGTTGTTCAATTGTTTATATTCTCATAGACGTATACAAGACACAATTGAAGCAAAATTGGTTTAATTAACTCGAATCAATTCA
TGAACATTATAGCCACGGTTTGCAAGAGATTGCATTCCTTATTATATAAATGTATTAGTTCATGGCACAACCAATTTTAGAACATAACCCGCTCAAGTAGGCAAACGGGTACGCCAACCTA
AGTAGCCGGACTTGGACTGTGTTCGATAGTACGCAAACGGGTATGCATACTGTCACACATCCAAACTTCAGTTGAACCCAGTACGCATACAGGTACGCATACTATAGTACCCGTACTTCAA
CTGACTTCGCCAGTACGTGTACATGTAAGCATACTCTAGATCCCGGACTTTACCTGACCTTGCCAGTACACATACGTGTATGCATACCACACCGGTCTTGGATCCACATATATGCAAGTAC
ACATACTATGTTTATAATCCAATTATGGTTAAGTGTTCTAAACTCCATTTCAATCATTGAAACATTCTTAGAAGACGACAATAGCTATCTCACACAGACTATTAGCTTCAAAGAAATTTTC
AAGTGATCGAATGATCAATACGAAACATTCCGAGTCTACATCAAATGACTGTCTCATACAAATCATATAAGATGTTACCAGATGGCTTTCACAAGATCATCTTTTGACTTTCGTCAAGAAT
ATAAAATGAACTTGGTTAAAGTGAAAGCTTACCAACATATATTTCGAGAAATATGTAAGCGAGTTAAACTCAGCTCGAAATATCAAATGTGTATAAATCGAATATTATATAGCTATACGAC
TTTTGTCTCAAATAGGAGATACAGTAGAAACAGACTTTTTGAGTGATATATTAGTTTCAGTCTCCACATACCTTTTGTTGATGAAGTTCCACAAGCTCCCCTTAGTAGTTATTCGTTTTCA
ATTGATGAATCCCGTGAATTCTAAAGATCAACTACACATTCTATCCTAGTCCGAGACATAGCTATAAGTAGACTAGAAATCAAGACTTATAGTTTTGATCACTAAACTTGACAAAAAAGCT
TGAGATAACAACGATTGCAAGTTTGACTGAGCAGTGCTCTAATAATTTCTGGATCAAAAAGTGGAAAAGGATCGAAGACCCTCCTTTAATTGTCCAAGCGCGATAGGCATGCGAGTCGCGC
TCCATGGTTCTCTTTGAATCCTCTCCTCGTCTAACAGGGTAGTGGCAGAAGAATCCTCTGGAACTGAGAGAATGAAACCACTCTTTTTAAGTTCTACTTGAAAAGTTTCAAAACTTTCTGC
GGTCTTCTTACCGGATTTTGGTTTCGAACTGCATGATTAGTATGGGAATAAACATTGGAATACTGAAATTGAATAACATAACAAAGCTTGGAATACGATAACAATGGAAAAGGAAAAATCA
CAGAAATTATAGTCATGTCGTAAGAAACGCAAAATGCAAATAACTGAGGCAAAATTAACTCACCGATAAGATGATGTAGCAAAAACAACAGTGGGAGCAGACGTATTTTGATCTGCCATGG
AAACATAGGCGAAGTTGAAGAACACGGAGGAAGAAGACAAAGTACGGAAGAAGAAGAAGTTGAACAGAGGAGGAAGACGGTAGAAAATAAAATATTTTTTTCTCTCTCCTGAGTATTTATC
CGATAAAAAGAAGTAAAGACGTGTCGCTTGAAAAGGAAAAACTTTCAACACGTGTAGCAGTTACGCGGAAAATACATAAGTGTGTCGGGTAATACAAAGAAGATGAAAAGAAAGGACTGAT
GCGAAGTTTGACCAAGAATTCTCGCATTATTCTACTTGCATAAGAACAACGCGACAAGAAGCAAAATGTAGGGTCTAAATATTGCACAAGGCGAACACACAACTTCGCTCTAGATAAAACA
ACTCCAAGCACGAAGATATTATAATAATTAGAAGACAGGGCGAAGAATAGTCGCGCGACAATAAGACAAGATTGCGAGAAGATAATCAAGTGATAAAGATAGCACGGAGACCATATTCAAA
GGAGTACCATGAAAGATACAACAGACAACACGATGCACATTCATACCTTCTCAAAAGTCTGGCGCATAAGACAAACTGAAATAGAATCAGTCAACTGAGATCTTGGACGTGAATCATGATT
GTCTTCTATTGATAAGTTTGCCTGTATAAGGACTTAGAAATCTAACAGAAAGAGAGGGGTCGATTAAGAGAGAGATAGCTAAGAAAATTATAAGAAATAAAGAAATGAATTTCATCTTTAA
ATTTGGCTTTCATCTTGTATTCCTCAATTGTTGTATGTGAAATCTACCTTGAATCTACAGTAAGAACAATACATATTCATTATCAAGAGATTGAGAAGATCATATAAGTGTCAAAGTGCTG
AATAATATCATTAATGTTTGAGTTGATCAAAATGACTTAGACTTAATGTTATTATCAATACAAGGGTCTAGTTGTGGGATTTTCCGCAACTACATAAATGTTTCACTGAAATGATGAAAAA
ATTCGTGTTCAGATTTTTTTTTTCCTTTATAACTTAACTGAAAATAGAGAAGAGATAAGGGGAACAGAAGCCGATATGTTTTGCTGGAAAAAAAAAATGAGTCATGTGAACCTTGAGTTTT
TTATTTTTATTTTTTTAGTAAAGTTTTGTCACTTGTACACGGTGGACTCGCCGCATCAGCGCGAGAGGAGAATAGAACTGCGAGAAGAGTGGAGAATATCATTACCGAAATTATACACGCA
AAAGTATCATAAAGTTGAGACGAAATGAAAATTCTTTCTGTATGAATCTTTAATTTAGTCGACAACTCAAAATCCAGACAGTGTCGTGTACTTGTGCCAATTAGGTTAAGCCAG
CACAACCCACGAAATTAACATGTTGGGTCGTGCTGGACTAAATCACGTGCTGATATGGCCTCAGATTTTAGCGTGTTGTGCTAGATTGTGCTGGTGCTGGTACAACCCAATTTACACCTCT
AAGCACCACCACCAAGCCTTCCTTTAGTAGCTGTTTTAGAAGCAATGGGATCAAGATTAGAGCGATAAACACAACAAGTAGCATGAGAAACCAACCAACGTACCCTAGCCTAACAGAGACT
TAACAGAAAATGAGAAGACAATCTTCAATCCAATACATTATTTATTACATAATATTAAGCACCATATGAATTCTTCAAAAGCAACTTATTTCAACCGTGATTACATCACTTGACCCAAACA
GAGTCTTACATATAGTGAGACGAGAATCTTCAAACCAATACATTATTATTACATAATAAATCTCTCAGTCATCCACGAAATTTAAGCACAATATGAATTCTTCAAAACCAACTTATTTCAT
CATACTTCAAAAGCAACTTATTTCATCATATATTAAACACAATGTGGAGTATAATAATATGTTCACACTTTCTCACATCCTCATGTAGTCGAGAAATGATTTTCCATCAAGCTTCCTTGAAA
GATTTTCCTTCAAAATATCCTCATACCTACCTCTTTTGAACAAAGCAGGTGTCTCTGGTGTGACCAAGCTCGAAATTGGGCCTATTTCTGACTCTAGTTTAGAGTCATGAAATGTCGCGAT
TGAGAGCCTCTCCTTTGTTGAGTTTACTACTGCCCGGTGCTCGACGCTACGGTAAATCCCATTAGTCATTATCTGTAAAAGGGAAACATTAAGTAACTAAGCAACTTTCGGGGTACAGTC
AAAACAGGTACTCACAACAAGGACCAGTAGGCAGGCCATATATGCCCTTAGCCTAAATACCTCCAAAATGTCTCCAACATTCACTATGAACGCATCAGGTAGAGGTTTGATTGAAATCCAC
CTCTCTTCTTTTCTTATTTGTAATCCTTCAACTTCATTAAGTTGAAGGAGAATTGTCAAACCGCTAAAATCCGAGTGTGACGTAAGACCAAGTACAAGCTCTGGTCGAGGACAAGGAGGAT
AATAGTTCATCCTCATTGTTTGCAACCCATCTTCAAATAAGTCTGTCATACCTTTAATCTCAACTAATTGTAGAGATTTTCCAACATCTCAAAGACAACCGTTTGATAGTTTTTCATTTT
TGATAGGTAGGATTCCAGTGTCTCCCTGCATAAAGGATGTAAATTAGGGATCAAATGCAACTTCAGTATATTTTCCATCTGACAGATTTTAGGGATGCATGATGTTTGGATTCGATATATT
TTGGAAACCATGCAATTCACAAACCCGGATTGCTTGATACTGCCTCCGAATGTGAATTTTAGGCAACCTAGACAGGATTTTCCATGGTTTATCCTAGCAACCTAGACGATATTGATGAAAC
TATTCTGAACCAACCAAAATAACATTAAAAAAAAAAATATCACGTGCGGCATTGAGAAAAAGCAAACGGAGAGAAACAATAACCTGAAAGGCAGAGGGAGTTCTGGAAACAAATGAGGCT
TCCTTAAATGGAGAGGAAGACTTAACATGCTAAACACTTCAGTCCAATCAAGTCTTTGGTCCTCCGATTCAATATAGGGTTGTCCAAATCCTTCAAAATCTCCATCTTGCTGTCCGTATTT
AGTTTTCTCATTCATTGGAAGGTTAAAGAAACCTTTAATTTCTGATTTTATATTGTCCATCAGTAAAGCGTCGACTCCATGGTTAACCAGCTGAAACATAAAGCAGATTAGCATGAACTAG
TAAAACACTAAATAGTCGTCAATCAACTACGTATAATCAAAATTTGTATACCTGAAAGAAACCCCATTCTTTGCAAGCAGAATGAAGCTTATCCAATTCTAACTTTCCAACTACGGGTTCT
GGAGATAGTAAATTTTGCAAATCGATGACAGGAACTGTTTCATCATCTGTTACAGACGCACCAATATTATTCAACGGGCTTTCACCGGTGCATGTGTATCGAGATGGAATTTCTGCAAGCG
TGAGTTTAGCCAATTCCTGAACACTTGGTATTGACAAACCATTGCCTAGCTTGATAAGTATTGGTGTCTCCATGAACTCTATATTTGATGAACTTTCTTGTCAGATCTTCAGATCATATAT
CAATCTTTCAGGTCTCTTTGTTTCTCTCAAGTGGATATGCCTTACTTATATCTATATTTCACTCGATGTCTAATGAATAAATGACTAGTGTTTTTTCAACGAAGGGTCTGAGCAACCACTG
CTTACCCTACCAGTTTTAACCTGATGATGTCTACGATCTAACCTGCCGGTCTAAGTTGCAAGATAATCGTGAAGTTTCGGTTGTTAGGTGATTAGGACTG
ATTTTTGTTGTTCTTCTCACTTTCTATATCTTTCTTGTTGCAGAACTTGGTACGATCTAAGGTATTAGATTGGGTCCATGCCTACATGAGCAGGAGGACGTGGCAAATGGTGTACGATCTA
AGGTATTTTCCCAAGCTAATGTAGCGACTGAGGAACGAGAACCACTACTTCAGTTCATGTTAATAATAAAGTTAATCCTTTTCTTCACCAAATCTCAATCAAATCCACACAAGCAACTGCT
TTTTAGTACTTTCTATAGGCCATATCTCCATACCTAAAGTTCCTAGAAACAAACATAAGATTGAATTTATGATAAGATGGAAATATGATAACTGACAGTTAGTGATAGAAGAACGGTGAAA

AAATTCACCTAGTAAGAATCGGTTGTTGATCATAGAAAATAAATAAAGTGAAGATCTCGATTAGTAATTTAGTACTATTTTCAAATATTACTACCTCTGCATTCAGATTGAAGATCCTACC
ACCACTACCACTACCTTTACCTCCGCAAGAACCATGTATTTACAGAATTAAAAAAAATCGAAAACCAAATTACTTAGACAAGCATCCAACACAACTACCTCTGCATTCAGATTGAAGATCC
TACCACCACTACCACTACCTTTACCTCCGCAAGAACCATGTATTTACAGAATTAAAAAAGTACTATTTTCAAATATTACTACCTCTGCATTCAGATTGAAGATCCTACCACCACTACCACT
ACCTTTACCTCCGCAAGAACCATGTATTTACAGAATTAAAAAAAATCGAAAACCAAATTACTTAGACAAGCATCCAACACAACTTAAACAGGATCTAGCATCAAAATCAAATTATAGATCA
AAATCACTTCAGTTCATGTTAATAATAAAGTTAAAATCATATCTTTATACCAAATTACTTAGACAAGCATCCAACACAACTTAAACAGGATCTAGCATCAAAATCAAATTATAGATCAAAA
TCACTTCAGTTCATGTTAATAATAAAGTTAAAATCATATCTTTATACACAATCAAAGCAAATCAGATCTCCTTTTCTTCACTAAATCTCAATCAAATCCGCACAACAAATCTTTACAAATA
CTCAACTGACACTTTACACAATAGGACAGAAGTAGAGTGTTGCAAAAGCCCAGTATCTTGCGGTTATCTTGCATACAAGAGGAATAGATAAATTGGTGGATTCCAGATGAACCACGATGAA
AATTATATTTCTGCAACAAATGCCTTGTGGATATCGTTTACGGTGATCAGATGGAAGAACACACATTTTCTCACTTTTGTGCGGAATTAAGGGGATGAAATTAACCTTACTATTTCGCTTG
GATATATATTTGTTAAACATGTATTTTCTTCGTCGTACCGATTAAAAAAAAAAAAGATTTGGTCGGCGATAAAAAAAAATAAAAATATTTAGGCAACGAGGCAATCTCTAGTACCTATAAT
GGACTATCTTGAAGCACAATGTCATATCTAAATATCTAGTGGGTGAATAAGATATTCAATAAAATCTAACATTTGTATTTATTTGTTCACATCAACAAAATGATAGATAGATTTCGCTTAA
GCACCGAGAAATGCATCGGGTAGAGATAAAGAGAAATGGACCCATTTCCACTGCTCTAATCCCAATTTTGAGAAGTTGATGGAGGCCGTTGGACAGGGACGGTATTAGAGGTTGCTTACCC
AGGCTGAAGCCCACCCCAAAATTTCTAAAAGTAAAAGTGAAAACTACGGCCACTCCCATTATTCAGAGATTTAGTTTAGGAAAAAATAGTTTTTGCTTTTCTGGTTCCCCAAACTGACCTT
CCCGATTACTCTTCTCTTTTTAGTTTTTTCCTCGTACAATTTTTTCAGAGAGTGAGAGAGTCAAGATAATATTGTAGGGTTATCTGAAGATCAGTAACCGAGAGATCAGATCTTTCTAAA
ATTTTTTTTCATCGTTTCAAAATAATTTTCTATCTTGTTATATTTTTCCTTTCTGTTATCGATTTGGTTTTGGTGTTTTTCAATTTTAACTTTTTTTTAAGGTTTTTTTTTTATGCGATCT
GTTAAGGTATTCGAAATTCTTTTATAAATTGACAGAGAAAATGTTCAACAATCATTAATTAAGTGAATCCAATTCTGTAGTTGGTTTTCGTTGCCGGGGCCTAATCATGGAGCAGGCGGAG
AAGGAGTCCTAACCTTCTAAAATGGAAAACGAAAAGGTATTTTAGTTTTCTTTTTTCAAGCTTTGATATTGGTTTTAACTTCTTAGTTTAAATAAATATTAGGTAAATTTCTGATTTTAGTA
CTTATTAAGAAGCATTTTATACTAATTGGTCTTTAATTTGATGTCAATGTATAAGTAAGTTTGTGAAATTTCACAAATTTAAGCTTTTGAGTGGGATTTGTGGATTTTTCTTTTTAATTTG
GTGTGGTTTAAGTTGAATATGTTTTTATGCAGTAGTACTCCTATTATGTGTGCTTTTGGCCCTTCAAATTCGTGATGCGAATTAGTTAATTAGACCATAAAATGATATTACTATACAAGGC
CTGTCTGTTACTGTAATATTAGCTTTACTGAGATGAAACTCTTGCGGATATTGAGTTTGTGATTCTGCAATTTTGCTTAAAAGAGTATGACTCTTTGTACATCTAGTGCATTAGTTGGTAA
GTCATAGTCTAAATCAAAAAAAGAAAAAGCAAGTTCATGTGATCTTGTTAATTGACAACTTGTAATAGTCAAAGGATCACAATCGAAAAGAGTTCTGGCTACTGTACGAAACAAGAAAAGG
AATGAGTTGAATGCATGTGTTCAATATGGATTGACCACTTGTACCTATCCTATAGATGGTTTGCCTATTTCGCAATGGATTTTTCTGACACTACTATTAGATTCTTGTTTTTTTTTTTATT
TTACTATTCATATAAATGCTCTTAAAAAAATGGTGCCTGTTTAGGCTGGTATACTTTAAGGTCGAAGAAGAAAGATGGTCGAGCGGGTCAGTATACATTCTTGAGCTATTGAGCTGAACTT
GTTTTTGCGCTAATCATCGTAAATCGGGCAATAAACAAATCACGACTAAACAATAAAAGGTCGATGGTATAAACTATAAACCGATGAGTTGAATTCTTGTTTGTAAGTTCTTAAAATTTAA
TTTCTATAACCTGTGCTTGGAACCTGTACATTGCAAAACTCATTATGATAGACAGGAGCATATGATAAGGCGAGTTATTTATCATTCTCCTACATAATTGACTACACACTTGAATGTATGA
TCTAATGGTCACTTTAATGTGGTGTGTTACCGCTCATGTGTAAACTTTCTTTTATTGAATGCTTAATGTAATTGTGTCGACGCTGACCATGTAATTTTACTAGTTGTAATTTCAGTTTTTT
CTTACGAAAACAGTGTTGTTAGAAGTTGCTGGTTAATTAGTTTATAGCCTGGGCAAGTATGTTCCTCCACATAAGAGAAATTTGAAAACTATGCGCCAAAAACTCATAACCTCATCGTAAG
TGAGTCATCCACCTGTTAAACTCTTTATGGACTTTTATGGCTGCTAATATGTTTTATTGGTGATTAGTGAAATGTATTTATAATGCGCAGGTTGGAATATTCTTTCTGAGTTTTGGTTTAG
TGGTCGGACAGAGGTGCTTGCTCATATTGATTTCGTCTGTTATTGATGTTATGTTGGCTATCTTCAGATTGTTCAAGCTGGTGATTAATTAGTGTTTGGTTCAACTGCTGTTGTCTTGCTT
GAGTATGCTAGTTGATGCTTCACTTCGGCGAGATGAGTGTTTGATCTTTTTTTTTGTTCATTTTTATTTTTATTTATCATGTATTGATGGTGAACAGGTTATGCAATAAACTGAAAGGGTG
TATAACAACTTATACATCCATCTTTAAAGGGTGCATAACCTGTTATGCAGATTAAAGTATCTGCATAACTTGTTATGCAACAAACTGAAAAGGCGCATAACAGGTTATGTATCCATCTTTA
AAGGGTGCATAACATGTTATGCATACCAAATTAAATGCATAACTAGTTATGCGGCTACATAACCGTGATAAATAAATATATATTTCTTTACTAAGTTATACAGCCATTTTGACAGTGGCGG
TGGTGCTGGTGGGGGTAGTAGAGGTGGGGGTGTAGATAATGGTGGAGGTGCTGGCGGTGAAGCTGGGAGAGGTCAAAAGGCAACAAGAGGTGATAAGAACGGTCGAGGAATCCGCCAGTTT
AGTATGAAAGGTGAGTCTTATTTGCTGATCATGGGGCTGAAAAAATTGATTGTGAAATGAAAATCTTCATGAGTTTTATGTTAAACTTTGCTGGGATTAGCAACGAAACAAAGATGGATGG
ATTTTTTAACTCGCCAATATAAATTTGAGATTAAGATTATGCCCCCTTGCTTATTGTAATATTATTGTTAGTCAGTAGGCTTTATTATATTTGTAACTTGTGTATTATGTTCAAGAATAAA
TGGTTTGTTGTTGGAAAACAAAATTGGACTTTTTTTTGAAACTATAAGTGTAATGTGTAACTGTATAATAAGATATGCAACTAGAGAAATGGCCGCATAACTTGTTACGCAGTCATGAAAA
TGGATGCATAATGCATCATGCCGTCATTTTTTTCTTCATTGGCACTGAAAAACGACTGTATAATCTGTTATGCAGCTATAAAAGTGTATGCATAACCTGTTATGCAGTGAAAAATGGTTGC
ATGATGCATTATGCATAACCTGGATGTATAATCTATATAGAGCTACAAATATGGCTGCATAACCTGTTATGCATCCGAGAAAATAGATGCATAACTTGTTATGCATTCGAGAAAATGGCAG
CATAACCCATTTATGCAACAACGTTTTGTTAGCATTAAAATTGATCAGAAAACGGCTGCATAACATGTTACGCAACTTCAAATTTTGTTGCATAACTTATTATATAATCCAGAAAATGAT
CGCATACTTTATTGTTCATTTTTTTTTTGGTGAAAATGTTACTTTTTAGGTTTATACCGCGTTTTAGCGGGTGCAAAATTAAATAAATGGATGCATGCACCAAAATATGGTTGCACAATTT
GTTATGCATCCTTTTTGGTGGTTGCATAATGTGTCATGCATCATCTTTCTTTCATTATTTTTAGGCATATCATTCGTTTTAGTGGTTATATAACCAACTAGTGGATGCATAAACCAAAAAT
ACGGTTGCATGATTTGTTATGCATCCTTTGTGGTCTTTGCATAATGTGTTATGCATCCTTTATGGTGGCTGCATAATGGTAATCTATCAAGTTTCTGAAATTTCTCCCTAAAATAACCGTT
TTTTTCGTAAAAAATTTAAATTTGATATTGTTGTTTGTACTCGTTTTGTGGAGTGCTTTAAAAGCTTTCCAACGAGATAAAATTTGTAAAATTCTAAGGCACGGTTTTTTAGATGTATTAT
ATTCAAGTTTGCATACCAATTATACCCCTGAAAAAATAGGACGCATAACTCGTCATGCAGATTTTTTACTGAACAGATAGAAAATCCAAAATTTCTTCCTTTTACAAGTAATATATTAAGG
GTAACTTTGTCTTGTTACTATCTATTTATAATTTTGAATACGGTCCGCGAGTTAAAACCCCAAAGATGTCACTATCCATACACAAAAACTCCAACAAAACAAATTGTTCACAAAAACCCCA
TTTAATATAATTGGTTTATATTACCCTCTTAGTTTAAATCACCCCAACAAAAACAAAAATCAACCCACTTTAATTCTTATTATATTGTCACCACCAACAAGCAACGCCACCTTCTACCACC
GCCACCGACCACCACCGTCGCTGCCGCCACCGACCACCACCGCCGCCACCACCTCCGACAGCCGACAAAAACAAAAACCAACCCACCCGAATCCGAACGACACCAACCACCACCAACAAAC
AACCACAACTGAACTACCACCGCCACCGACCACCACCGCCGCCGCCCCCAACCACCACCACCACCACCACCACCGCCGACACCCACCACCACCTCCGACAGCCGCCGCCACCTCCGACCGC
TGCCGCCGCCTCCGACCGCCCACCAACCACCACCGCCGCCGCCACCGACCACCACCGCCCACCTCCGACTGTCACCGCCGTCGCCGACCACCGCCGCCCACCGACCACCACCGCCACCACCGAC
CAAAATTTAGATGAAACCGATTTTTGTTTCATCATAAATCCGATGAAACCGATTTTGGTTTCATCATAAAATAAAATGAAACCCTAATTGGTTTCATCAGAATTCAACTACAAGAAAAATT
CAACAAGGATCTTATGTTGCAGTTCTCTTGGATATAGCAACATCACAACAACACAGATCCATGCTCTATGCTACAACACATATTATCAACTTCCAGCACCACCATTGTCGACCACCGTCGA
CCACCACCGACTAAAACCAGCAACATCGCCAACACCCGCCGACGACCACCACCACCACTGACCAAAAATTAGATGAAACCAATTTTGGTTTCATCATAAATACGATGAAACCGATTATGGT
TTCATCATAAATAAGATGAAACCATAATTGGTTTCATCAAAATTCATCAGTAGCTGCACTTCATTTAAGCTTCATTTCTTCTCTAGTTTACAGCATCATCTGCTATAAATCAGCTCTACTT
TCATCCTCCGTCTATCCACAACAGCACCATCTTCTTCATTGACCATCTTCTTCTCATCCTTGTAGCCCTAGCCAAGACAGCAACAACACAGCAACACAGATCCATGCTCTATGCTATATCA
AAAATCAATATGTTACATATTTAGCTCCTAAATCATCAGAAAAACGTTTGTTTGAAACATGAACATGATTTTTACAAAACAGAAAAACGTTTCTGAAAAATTCTGACGAGCAGAAAATTGT
TGTTAACTAAATTAATTTTTTGTGAGGGTAACCATGAGTTTTTTGAAACGCTGATTCAAACGAAGTTTGTATATATAGATGTTATCTTCAAAACTCATATTATACTTTCTGATTTGGAATTG
TGATTTGTTAATAAAGGTGTCATCTCCGAGGGACATGGCTAATAGCCGCATGTGGTTGGAAACAAATCTTCCAAAGATGGCAAAGGTGAGAACATCGAACGCAACTATAAGCATGGTCTTC
ATGATAAAGGTATATTTCGTAAACCTGAATCCAGAATTACTTTAGTTAAGGGTGAGTGTTATGTTTGTAAAATTCCTGGCCACGCGGCGAGTAAATTGTAGACAACATAAAAACCTTAATAA
GAAGAAAGTTAATGCTAATTTAGTTGAAACAAACTAGAACGAGTTTGGTGGCGTGATGTCGGAAGTTATTTTAATAACCAATGTGAGAGACCAGTAGGTGGACTCTGGAGCCACCAAGAAT
GTTTGTTGAAACAGAGACCTGTTCACCTCCTATCATAGGATAGGGGATGTCGAGAAACTCTTATTGAGTAACTTATATACAATAGAGGTTGCATAAAAGGAAAAGGTCGAGCAGAAGCTCA
TATCTGTAATATTCTCACATTGAATGAAGTTTTCATGTTCCGAGCATATGCGAAAATCTTGTATCTTGTTCTGTTGTAGATGGAAAAATATTTAAGATCTTAATTGAATCTGGAAAACTTG
TTGTAACTAGGCAGTGATTTTTAAGCAAGAGTTATAGGACTTTGGGTCTATATAAGCTTAACGGAAAAACTGATGATGTGAACGTAATTGATTCTTGTGATATCTTTGTGTGATTGATTG
TTTTACGTGGTAGACTTGGAACCGTAAACTTATAAGTCAATGCTTAACTGGCTAGCATAGGCTTCGTACCCAAATTTAGTTTGGATTTTGAACACAAAAGTGAAATCTGTGAACACAAAAT
TTGCATTACGACCTGTATAATTACGTGTTGTTTGGTTAGTTTGATTGCGTGCTTTGTCAATCGAGCGAAAAATAAATTCATGCTTTGTGCGTAATCTGTGGTTTTCTATTATGTAATGTCA
AGTATGAGTATTATATATATGTTTCATGCATCTGTGTTCTTCTTTTTTATGTGATTTATTCATCTTTTCCTCTTACCTGATTTATTCTATTATTTCCCTGTGATTCCTCTTACTTATGAGGTC
TTAAACAT

>PS_89C05_000087 (11.5 kb)

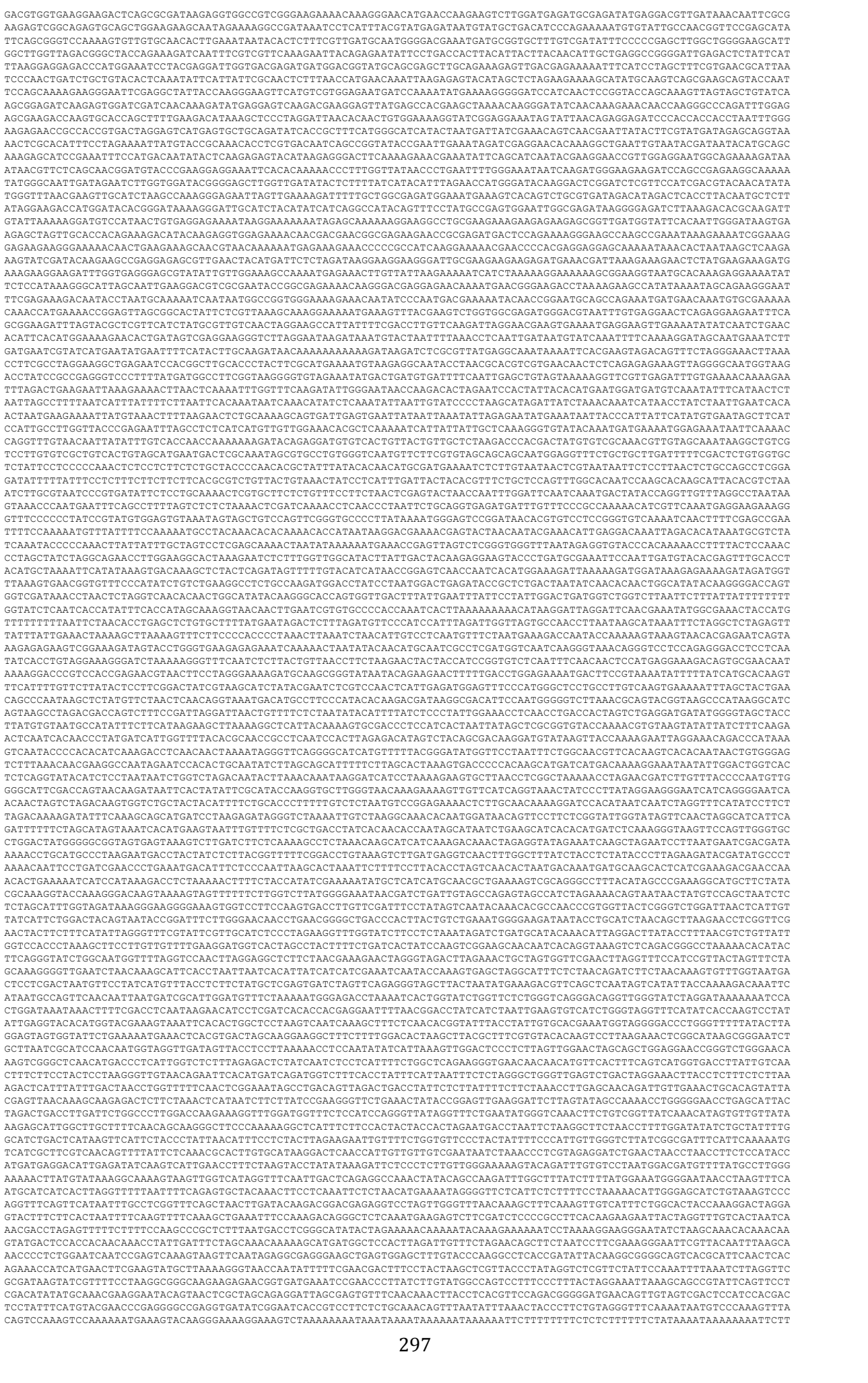
ACTTGATAAACAAATCTGTCTCCCACAGAAATACCTATGAGTTTTGTTTCGTCTTATGATAAATCAAGGTGCACATGAACCAATTGATAAACCAGACTTATATTCCCGAAAAACATCCAAG
TATTATCAATCATCTCACAATAATGTTAATCGTATGGAAGAGAAACAAGATATTGTGGAATCATAAACGATGAGACGAAGATGTTCGTGACTTATTTTAATCTTACCCATTGGAGATAAAT
CTCGAGAAAATCTTAGAGAAGATAGTACTCAATCACGATAGTATAAAGTAAGATTAGAACACGCAACTACAAAAAAAAATAGCTGGGTCTGGCTTCCACAATCCCAATGAAGTCTTTAAGTC
GTTAACCAACAGGGTTTTGGAATAAACCTAAGGTTGAAGGAGAATCGACTCTAGTCGCAACTAGTATCACACAAGAGGTGTGTGGATTAGGTTTCCCAGTTGCTAGAGTTCTCCCTTATAT
AGTCTTCAAATCAGGGTTTGATAGCTTTGGAACAAAGCAAACAATATTCACCGTTAGATGAACTCCTGATTAGATTCATGCTAATATCTTTCAACTTTTAGATCGAACTTAGCTTGTTACA
CACAAATAAAATGTACCTTCAGTTAGATATGGGTAACCGTACCTAAATGTACCTAAACATGTAAACCAAGTTGGCTCAACAATAGTTCAACTGTTAGATCGAACTTAGCTTGTTACACACA
AATAAAATGTACCTTCAGTTAGATATGGGTAACCGTACCTAAACATGTAAACCAAGTTGGCTCAACAATAGTTAACTGAAGTTATCCATATGAACATGATCCATGCTAATATCTTTCAACT
TTTAGATCGAACTTAGCTTGTTACACACAAATAAAATGTACCTTCAGTTAGATATGGGTAACCGTACCTATATGTGTACACCAAGTTGGCTCAACAATAGTTAACTGAAGTTATCCATATG
AACACTTTCATATCAATCTAATTCATCTTAACCAAAACTAGTTCATATGACTCAAATGAAACTAGTTATAGAGTTGTTCAATTGTTTATATTCTCATAGACGTATACAAGACACAATTGAA
GCAAAATTGGTTTAATTAACTCGAATCAATTCATGAACATTATAGCCACGGTTTGCAAGAGATTGCATTCTTTCCCTACACCTAGTTATAGAGTTGTTCAATTGTTTATATTCTCATAGAC
GTATACAAGACACAATTGAAGCAAAATTGGTTTAATTAACTCGAATCAATTCATGAACATTATAGCCACGGTTTGCAAGAGATTGCATTCCTTATTATATAAATGTATTAGTTCATGGCAC
AACCAATTTTAGAACATAACCCGCTCAAGTAGGCAAACGGGTACGCCAACCTAAGTAGCCGGACTTGGACTGTGTTCGATAGTACGCAAACGGGTATGCATACTGTCACACATCCAAACTT
CAGTTGAACCCAGTACGCATACAGGTACGCATACTATAGTACCCGTACTTCAACTGACTTCGCCAGTACGTGTACATGTAAGCATACTCTAGATCCCGGACTTTACCTGACCTTGCCAGTA
CACATACGTGTATGCATACCACACCGGTCTTGGATCCACATATATGCAAGTACACATACTATGTTTATAATCCAATTATGGTTAAGTGTTCTAAACTCCATTTCAATCATTGAAACATTCT
TAGAAGACGACAATAGCTATCTCACACAGACTATTAGCTTCAAAGAAATTTCAAGTGATCGAATGATCAATACGAAACATTCCGAGTCTACATCAAATGACTGTCTCATACAAATCATAT
AAGATGTTACCAGATGGCTTTCACAAGATCATCTTTTGACTTTCGTCAAGAATATAAAATGAACTTGGTTAAAGTGAAAGCTTACCAACATATATTTCGAGAAATATGTAAGCGAGTTAAA
CTCAGCTCGAAATATCAAATGTGTATAAATCGAATATTATATAGCTATACGACTTTTGTCTCAAATAGGAGATACAGTAGAAACAGACTTTTTGAGTGATATATTAGTTTCAGTCTCCACA
TACCTTTTGTTGATGAAGTTCCACAAGCTCCCCTTAGTAGTTATTCGTTTTCAATTGATGAATCCCGTGAATTCTAAAGATCAACTACACATTCTATCCTAGTCCGAGACATAGCTATAAG
TAGACTAGAAATCAAGACTTATAGTTTTGATCACTAAACTTGACAAAAAAGCTTGAGATAACAACGATTGCAAGTTTGACTGAGCAGTGCTCTAATAATTTCTGGATCAAAAAGTGGAAAA
GGATCGAAGACCCTCCTTTAATTGTCCAAGCGCGATAGGCATGCGAGTCGCGCTCCATGGTTCTCTTTGAATCCTCTCCTCGTCTAACAGGGTAGTGGCAGAAGAATCCTCTGGAACTGAG
AGAATGAAACCACTCTTTTTAAGTTCTACTTGAAAAGTTTCAAAACTTTCTGCGGTCTTCTTACCGGATTTTGGTTTCGAACTGCATGATTAGTATGGGAATAAACATTCGGAATACTGAAA
TTGAATAACATAACAAAGCTTGGAATACGATAACAATGGAAAAGGAAAAATCACAGAAATTATAGTCATGTCGTAAGAAACGCAAAATGCAAATAACTGAGGCAAAATTAACTCACCGATA
AGATGATGTAGCAAAAACAACAGTGGGAGCAGACGTATTTTGATCTGCCATGGAAACATAGGCGAAGTTGAAGAACACGGAGGAAGAAGACAAAGTACGGAAGAAGAAGAAGTTGAACAGA
GGAGGAAGACGGTAGAAAATAAAATATTTTTTCTCTCTCCTGAGTATTTATCCGATAAAAAGAAGTAAAGACGTGTCGCTTGAAAAGGAAAAACTTCAACACGTGTAGCAGTTACGCGG
AAAATACATAAGTGTGTCGGGTAATACAAAGAAGATGAAAAGACAGGACTGATGCGAAGTTTGACCAAGAATTCTCGCATTATTCTACTTGCATAAGAACAACGCGACAAGAAGCAAAATG
TAGGGTCTAAATATTGCACAAGGCGAACACACAACTTCGCTCTAGATAAAACAACTCCAAGCACGAAGATATTATAATAATTAGAAGACAGGGCGAAGAATAGTCGCGCGACAATAAGACA
AGATTGCGAGAAGATAATCAAGTGATAAAGATAGCACGGAGACCATATTCAAAGGAGTACCATGAAAGATACAACAGACAACACGATGCACATTCATACCTTCTCAAAAGTCTGGCGCATA
AGACAAACTGAAATAGAATCAGTCAACTGAGATCTTGGACGTGAATCATGATTGTCTTCTATTGATAAGTTTGCCTGTATAAGGACTTAGAAATCTAACAGAAAGAGAGGGGTCGATTAAG
AGAGAGATAGCTAAGAAAATTATAAGAAATAAAGAAATGAATTTCATCTTTAAATTTGGCTTTCATCTTGTATTCCTCAATTGTTGTATGTGAAATCTACCTTGAATCTACAGTAAGAACA

ATACATATTCATTATCAAGAGATTGAGAAGATCATATAAGTGTCAAAGTGCTGAATAATATCATTAATGTTTGAGTTGATCAAAATGACTTAGACTTAATGTTATTATCAATACAAGGGTC
TAGTTGTGGGATTTTCCGCAACTACATAAATGTTTCACTGAAATGATGAAAAAATTCGTGTTCAGATTTTTTTTTTCCTTTATAACTTAACTGAAAATAGAGAAGAGATAAGGGGAACAGA
AGCCGATATGTTTTGCTGGAAAAAAAAAATGAGTCATGTGAACCTTGAGTTTTTATTTTTATTTTTTAGTAAAGTTTTGTCACTTGTACACGGTGGACTCGCCGCATCAGCGCGAGAGG
AGAATAGAACTGCGAGAAGAGTGGAGAATATCATTACCGAAATTATACACGCAAAAGTATCATAAAGTTGAGACGAAATGAAAATTCTTTCTGTATGAATCTTTAATTTAGTCGACAACTC
AAAATCCAGACAGTGTCGTGTACTGTGCTACTGTGCCAATTAGGTTAAGCCAGCACAACCCACGAAATTAACATGTTGGGTCGTGCTGGACTAAATCACGTGCTGATATGGCCTCAGATTT
TAGCGTGTTGTGCTAGATTGTGCTGGTGCTGGTACAACCCAATTTACACCTCTAAGCACCACCACCAAGCCTTCCTTTAGTAGCTGTTTTAGAAGCAATGGGATCAAGATTAGAGCGATAA
ACACAACAAGTAGCATGAGAAACCAACCAACGTACCCTAGCCTAACAGAGACTTAACAGAAAATGAGAAGACAATCTTCAATCCAATACATTATTTATTACATAATATTAAGCACCATATG
AATTCTTCAAAAGCAACTTATTTCAACCGTGATTACATCACTTGACCCAAACAGAGTCTTACATATAGTGAGACGAGAATCTTCAAACCAATACATTATTATTACATAATAAATCTCTCAG
TCATCCACGAAATTTAAGCACAATATGAATTCTTCAAAACCAACTTATTTCATCATACTTCAAAAGCAACTTATTTCATCATATATTAAACACAATGTGGAGTATAATAATGTCACACT
TTCTCACATCCTCATGTAGTCGAGAAATGATTTTCCATCAAGCTTCCTTGAAAGATTTTCCTTCAAAATATCCTCATACCTACCTCTTTTGAACAAAGCAGGTGTCTCTGGTGTGACCAAG
CTCGAAATTGGGCCTATTTCTGACTCTAGTTTAGAGTCATGAAATGTCGCGATTGAGAGCCTCTCCTTTGTTGAGTTTACTACTGCCCGGTGCTCGACGCTACGGTAAATCCCATTAGTCA
TTATCTGTAAAAGGGAAACATTAAGTAACTAAGCAACTTTTCGGGGTACAGTCAAAACAGGTACTCACAACAAGGACCAGTAGGCAGGCCATATATGCCCTTAGCCTAAATACCTCCAAAA
TGTCTCCAACATTCACTATGAACGCATCAGGTAGAGGTTTGATTGAAATCCACCTCTCTTCTTTTCTTATTTGTAATCCTTCAACTTCATTAAGTTGAAGGAGAATTGTCAAACCGCTAAA
ATCCGAGTGTGACGTAAGACCAAGTACAAGCTCTGGTCGAGGACAAGGAGGATAATAGTTCATCCTCATTGTTTGCAACCCATCTTCAAATAAGTCTGTCATACCTTTAATCTCAACTAAT
TGTAGAGATTTTTCCAACATCTCAAAGACAACCGTTGATAGTTTTTTCATTTTTGATAGGTAGGATTCCAGTGTCTCCCTGCATAAAGGATGTAAATTAGGGATCAAATGCAACTTCAGTA
TATTTTCCATCTGACAGATTTTAGGGATGCATGATGTTTGGATTCGATATATTTTGGAAACCATGCAATTCACAAACCCGGATTGCTTGATACTGCCTCCGAATGTGAATTTTAGGCAACC
TAGACAGGATTTTCCATGGTTTATCCTAGCAACCTAGACGATATTGATGAAACTATTCTGAACCAACCAAAATAACATTAAAAAAAAAAAAATATCACGTGCGGCATTGAGAAAAAGCAAAC
GGAGAGAAACAATAACCTGAAAGGCAGAGGGAGTTCTGGAAACAAATGAGGCTTCCTTAAATGGAGAGGAAGACTTAACATGCTAAACACTTCAGTCCAATCAAGTCTTTGGTCCTCCGAT
TCAATATAGGGTTGTCCAAATCCTTCAAAATCTCCATCTTGCTGTCCGTATTTAGTTTTCTCATTCATTGGAAGGTTAAAGAAACCTTTAATTTCTGATTTTATATTGTCCATCAGTAAAG
CGTCGACTCCATGGTTAACCAGCTGAAACATAAAGCAGATTAGCATGAACTAGTAAAACACTAAATAGTCGTCAATCAACTACGTATAATCAAAATTTGTATACCTGAAAGAAACCCCATT
CTTTGCAAGCAGAATGAAGCTTATCCAATTCTAACTTTCCAACTACGGGTTCTGGAGATAGTAAATTTTGCAAATCGATGACAGGAACTGTTTCATCATCTGTTACAGACGCACCAATATT
ATTCAACGGGCTTTCACCGGTCGTATGTATCGAGATGGAATTTCTGCAAGCGTGAGTTTAGCCAATTCCTGAACACTTGGTATTGACAAACCATTGCCTAGCTTGATAAGTATTGGTGTC
TCCATGAACTCTATATTTGATGAACTTTCTTGTCAGATCTTCAGATCATATATCAATCTTAAGTTAGAATTGGATAAGCTTCATTCTGCTTGTAGCCAATTCCTGAACACTTGGTATTGAC
AAACCATTGCCTAGCTTGATAAGTATTGGTGTCTCCATGAACTCTATATTTGATGAACTTTCTTGTCAGATCTTCAGATCATATATCAATCTTTCAGGTCTCTTTGTTTCTCTCAAGTGGA
TATGCCTTACTTATATCTATATTTCACTCGATGTCTAATGAATAAATGACTAGTGTTTTTTCAACGAAGGGTCTGAGCAACCACTGCTTACCCTACCAGTTTTAACCTGATGATGTCTACG
ATCTAACCTGCCGGTCTAAGTTGCAAGATAATCTAACAACAACGAATATGGTGAAGTTTGGTTGTTGTTTAGGTGATTAGGACTGATTTTGTTGTTCTTCTCACTTTCTATATCCTTCTT
GTTGCAGAACTTGGTACGATCTAAGGTATTAGATTGGGTCCATGCCTACATGAGCAGGAGGACGTGGGCAAATGGTGTACGATCTAAGGTATTTTCCCAAGCTAATGTAGCGACTGAGGAA
CGAGAACCACTACTTCAGTTCAGATCAATAAAAAAAAATAAGACCTTTCTTCACCAAATCTCAATCAAATCCACACAAGCAACTGCTTTTTAGTACTTTCTATAGGCCATATCTCCATACCT
AAAGTTCCTAGAAACAAACATAAGATTGAATTTATGATAAGATGGAAATATGATAACTGACAGTTAGTGATAGAAGAACGGTGAAAAAATTCACCTAGTAAGAATCGGTTGTTGATCATAG
AAAATAAATAAAGTGAAGATCTCGATTAGTAATTTAGTACTATTTTCAAATATTACTACCTCTGCATTCAGATTGAAGATCCTACCACCACTACCACTACCTTTACCTCCGCAAGAACCAT
GTATTTACAGAATTAAAAAAAATCGAAAACCAAATTACTTAGACAAGCATCCAACACAACTTAAACAGGATCTAGCATCAAAATCAAATTATAGATCAAAATCACTTCAGTTCATGTTAAT
AATAAAGTTAAAATCATATCTTTATACACAATCAAAGCAAATCAGATCTCCTTTTCTTCACTAAATCTCAATCAAATCCGCACAACAAATCTTTACAAATACTCAACTGACACTTTACACA
ATAGGACAGAAGTAGAGTGTTGCAAAAGCCCAGTATCTTGCGGTTATCTTGCATACAAGAGGAATAGATAAATTGGTGGATTCCAGATGAACCACGATGAAAATTATATTTCTGCAACAAA
TGCCTTGTGGATATCGTTTACGGTGATCAGATGGAAGAACACACATTTTCTCACTTTTGTGCGGAATTAAGGGGATGAAATTAACCTTACTATTTCGCTTGGATATATATTTGTTAAACAT
GTATTTTCTTCGTCGTTACCGATTAAAAAAAAAAAGATTTGGTCGGCGATAAAAAAAAATAAAAATATTTAGGCAACGAGGCAATCTCTAGTACCTATAATGGACTATCTTGAAGCACAAT
GTCATATCTAAATATCTAGTGGGTGAATAAGATATTCAATAAAATCTAACATTTGTATTTATTTGTTCACATCAACAAAATGCTAGATAGATTTCGCTTAAGCACCGAGAAATGCATCGGG
TAGAGATAAAGAGAAATGGACCCATTTCCACTGCTCTAATCCCAATTTTGAGAAGTTGATGGAGGCCGTTGGACAGGGACGGTATTAGAGGTTGCTTACCCAGGCTGAAGCCCACCCCAAA
ATTTCTAAAAGTAAAAGTGAAAACTACGGCCACTCCCATTATTCAGAGATTTAGTTTAGGAAAAAAATAGTTTTTGCTTTTCTGGTTCCCCAAACTGACCTTCCCGATTACTCTTCTCTTTT
TAGTTTTTTTCCTCGTACAATTTTTTCAGAGAGTGAGAGAGTCAAGATAATATTGTAGGGTTATCTGAAGATCAGTAACCGAGAGATCAGATCTTTCTAAATTTTTTTTCATCGTTTCAAA
ATAATTTTCTATCTTGTTATATTTTTCCTTTCTGTTATCGATTTGGTTTTGGTGTTTTTCAATTTTAACTTTTTTTTAAGGTTTTTTTTTTATGCGATCTGTTAAGGTATTCGAAATTCTT
TTATAAATTGACAGAGAAAATGTTCAACAATCATTAATTAAGTGAATCCAATTCTGTAGTTGGTTTCGTTGCCGGGGCCTAATCATGGAGCAGGCGGAGAAGGAGTCCTAACCTTCTAAA
ATGGAAAACGAAAAGGTATTTTAGTTTTCTTTTTTCAAGCTTTGATATTGGTTTTAACTTCTTAGTTTAAATAAATATTAGGTAAATTTCTGATTTAGTACTTATTAAGAAGCATTTTATA
CTAATTGGTCTTTAATTTGATGTCAATGTATAAGTAAGTTTGTGAAATTTCACAAATTTAAGCTTTTGAGTGGGATTTGTGGATTTTTCTTTTTAATTTGGTGTGGTTTAAGTTGAATATG
TTTTTATGCAGTAGTACTCCTATTATGTGTCTTTTGGCCCTTCAAATTCGTGATGCGAATTAGTTAATTAGACCATAAAATGATATTACTATACAAGGCCTGTCTGTTACTGTAATATTA
GCTTTACTGAGATGAAACTCTTGCGGATATTGAGTTTGTGATTCTGCAATTTTGCTTAAAAGAGTATGACTCTTTGTACATCTAGTGCATTAGTTGGTAAGTCATAGTCTAAATCAAAAAA
AGAAAAAGCAAGTTCATGTGATCTTGTTAATTGACAACTTGTAATAGTCAAAGGATCACAATCGAAAAGAGTTCTGGCTACTGTACGAAACAAGAAAAGGAATGAGTTGAATGCATGTGTT
CAATATGGATTGACCACTTGTACCTATCCTATAGATGGTTTGCCTATTTCGCAATGGATTTTTCTGACACTACTATTAGATTCTTGTTTTTTTTTTTATTTTACTATTCATATAAATGCTC
TTAAAAAAATGGTGCCTGTTTAGGCTGGTATACTTTAAGGTCGAAGAAGAAAGATGGTCGAGCGGGTCAGTATACATTCTTGAGCATTTGAGCTGAACTTGTTTTTGCGCTAATCATCGTA
AATCGGGCAATAAACAAATCACGACTAAACAATAAAAGGTCGATGGTATAAACTATAAACCGATGAGTTGAATTCTTGTTTGTAAGTTCTTAAAATTTAATTTCTATAACCTGTGCTTGGA
ACCTGTACATTGCAAAACTCATTATGATAGACAGGAGCATATGATAAGGCGAGTTATTTATCATTCTCCTACATAATTGACTACACACTTGAATGTATGATCTAATGGTCACTTTAATGTG
GTGTGTTACCGCTCATGTGTAAACTTTCTTTTATTGAATGCTTAATGTAATTGTGTCGACGCTGACCATGTAATTTTACTAGTTGTAATTTCAGTTTTTCTTACGAAAACAGTGTTGTTA
GAAGTTGCTGGTTAATTAGTTTATAGCCTGGGCAAGTATGTTCCTCCACATAAGAGAAATTTGAAAACTATGCGCCAAAAACTCATAACCTCATCGTAAGTGAGTCATCCACCTGTTAAAC
TCTTTATGGACTTTTATGGCTGCTAATATGTTTTATTGGTGATTAGTGAAATGTATTTATAATGCGCAGGTTGGAATATTCTTTCTGAGTTTTGGTTTAGTGGTCGGACAGAGGTGCTTGC
TCATATTGATTTCGTCTGTTATTGATGTTATGTTGGCTATCTTCAGATTGTTCAAGCTGGTGATTAATTAGTGTTTGGTTCAACTGCTGTTGTCTTGCTTGAGTATGCTAGTTGATGCTTC
ACTTCGGCGAGATGAGTGTTTGATCTTTTTTTTTTGTTCATTTTTATTTATCATGTATTGATGGTGAACAGGTTATGCAATAAACTGAAAGGGTGTATAACAACTTATACATCCATCTTTAA
AGGGTGCATAACCTGTTATGCAGATTAAAGTATCTGCATAACTTGTTATGCAACAAACTGAAAAGGCGCATAACAGGTTATGTATCCATCTTTAAAGGGTGCATAACATGTTATGCATACC
AAATTAAATGCATAACTAGTTATGCGGCTACATAACCGTGATAAATAAATATATTTCTTTACTAAGTTATACAGCCATTTTGACAGTGGCGGTGGTGCTGGTGGGGGTAGTAGAGGTGG
GGGTGTAGATAATGGTGGAGGTGCTGGCGGTGAAGCTGGGAGAGGTCAAAAGGCAACAAGAGGTGATAAGAACGGTCGAGGGGGAGTTCAGACGTGTAGCTCTTTCTTTACTAAGTTATAC
AGCCATTTTGACAGTGGCGGTGGTGCTGGTGGGGGTAGTAGAGGTGGGGGTGTAGATAATGGTGGAGGTGCTGGCGGTGAAGCTGGGAGAGGTCAAAAGGCAACAAGAGGTGATAAGAACG
GTCGAGGAATCCGCCAGTTTAGTATGAAAGGTGAGTCTTATTTGCTGATCATGGGGCTAGTAGAAGTGGGGGTGTAGATAATGGTGGAGGTGCTGGCGGTGAAGCTGGGAGAGGTCAAAAG
GCAACAAGAGGTGATAAGAACGGTCGAGGAATCCGCCAGTTTAGTATGAAAGGTGAGTCTTATTTGCTGATCATGGGGCTGAAAAAATTGATTGTGAAATGAAAATCTTCATGAGTTTTAT
GTTAAACTTTGCTGGGATTAGCAACGAAACAAAGATGGATGGATTTTTTAACTCGCCAATATAAATTTGAGATTAAGATTATGCCCCCTTGCTTATTGTAATATTATTGTTAGTCAGTAGG
CTTTATTATATTTGTAACTTGTGTATTATGTTCAAGAATAAATGGTTTGTTGTTGGAAAACAAAATTGGACTTTTTTTTGAAACTATAAGTGTAATGTGTAACTGTATAATAAGATATGCA
ACTAGAGAAATGGCCGCATAACTTGTTACGCAGTCATGAAAGTGAGTCTATTTGCTGATCATGGGCTCATTTTTTTCTTCATTGGCACTGAAAAACGACTGTATAATCTGTTATGCAGCTAT
AAAAGTGTATGCATAACCTGTTATGCAGTGAAAAATGGTTGCATGATGCATTATGCATAACCTGGATGTATAATCTATATAGAGCTACAAATATGGCTGCATAACCTGTCATGCATCCGAG
AAAATAGATGCATAACTTGTTATGCATTCGAGAAAATGGCAGCATAACCCATTTATGCAACAACGTTTTTGTTAGCATTAAAATTGATCAGAAAACGGCTGCATAACATGTTACGCAACTT
CAAATTTTGTTGCATAACTTATTATATAATCCAGAAAATGATCGCATACTTTATTGTT

>PS_109H06_000063 (54 kb)

TCCTTACCTTTTTACTCTCAAACCAACCGCTCATCGATGCCACCTGCAACAACACTTCTTGAGCACTAACAGTAAAAGCCAAAGCTTGCTTTCATCTCCATTCGAGTACACAGCCACAAGC
TAATGGCTACAGCACCAGTTCATTCAGCCAACTCCCGTCGCCTGCAACCATCCATTCATCACCAAAGTACATCTACTGCCTGTTTGTCCGTCCAACTTCAGATCCTTCTCAGACTCATCAC
CGGCTTTACTGATTTGAACTCCTAACTTGGCTGCAAGAATATTGCAATTGAATCTGTAATGTACTGCGAAAAATCGACACCTTGAATCATCAATACCGGTTGCTCCTGTTCCATTCTGTGT
TCTCGTCATCAGTATCTCCTCTGCCACGAAACCATCAAAAGATTCTGCCATCATCCGTGTGATGGCAGTTTCGGCCATAGCACCAACCTGACTTGTAAACTTTGTCAGTCTCAGCCACACA
ACAGCTGCTCCCTTGCTGTTCATCCATCTGCCAGCAGCTCCATTCTCGATCAATTCCAACGACAACTCAAATAATACACAGACCCTAACTTCTTCTGTGAGACACTTTAGTTCCAGAAGAA
ATAATATTTTACTTGATGCTCTTGGTCAAAATTCGAGTGATAAGAAGAAAAGGCTTAGACAAATATAGCGTAGGCCACAATAGTGTGAACTCCCTTTGTCAAGAAGTGCACAAATGTGCT
CAGAACTTCAGCATCTCGTGTGTAGTGGAGTTTCACGCGCTTGCTTATACAAAGCAACAACTTCTTAGTCCCACGTTTGGTTTCTGGATATGAAATGTTATTGGTGCGGCAACCGGCTTTG
TTTACTTGGCAGTCACAGTTAATTGAGAAGTTCCCTTTTGGTGCAAGTGAGAAGCAACTTTTGTTGTTGTAGTCGCAAACGTCAATTAAGCTACACCTTAGGCATTAATTTCCATGATGAC
CGAACTCACTTGTACTTACTATAAAAGCATTAAAATAATATTGTTAGCCAAAATATCGAGTCCTATCTAATATAAAATATGAATATAAAATATAGTACTTACGTGCTTATCAACACCCCCAC
ACTTATATTTTGTTAGTCCTCGAACAAAACAGAAGAAGACCCGCTACACAGCTGGTCACATCATAAGAGAGAATTAGACCCGCTACACAGCTGGTCATAGAAAATACCCTACAAAAGACCC
GCTACACAGCTATTCATGACAAATACCATACACAAGACCCTCTACAAAGCTGGTCATAACGCTAAAAACCATGGTGATAATTAATTTTTTAAGACTCGTTACACAGCTGATCATCTTTTTT
AAATTTGTTTTTTAGACCCGCTACACAGCTGGTCATAATATACCTGTAAACCTACAATAGACCCGCTCCACGGCTAGTCATAAATTACAAGAAAATTCCTGAAAAAGAAAAATGAAAAAGT
TAACCACTCCCCCACACTTAAACATTACATTGTCCTCAATGTAATTGAATCATCCAACATAAAATTTAAGATGGAAAATTCATTAATATGCGAAAAAAAACAAATAAAATATAAAATAAGAAA
ATAGAATTACCTACTGGAATATACATAAAGGATTCTCCATATATCAACAACCTAAATAGTTGGGATCAACCCAAAGAACAATATATACAACTAGTGATAAACTCAGATCTCCGTTGTAGCA
ACGTAAAGATATAGAGATTATCAAAAGGAAAGAGCTGCCCACATACCTTGTTTTTGCAACTGATAAATCGTCAGTAACAAGAATTACCGACTCAAGTGGTTCGTTGTGGGTGAGAACTGAA
ACCATTTGGGAGACATGCACAGGAAAAGAGAAGTCACGAGATGCAAAATAATCTAGTAATACGGAGAATGACGCATTTCTAATCCAAAATTAGGAATCTTCCGAAAAGTTGTGCTAACAAT
TCTAGGTATATGAACTTCTCCACTGGGTTTGTGACCCTGTACTAAATCATCACCATAATAATCATCAACCAAGGAATCAATATGAATATCACAATAAAGATGATTAAAGGAAACCTTATAA
ATAGAGTGCAATCTCTCTAAGTCGTTGGACTTAACTACAGTTTCATCCGAATTAAAACATTCTTCATAGGCATCATTTTTATAACACTAATTTGTCACACCGTATAAAAGAATGGGTGTCC
TCAATCCTAAACTCATCCTTTTAATAGGCGAGTGATCATTAACACGATCAGGAGTTGGACTAAATACACCACCTTCACTTGGTATGACAATGTCTCTATTCCTATGATCACAACTATCAAG
TTTACCATCATTATCCATACACATTGATTCATTAGTTAAAGCGGTTTCATCATAAATATCAAGAGTATAAGGAAAATTATAACTTGTTATCCTCAGTAATACCACGATCTTTTACCGACTGC
ATTAGTTTAGCAATCGATCAATCAGTTTCTTCAGAAAGTTCAGCAATATAAGCTCGGGTTTTACTCAGCTCAACAAAAAATAACTCATTAACTTTGGTCTGTTCAGCAATTTCAATATTGA
GGCGTTCTAGCAAAGATAGTTGTGATGCCGATGGGTCACAGTAAGATTGATCATCTCTCTCATACTGTGACGGAGAAATCTGTGAATAATAATTGGAATCAACAAATTGTGGCTCATAACC
AGGGTAAGGTTGGTAATGATCATAACCTTCACAAGGTCTGGAATGATCATACCTATGCAAATAATTCCAATCATTATTTACACACTCGGCAGGTTCTTGATAATGAGAAGGTGCATATGCA
TACTCCTCAGGTGAAATATCAAATGACTCCTCTGTCTCAACTCTTTTGGTAAGATTCCAAAAAGGTATAAAAGAAGTCAACTCACTAGAAACCAATATACAAAGTAAAGTTTTCCCAACAA
CTCACCTTGCAAAAGCTTTGTGAGAATATCGAGTACTTACTTTAAGCCCGTTGTTGACTCCAATCGCCGCTCTTCAGACTAAGGTACCTAGGAAAAAATCAAAAGCTACGTAAAAAGAACT
AAAAGAAAATATAAAAATAAAAGAAAAATAAGTTATGTACAAAATACCAATTACAATATACCACAACCACTCCCCGGCAGCGGCGCCAAGAATTGATGGGCTCCGGATATGCCTATATTAT
AATAACCTCAATCGCACGGTGTCCAACTAGCAGACAGAGGCAAGTACGGGTCGTTCCCACAAGGAGCGGTGGAAATACTGTAATCAATATCTCTAATAATCAGAAATGATGTTTTTGAGAT

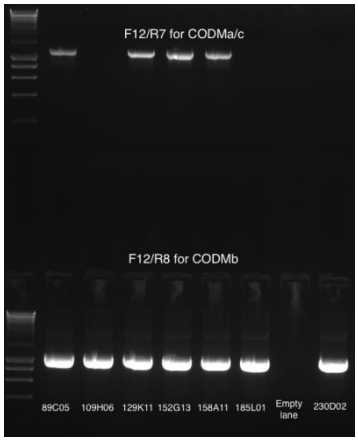
TTTTATAATTGTGAAATGTAAAATGAAATAATAAAATTAAACTAATAAAATAATATAACTGAGAAAGTTGGTAACCAGGGTTTACAGTTTCCACCACTATTTCTATTCAAAAACTGAAATG
AAACATAATTCATGAATTATTTATTGTTCGTTCTATTGACATTACCTATTATATGTATTAGAATCGCAAATATCATTAGGACGCCCTAAGAATGGCACATCAAAAGACTAAACCGAAGCAT
GCACCATCAAATTGCATTAGCGAGAAATTTATACGAAACTAATTATATGTTACAAATAATGCCTAGCTATTCTAAGTATACCCCATCAAAAGATATAACCCAGCATGGAATATCAAATAA
GTAGACAAATATATTTGTTATCCTAACAATTATGCACAAAGGTTATACAAAACCGGTGAAGCAACAACCCATATTTTGATTAATCAATAAATATGATCATGGAATTAATCAAATCAAACCA
TAATAGTCATTGCTCAATATTATTCAAATCAAAATAGCCTAAAACTCCAAGTGGTTTCAACATAAATCCTAGTTACACAAATCTAGCAATACATGGCTTTCTCAAAAGCCATGACATATAA
AAGAGAAGCACAAATAAAACAAAATTGATGCGAACAATGAAACTTCAAACCGTAATCCTTTTATCGCAGTCGTCCTCTCTTCTCGTTTCCTGGTTTTGGCTTGTAGCCGCTACCTCCAGTC
CTGTCTCTCCCTTTGTCTGTCTGTAGCTAGCTCTTTATAGAAATCCAGTTCCTCCTCTTACGATCCATGCCAGCACTCACCAAAGTCCAGCCTAATAAGCAAAATTATCCAATTATAAA
CCCAAATCCGCTTTTTATTTTCTAGCCCTTTCAAGACTCCGACCATTGGAACATCACCACTGGTTGTTATAACACCCGAAGCTACTCCAGATTGAATCGATTCATACAGCCCATATCTTCT
TCTTCCTGTTTCGACTCACGTCTTCGCATCCACTGGCTCACGGCTTAAACAGGCTCTGACATAGCAACAACTCCAGCCCCGGTTACAGCAGAACTTTCTGTTCAAATCGTCACCAATCTAC
TCGAACCAAAACCCCACACACAGAATTCATTCAACCAACACACGATCTTTTTTCATGCCTTTGTTCTGATCCTTTCCATCTCCATTTCTTCAGCTCAAGAACATCAATGAAATTCAGTTTC
ATTCCATAGTTGGCCACAGCACCACTGATCACTACTACCACGAACTCCAGCTTGAAATCCCAATCTCGCATCATGGTCTCAACTCTGCCAACATCATCAACAGCTCATATAAACCATCTAA
CTCACGTCCAAGTTCTCCATGCAACCATCCTTACCTTTTTACTCTCAAACCAACCGCTCATCGATGCCACCTGCAACAACACTTCTTGAGCACTAACAGTTAAAGCCAAAGCTTGCTTTCA
TCTCCATTCGAGTACACAGCCACAAGCTAATGGCTACAGCACCAGTTCATTCATCACCAAAGTACACTCTACTGCCTGTTTGTCCGTCCAACTTCAGCTCCTTCTCAGACTCATCACCGGCT
TTACTGATTTGAACTCCTAACTTGGCTGCAAGAATATTGCAATTGAATCTGTAATGTACTGCGAACAATCGACACCTTGAATCGTCAATACCGGGTGCTCCTGTCCCATTCTGTGTTCTCG
TCATCAGTATCTCCTCTGCCCACGAACCCATCAAAAGATTCTGCCATCATCCGTGTGATGGCAGTCTCGGCCATAGCACCAACCTGACTTGTAAACTTTGTCAGTCTCGGCCACACAGCAG
CTGCTCCCTTGCTGTTCATCCATCTCGCCAGCAGCTCCATTCTCGATCAATTCCAACGACAACTCAAATATACACAGACCCTAACTTTTCTTTGAGACACTACAGTTCCAGAAGAAATAAT
ATTTTACTTTATGTTCTTGGTCAAAATTCAAGTAATAAGAAGAAAAGGCTTAGAAAAATATAGTGCAGGCCGCAATAGTGTGAACTCCCTTTTGTCAAGAAGTGCACAAATGTGCTCAGAA
CTTCAGCATCTCGTGTGTAGTGGAGTTTCACGCGCTTGCTTGTACAAAGCAACAACTTCTTAGTCCCACCTTTGGTTTCTGCATATGAAATGGTTATTGGTGCGGCAACCGGCTTTGTTTA
CTTGGCAGTCACAGTTAATTGAGAAGTTCCCTTTTGGTGCCAGTGAGAAGCAGCTTTGGCTGTTGTAGTCGCAAACGTCAATTAAGCTACAGCTTAGGCATTAATTTCCATGATGACCGAA
TTCACTTGTACTTTCTACAAAAGCACTAAAATAGTATTATTAGCCAAAATATCGAGTCCTAACTAATATAAATATGGGTATAAAATATAGCACTTACGTGCAACACCCCCACACTTATATT
TTGCTAGTCCTCGAGCAAAATAGGAGAAGACCCGCTACACAGATAGTCACATCATAAGAGAGAATTAGACCCGCTACACACCTGGTCATAAAATGAAAATTTGAAATATCATACACAAGAC
CCGCTACACAGCTGGTCATAGAAAATACCCTACAAAAGACCCGCTACACAGCTGGTCATGACAAATACCATACACAAGATCCGCTACACAGCTGGTCATAACCCTAAAAACCGTGATGATA
ATTAATTTTTTTAATTTTTTTTAAGACCTGCTACACAGCTGATCATCTTTTTAAAAAAAAAATTTTAGACCCGCTACACAGCTGGTCATAATATACCTGCAAACCTACAATAGACCCGCTC
CACAACTGGTCATAAATTGCAAGAAAATTCCTGAAAAAGAAAAAAGAAAAAGTTAACCACTCCCCACACTTAAACATTACATTGTCCTCAATGTAATTGAATCATCCAACATAAAATTTAA
GATGGGAAATTCATTATATGCGAAAAAAAACAAATAAAAATAAAATAAGAAAATAGAATTACCTACTGGAATATGCATAAAGGATTCTCCATATATCAACAACCTAAATAGTTGCGGGATC
AACCCAAAGAACAATATATACAACTAGTGATAACCTCAGATCTCCGTTGTAGCAACGTAAAGATATAGAGATTATCAAAAGGAAAGAGCTGCCCACATACCTTGTTTTTGCAAAGGATAAA
TCGTCAGTAACAAGAATTACCGACTCAAGTGGTTCGTTGTGGGTGAGAACTGAAACCATTTGGGAGACATGCACAGGGAAAGAGAAGTCACGGGATGCAAAATAATCTAGTAATACCGGAA
AAGGACGCAATTCTAATCCAAAATTAGGAATCTTCCGAAAAGTTGTGCTAACAATTCTAAGTATATGAACTTCTCCACTGGGTTTGTGACCCTGTACTAAATCATCATCATAATCATCATC
AACCAAGGAATCAATATGAATATCACAATAAAGATGATTAAAGGAAACCTTATAGATAGAGTGTAATGTCTTTAAGTCGTTGGAGTTAACTATAGTTTCCTCAGAATTAAAACATTTTTCA
TAAGCATCATTGTTATAACACTGATATGTCACACCGTCCAAAAGAATGGTGTCCTTAATCCTAAACTCATCCTTTTGATTATTAGAGACTTAACAGAAAATGAGATGACAATCTTCAATCC
AATACATTATTTATTACATAATAAGTTAACACAAAATTGGTTAAAAAGACCAAAATCAACAATTCCTAGGTGAAAAGGACATTTAGATTTTGATACTGTTTAAATGGACAAAAATGTAAAA
ATAGCCAGGATATAAACAGTTTCATCCTACCCATTTTCAAATACTTTTCTTATTTTTAATTTATATCAGGATGCATCCAGTTTCATCCTTACTATTTTTTAAGTTTAAGTCAGGATGAAT
CCAGTTTTATCCTTGCTATTTTTTTGGTGTCCGTTTCACCCATAATAATTTTTACTCGTCCATTTGAACCATGTTTTAAAAATATTTGGACAAATGACCAATTTTCCGATAAATTAAGCAC
CATATGAATTCTTCAAAAGCAACTTATTTCAACCGTGATTACATCACTTGACCCAAACAGAGTCTTAACATATAATGAGACGAGAATCTTCAGACCAATACATTATTATTACATAATAAAT
CTCTCAGTCATCCACGAAATTTAAGCACAATATGAATTCTTCAGAAATGATAGGCATGCGAGTCGGCGCTCCATGGTTCTCTTTGAATCCTCTCCTCGTCTAACAGGGTAGTGGCAGAAGAA
TCCTCTGGAACTGAGAGAATGAAACCACTCTTTTTAAGTTCTACTTGAAAAGTTTCAAAACTTTCTGCGGTCTTCTTACCGGATTTTGGTTTCGAACTGCATGATTAGTATTGGGAATAAAC
ATTGGAATACTGAAATTGAATAACATAACAAAGCTTGGAATACGATAACAATGGAAAAGGAAAAATCACAGAAATTATAGTCATGTCGTAAGAAACGCAAAATGAAAATAACTGAGGCAAA
ATTAACTCACCGATAAGATGATGTAGCAAAAACAACAGTGGGAGCAGACGTATTTTGATCTGCCATGGAAACATAGGCGAAGTTGAAGAACACGGAGGAAGAAGACAAAGTACGGAAGAAG
AAGAAGTTGAACAGAGGAGGAAGACGGTAGAAAATAAAATATTTTTTTCTCTCTCCTGAGTATTTATCCGATAAAAAGAAGTAAAGACGTGTCGCTTGAAAAGAAAAAACTTTCAACACGT
GTAGCAGTTACGCGGAAAATACATAAGTGTGTCGGGTAATACAAAGAAGATGAAAAAACAGGACTGATGCGAAGTTTGACCAAGAATTCTCGCATTATTCTACTTGCATAAGAACAACGCG
ACAAGAAGCAAAATGTAGGGTCTAAATATTGCACAAGGCGAACACACAACTTCGCTCTAGATAAAACAACTCCAAGCACGAAGATATTATAATAATTAGAAGACATGGCGAAGAATAGTCG
TGCGACAATAAGACAAGATTGCGAGAAGATAATCAAGTGATAAAGATAGCACGGAGACCATATTCAAAGGAGTACCATGAAAGATACAACAGACAACACGATGCACATTCATACCTTCTCA
AAAGTCTGGCGCATAAGACAAACTGAAATAGAATCAGTCAACTGAGATCTTGGACGTGAATCATGATTGTCTTCTATTGATAAGTTTTCCTGTATAAGGACTTAGAAATCTAACAGAGAGA
GAGGGGTCGATTAAGAGAGAGATAGCTAAGGAAATTATAAGAAATAAAGAAATGAATTTCATCTTTAAATTTGGCTTTCATCTTGTATTCCTCAATTGTTGTATGTCGAAATCTAGTTGAA
TCTACAGTAAGAACAATACATATTCATTATCAAGAGATTGAGAAGATCATATAAGTGTCAAAGTGCTGAATAATATCATTAACGTTTGAGTTGATCAAAATGACTTAGACTTAATGTTATT
ATCAATACAAGGGTCTAGTTGTGGGATTTTCCGCAACTACATAAATGTTTCACAGAAATGATGAAAAAATTCGTGTTTAGATTTTTTTTTTCCTTTATAACTTTAACTGAAAATAGAGAAG
AGATAAGGGAACAGAAGCCGATATGTTTTGCTGGAAAAAAAAAATGAGTCATGTGAACCTTGAGTTTTTTATTTTTATTTTTTTAGTAAAGTTTTGTCACTTGTACACGGTGGACTCGCCGC
ATCAGCGCGAGAGGAGAATAGAACTGCGAGAAGAGTGGAGAAATCATTACCGAAATTATACACGCAAAAGTATCATAAAGTTGAGACGAAATGAAAATTCTTTCTGTACGAATCTTTAAT
TTAGTCGACAACTCAAAATCCAGACAGTGTCGTGTACTGTGCTACTGTCGCCAATTAGGTTAAGCCAGCACAACCCACGAAATTAACATGTTGGGTCGTGCTGGACTAAATCACATGCTGAT
ATGGCCTCAGATTTTAGCGTGTTGTGCTAGATTGTGCTGGTGCTGGTACAACCCAATTTACACCTCTAAGCACCACCACCAAGCCTTCCTTTAGTAGCTGTTTTAGAAGCAATGGGATCAA
GATTAGAGCGATAAACACAACAAGTAGCATGAGAAACCAACCAACGTACCCTAGCCTAACAGAGAACTTAACAGAAAATGAGAAGACAATCTTCAATCCAATACATTATTTATTACATAATA
AATTAAGCACCATATGAATTCTTCAAAAGCAACTTATTTCAACCGTGATTACATCACTTGACCCAAACAGAGTCTTACATATAGTGAGACGAGAATCTTCAAACCAATACATTATTATTAC
ATAATAAATCTCTCAGTCATCCACGAAATTTAAGCACAATATGAATTCTTCAAAACCAACTTATTTCATCATACTTCAAAAGCAACTTATTTCATCATATATTAAACACAATGTGGAGTAT
AATA**A**TATGTTCACACTTTCTCACATCCTCATGTAGTCGAGAAATGATTTTCCATCAAGCTTCCTTGAAAGATTTCCTTCAAAATATCCTCATACCTACCTCTTTTGAACAAAGCAGGTGT
CTCTGGTGTGACCAAGCT**A**GAAATTGGGCCTATTTCTGACTCTAGTTTAGAGTCATGAAATGTCGCGATTGAGAGCCTCTCCTTTGTTGAGTTTACTACTGCCCGGTGCTCGACGCTACGG
TAAATCCCATTAGTCATTATCTGTAAAAGGGAAACATTAAGTAACTAAGCAACTTTTCGGGGTACAGTCAAAACAGGTACTCACAACAAGGACCAGTAGGCAGGCCATATATGCCCTTAGC
CTAAATACCTCCAAAATGTCTCCAACATTCACTATGAACGCATCAGGTAGAGGTTTGATTGAAATCCACCTCTCTTCTTTTCTTATTTGTAATCCTTCAACTTCATTAAGTTGAAGGAGAA
TTGTCAAACCGCTAAAATCCGAGTGTGACGTAAGACCAAGTACAAGCTCTGGTCGAGGACAAGGAGGATAATAGTTCATCCTCATTGTTTGCAACCCATCTTCAAATAAGTCTGTCATACC
TTTAATCTCAACTAATTGTAGAGATTTTTCCAACATCTCAAAGACAACCGTTGATAGTTTTTTCATTTTTGATAGGTAGGATTCCAGTGTCTCCCTGCATAAAGGATGTAAATTAGGGATC
AAATGCAACTTCAGTATATTTTCCATCTGACAGATTTTAGGGATGCATGATGTTTGGATTCGATATATTTTGGAAACCATGCAATTCACAAACCCGGATTGCTTGATACTGCCTCCGAATG
TGAATTTTAGGCAACCTAGACAGGATTTTCCATGGTTTATCCTAGCAACCTAGACGATATTGATGAAACTATTCTGAACCAACCAAAATAACATTAAAAAAAAAAAATATCACGTGCGGCA
TTGAGAAAAAGCAAACGGAGAGAAACAATAACCTGAAAGGCAGAGGGAGTTCTGGAAACAAATGAGGCTTCCTTAAATGGAGAGGAAGACTTAACATGCTAAACACTTCAGTCCAATCAAG
TCTTTGGTCCTCCGATTCAATATAGGGTTGTCCAAATCCTTCAAAATCTCCATCTTGCTGTCCGTATTTAGTTTTCTCATTCATTGGAAGGTTAAAGAAACCTTTAATTTCTGATTTTATA
TTGTCCATCAGTAAAGCGTCGACTCCATGGTTAACCAGCTGAAACATAAAGCAGATTAGCATGAACTAGTAAAACACTAAATAGTCGTCAATCAACTACGTATAATCAAAATTGTATACC
TGAAAGAAACCCCATTCTTTGCAAGCAGAATGAAGCTTATCCAATTCTAACTTTCCAACTACGGGTTCTGGAGATAGTAAATTTTGCAAATCGATGACAGGAACTGTTTCATCATCTGTTA
CAGACGCACCAATATTATTCAACGGGCTTTCACCGGTGCATGTGTATCGAGATGGAATTTCTGCAAGCGTGAGTTTAGCCAATTCCTGAACACTTGGTATTGACAAACCATTGCCTAGCTT
GATAAGTATTGGTGTCTCCATGAACTCTATATTTGATGAACTTTCTTGTCAGATCTTCAGATCATATATCAATCTTTCAGGTCTCTTTGTTTCTCTCAAGTGGATATGCCTTACTTATATC
TATATTTCACTCGATGTCTAATGAATAAATGACTAGTGTTTTTTCAACGAAGGGTCTGAGCAACCACTGCTTACCCTACCAGTTTTAACCTGATGATGTCTACGATCTAACCTGCCGGTCT
AAGTTGCAAGATAATCTAACAACAACGAAATATGGTGAAGTTTGGTTGTTGTTTAGGTGATTAGGACTGATTTTTGTTGTTCTTCTCACTTTCTATATCTTTCTTGTTGCAGAACTTGGTA
CGATCTAAGGTATTAGATTGGGTCCATGCCTACATGAGCAGGAGGACGTGGGCAAATGGTGTACGATCTAAGGTATTTTCCCAAGCTAATGTAGCGACTGAGGAACGAGAACCACTACTTC
AGTTCATGTTAATAATAAAGTTAATCCTTTTCTTCACCAAATCTCAATCAAATCCACACAAGCAACTGCTTTTTAGTACTTTCTATAGGCCATATCTCCATACCTAAAGTTCCTAGAAACA
AACATAAGATTGAATTTATGATAAGATGGAAATATGATAACTGACAGTTAGTGATAGAAGAACGGTGAAAAAATTCACCTAGTAAGAATCGGTTGTTGATCATAGAAAATAAATAAAGTGA
AGATCTCGATTAGTAATTTAGTACTATTTTCAAATATTACTACCTCTGCATTCAGATTGAAGATCCTACCACCACTACCACTACCTTCATTGCGCTATTTGTTGAATTATCGCCGACTTTG
GCCCACTTGTCCTTCGAGAACGTATTAGCCACCACTCTTGGTCTCACACACCCATGTCCTGCATTATCTTATGCATAAAGATTCGTTGGGCAATGATTTACGATTGCATGGTGATGCGC
AATTGGGTATGTAATTATTAACGATTTCTACAATCATAGCATTACCTGCTCATGAAGAAAGCAATGAGATTGTCCCTCATTTTACAATGTGTGTGGTACACGAAATAGACTCGCCACGTGT
CGAACACAAAAGGAAGAAGCATGATTAGGAAATCTCACTTAAAGATCTCAGTCAGGACTTGTCAATTCGAATGTCAGAATTACCTCGTCACCGATCAATCCCGCGAAGTATCATATAGAAA
CAAGGAGACTTACACGAAGAATATAAGTTACGAAAGACCAAGCTAGAGTACCCAACAAGATTCTAATGAAGAAACCAAAAGGGAGAAGTACGGTAACTTGGTTGAAAAGATAGATCGCGAA
GTAACTATATTGGGGAGCAACAAAAGAGATAGGTGTCCCGACAAACCCATGTGAAGATGGCGAAAGAAGGGAGAAAATTTTATGGAAAATGGATTGGCGAAGTATAAAGCACCCCCGAAAG
AACTTGGGAAGTAAAATGAGTTGATACCCATTAGAGTTGATTGGCCTATAAATAAAGGTCCTTGGACAAGGCATGGAGGGCGGATTTTGGGAGAGAGAAAGAGAGCTTAGTGTAGAGAGT
GAAATTAGGGTTTCCTTGTGTTCAAGAACGGGTATTTTTGTTTCCTTAGAATTCATCAATAAAGATTCTTACACTACAAAAAAATTACTTTGTAGCCACACCACCGACTAGAGGGTATGTT
GCAAAATGTTTATATTGCCACAGTTATATCATGTTGTGGAAACATCTGATTTCTAGCCACATGGAAACCACTGGTGACGCAAAAAGTTTAGCTTTTGCCGCAGCAAGGTCCGGTGTGCGGC
AAAAAGACAGTTTATTGCTACGATTCATGTTATACGTGTGGCAAAAATGTATTAATATATTTTTATTCTTCATAAAATGGAAAGAAAGAGAATTAAACGTATTAATATATTTTTATTCTTC
CTAAAATGAAAAGAAAGAGAATTAATATATTCCGAGGAGTCAAAGTCTAACTCCACGTATACACCATTTACTACATATAACTTCTCCGGTTCTCATTCCCTCTTCTCTCTTAACAACTC
TGGTTCTTTTTCTCCCTTTTACGGGTTGAGATCCAAAAAAAGACTTCATATTTTGTTGAAAATCAATGGTATTCAAATTGTTAAAGAAGTTTTCGAATTTTTGGTTTCATTGCCGCTTGAG
AATTTAATTTATTCAATTATAGTGTTGACCAGTTCCCCTTCAAGAAGATTAACTTGATTTGGTAAGAACTAAGTGGAAATATTATTTAATTTCTTTAAGGTTTGATGACTGTTAGCGAAAT
AATTCACAGAGTATTCAACGATTTTACGAATTGTTTATTGTAACCTGATTTATGTTCACGGGTTTGGTAGTTTACATTAAAATATATTAAGTGTTTCTTTAGATTTAAATTGATAGTGTT
TTGATCTTTGTTTGCTTCCCTTCATTTTTCGTTTTTTAACTAGTTTAATAAAATCTGTTGATTTACAGCTCTTGTGGATATTGTTAATCAATTGTTGGGATATTTTCCATTGCAAGGTAAT
TTTCATCTTCTAGAAATACATGTACGATTTCGAATTATGATAGTAATTATCTTCCTTTGAACATTGTTTTTGCAGGAATGAACTACAAGATCGGAATGAAGTTTCATTTTTATTTTTATTT
TTTGTTTTCTCTAAATTATGATGTATCATTAATCTATTGTAATGAATGTATTGAGTAAATTATATTGTATATATACCAAATTGCATATGAAATAAAACCAAAAAGAGAATTTTATTCCTG
TTAAACATTGGTGTATTAAAATTTTGTTCCTGATAATGGTATAGAATCTTATAAAAGATCATAAGAAAAGATTTGGATCTAAGTAATATCACAATAACCAATATAATAGAGAAAAAAAACT
GAAATAAAATTCCATTTATAGAAAAAAAAGCCTATCAACAGATTTTAAAAGTGAAACAAATTAATTAAAATTAAATTACAAAACACAAAATTGGTTAAAAAGACCAAAATCAACAATTCCT
AGGTGAAATGTACAGTTATATTTTGATACTGTTTAAAAAGACAAAAATGTAAAAATAGGCAGGATGTAACCAGTTTCATCGTTCCCATTTTCAAAAATATTTTCTTATTTTTAATTTACAC
AGGATGTATCCAGTTTCATTCTCACTATTTTTTAAATTTCAGTTAGGATGAAACTGGTTTCATCGTTACTAATTTTTTTTTTGCCATTTCACCCAAACTATTTTTTACTCGTCCATTTGAA
CTGTGATTTAAAAATATTTGGACAAATGACCCATTTTCCGATTACAAAATACCTTTTTGCCTCATGATATTGCTTCACCGTTTGGTGTGGCAAAAATCCTAGCTTTTGCTGCACTACAATA
AAGGTGTGGCAAAATTTTTGTATGTGGCAAAAGGATTTAGCCACGTGAAATATGGGTTGTTGCCTCACCAATGCCGTGTTGCAATACCCCATTTTCTTGTAGTGTTAGTTCATATTACTT
AAACATGTCTTAGATCTTGATTGTTTTCTTTTTTAGGTGTACTACCGGATTTCAAGTAGTTACATTTTGGCGTCCAGAAACAGGCAGACTTAGATTGGGGATTCATCCTCATCTAAGGAGT
GGAGAGAATCTGAAGTTGTGGGAGTGAGAGATCTTAGCGAGAAGTTCCTTTTAAGATTAGGGTTTTAAGAGTAATTGATATTGAAACTAGGATAAGTTGATGACTAGAACACCAGATCGAG
CGGAACACATGATAGCGACAAGGAGAAGCAAGCGATTGGAGGCCAAAAGAGGAGGAAGAATGGAAAGAGGAGGAACATCAGCTACAGGAGAAAGGCAAGGCGTGAACCAGCAACGGACTGA
GGAGGCCAACGAAGATAACTTCAGAGAAGGTTCTGTACACACTTCCGACACCATTGCAGGAGAAGAGATGACTCATGAGGAAATGAAGGAGAATATCAGTGAAGAATTACATGACG
TTGGAGCAACTAAGGGATACACTTATCCAACTTTGACTCCTCGGAATATATTAATTCTGCAAAATCGCCAACTAAACGAGGAAGCTGCAGCTGACGTCACGGATTCAGAGGTGGACATGAT
ACCATTAGGAGAAGAAGATTTACGACGAAGGAGACATACTCGCGAGGAAGACGGAGGAGAGCGAAGAAGGAGACGACGAAGAGAAAGCGGTGAAGAAAGAGAACTGAGAAGGGCATTGGAA
ACATCAAATTGGGAAGATCAAAGGCGGAGATATGAAGAAGAGCGAAGATATGAGGAAAATGAGAGGAGGCAAGAAGATGAAAGACAACGCGAATTCAACCGGAAGAATGAGGAAGAAAAGA



TTTCCATGTTATCCATCAAGCTTACACAATACGCTTGTCATTACTCGTTTTCATCGAGATTCTATCCATTTCCTCGGTTATCTATAAATTTTATCGAATGTCCATAGATCTTACCCAATTT
CACCCGGCTTCACATGGATTTCATCAGTTTCTCAAGTTTTCATCGATATCACAAAATATCGCTCGACTTTGCTCAATTTCTCTTGGATATCTTCGAATTTTCTTGTTATCCATTAATTTCA
CATATCTACCGACTTCAATTAATTTCATCAATTGTGTCTGGTACCGGTAGCATAACAACCCGACTTGCTCCGACTTAACTTTTTCTTTTCTCCTTTTTAAGTGAAAAGATCTTACTGATTA
CAAGTGATACAACTCTCTTAACGAAGCCATCTACATGGGGCAAAAATATAATTGACACAAACATGTATGTTTTTTTTTTAATCAACTTTTTTTAAAAAAATATTATACGTTACTTTTTAAA
AATAGCGTTCTAAATTTCCACCAATTTTTTCGCTAAAAATCTAATCCCCTCCGCCTATTTTTACATCCCGTCCCATCTCTTTTAAATCCAACGACAATAAACTATGACAAGTCATTGATCC
GTATGCTCTCGTCGAGTTGTTTCACAATTGCAATCAGTAATGGCGGGAATAACTAAGTTTTGGCGCACTAAATCCCCAAAATTTTACAACAAAATTTTACTAAAATTCTTTTCCCTCCATT
GGAATCCTACTAGCCCCATCGTTGCCTATTATTTTTCTTCTTCGTCATTGCTACTTAAGGCTGTAATAAAACCCACAGACAGAATTTCATAGTTAATTTTCTCTCACCAAAATCAAAAACC
CAAAATCTACCCTAAACCGAAACCCATCTTCGAAACTTCATTTGAATGATTTTTTTTCGACAAAAAGAAATCATCATCAACAACAACGACAGGAACAAGTTTTTTCCCATTAGAAACCCAC
CATAGAATTCCCCTCATAAACATAGAATACATCTGATCTTTAGGACATGTTCTTGTGTAAATTCTATTACTCGATTTCGAATCCGATCTCAGATTAGTATGTCTGTAATCTGCATT
TACCCATAAAATGGTTAGGCAAATCATTTACAAATTACAGTTTTTCAATTTAATGGTGGAAATTGGAGTTCATCCGATTGTAACCGTAGAGATCCATTAATCAGGTAAGTAAGTTTGATAT
GTTCATCTATTTATTTCAAATACGTTTTATGTGGATTAAATATTTTGAAGCTTTGGTACCAATTACATCTATGTTTTTCTAATTTCTTATGATTTCCGTTAGTTTCTCAGAAATCTATTCG
AATCTTGTATTGTTCTAAAGAACAGTAATTGATTTGCTAATTGTGTTTTTATTTTGTTGCAGCTAGGACTGAGTTTTGGTTCGTTTGAAGGATTTGGGTATCCATCAGCAGCTGAGTGAAT
CACCAGTTTTGGGTAATTTTCTTCTTGTATTTCTTTGCTTCCTCGTTTGAGTTAATCTCAGTCATCATTTTGATACGTGTTACCTGTTTGAGATAATGTCCCACTTGCAAGTAATTATGGAT
TGCTTTCTCAACAGTAGTTATAGTGGTTAAGAAGATGCAGCTCTGGTTTTTATTTCTTTAGTCATTCCTCAATTTGTTGTGATTTGCTTCCGTACTTATGCAGTAAGTGCTCACCTGATTT
TGTGACAATTCCAAGTGGATGATATGTAATCCTAACTTTTATGAGACTGAGTGTAGACCATATATTATCTACCAATTATTTTCACCAAATCTCTTCTCAACAGCCATGCTAAGTGAGTCTA
CTTTCATATTTCTTGTAGGGCCATCTGAGGTTCCTGAATAGAAGCTAGAATTCAAGCATAGGCTTCCAATAGCTCTTGGAGCAGCTAAAGGTGATCCTCTGGAAAAATCAGAGAGCATTGA
TGGCTTCCCTGTATTGGTTTTGGTTATTTTAGTAAGAAAAAGCATTAGAAGGTTCTTTGTTCGCCCATGCTATAGATGCTGCCTCTAAGGCATTAATTTAGTCAAGAACGTCCTCTTTATG
GACTAGAATCTTGAAGCAGTACTCTCTGGATAACGTCTTTCTTATAATTTCCATGCAGGTTTAGCCCACCTGCATACTATGAGTCCTCCTTTGGTACATAAGAATTTCAAAACATGCAGAT
GTTTTACAATGCGACACCTCAAGCATTAGTAAGGACACTTGTTAGTGTTTGCCACGTCTTACTTTATAACGACACATCAGGGTGACATAATTTCAAGTCGATCTAGTAGTATTAGTTATC
AGTTTGATTGCCAAACTAGGTTATTTGCATTTCTAGGGATTCTGTCGACAGTCTCAGTGATGATGTCTCAAAACTTGCCAGAGGATTGTCAAATATTGGTGAATTGGTTTCTATAATGGAT
TCTGCTAAAAAGGTCATTAGTTCATTAAGCACTTCGTTTTTGTATTTTGTTATAACTAAAGCAAGAAAAGAAGTAAAAGGATACAAAAGATAAAATCCCATCGCGGTTGATGTTTTGCATC
AGATTTCATTGTGGGTAAGTTTGTCACATGAATCGGATAGTTGGAACCAATATCATCAATCAAGCATTCAAGCCATATGAAGGTTTGGGATTTTTGGCAATAAGCTTTTGACTGTTCCA
TGTTTCCCTTCCTCTATGTGAGTAGGACGGAGTGTAGGGGTGTTCTTCAACAATTTTTTTTTTTTTTTTGGGATTTCCTCTTTGGTAGAGTCTGAAGCATCCAGCAATTGCATTCCCTACC
TGACTCTTTTGTTTTTAATTGTAGGATGGAAAATACGGTGTGCCGAATTATGCGGTTTCTCGTGTTACTTCGAAGGTAGTCGAAGCTCCATCATGATCAAATCTGCTGATGGGAAGATTTC
TGATACCCTCGTGAATGGGGGTTTGGGTCATCATAAAGATTCTATTTCTGTAAACGGGAAGTTAGAGACATTCAAGGACCTTATCATCTAGCCTTGCTCTGCATGTCTTCTTTTCATTTA
GTCCGGAGTTCAAAGTCATCCGCATAAAGAAGTAATGAGAACAGAAGTCTGTAGAATGCTGAGACTGATTTCAACAAATGTACCGAAATCATCTTTGAGTTGAAGAGGAAAAGAGGTGAAA
GAAGATGATGGGAGGCGGGTCAAGTTTGTGATGATTCCCAAGATTTTGTCAGCACAATTGCATTAAGCATTTAGCTGGTTCTATGAGATTTTACTTTAGTTTATTTTTTCTAGTTTCCCGC
CATTTTTATTCTGACCAGTTCCAGAAGTGAATTCCAAACCAAGATGTGTGTAATTAGTTCCTTTCTATTTATTTTGTTATATTTCCATGGAGGGGTTCCTTTGCAATAATAAGTTGGTATT
CTTATTTGGTGGCCAACCTTAAGGGTGCAGGGAAGAGATTCTCTTCGCGGCAGTTTCCATGTAAAATGTGACGGTGAAAATCAGAAGTTTTATTGATCAAATCTTGGGATTTTGTAGAT
CATTCTTAGTAACTTATATCCCTATTAGACCGGTATATATCCTAGTTAGTTACTAGATAGAAGAAAGTATTTCTCAGTTGCTTGACTTGTCAATCTCGTGTTTAGTACTCAAATATGCGAA
TCCCTGTCTCTTGTTTTTTGTTGTAATTTCCCTTCTTATAGATTATGCTCACAGAATAATAATCACCTATTGATATTAGTATAATAGGAGGGATATTGGGTTGTTCTTGAGGTTCAAGTCC
ATTGTTTGAGATCTAGGTGGACCTTCGGGTGATTCAGGTTCAGCTGCTGCCTGTGAGGAGTGTCCTTGTTAAATCCATCACGTATGCCTGTGGATGCCAGTCATACTCTGAAGCAACATCA
CCACTTCAATTCCTAGAACTGAGTAGACAAAATCACAAGCAGTCACTGAATATTCGGATTTGGTGACTATAATTTGTCTTGCCACTACAAAGTTTTAACGGTAAAATTTCATATTTGCTTC
GTATTCATTTAGATGACGAATCAATTTTTGTTTGCACAAATAGATATAAGGACAAAAGTTACACATTATGAGCAAATAGTTCTTTCGGAGACCAATTTTTTCTCGTCAGCAGAAACAATCC
ATTTGCGACAACAATTTGTTTGGATACTAATAGTTCTAACAGCAAAATCCTTATTTGACGGGCATTTATTCGGCAACAAACTAATTTTGTCAGTGTAAGATACGTTCAGGGACAAAAGTC
TCGGATTGTTAGCAATAACATGTTTGGAGACTATAATTTTCCTTTGTCACTAGAAACCAAGTCCGTGACCGAAATAATTGGTCACAAAGAAATATTTTGAGGCGGAGACAAAATTGGTCAC
TATATTTGGTCACAAGAAAAACGTTTTGGGACGGAGACAGGATTGGTCACTATATTTGTTCACAAAAATTTTCAGTGACAACTTTAAAGACTAAATTTACCGTTTGTCGCCGAAAAAATAT
ATATGGTGACCAAAAAAATATTCCTCCTTAAACACTTGATTGGTAGACTAAATTTTATTGGTCACATAAAATTTTGTCACTAAACGTGTGTTTTGTTGTAGTGAAAAAATCAGCTAAAAAT
ATACAAAATGAGTAATTTTATTATCTTCAAAAGAGGTTTGGTGAACCCTGGTACATTTGAGTGAGGTTTCATGACCTTCGTTACAAGTAGGGATTAGGCCCTATTCTTGGATTTTGATTAT
CTAGTAAGTGGTCCTGTTAAGCATTGTAAGATTTTAGGTACTTGATGTTCCATGGGCGTTCAAGTTCTTCACCGTTCGATTTGAGCAATCTATATGCTCCATCTCCTGCCTTAGCTAGGAC
TATCAGTGGGCCTTCCCATACGGGTGCGAATCTTCCTGACTCGCGCTTGTATGGAAAAATCTCGCGCCAGACCTTATCGCCTAGGCTGAATTCTCTCTGGTGAACACGCTTGTTGTATTCC
CGAGCTAACCTTCTGTGATAGTTTTCGATCCTCTAGAAAGCGATTTCTTTATTCTCCTCCAAATCGTCTAGTCTGGAGAGTATCATGTCCGAGGTTAGGTTTTTCTCCCATGCCTCCGTCT
TGGTTGTTGGGATCATCGCTTCTGTTGGTATGATCGCCTATGTGTCGTATGTGAGAGCGAAGGGGGTCATCCTTGTTGCTTCCCTCCTGGTGGTCCGATAGGCCCACAAGATGTTATGTAA
TTGTTCACACTAACTCGCGCCATGCACATCGAGTTTCTTTTTGAGGATGTCAGTGATGGTTTTGTTAGTTGCCTCTGCCTGGCCATTGCTTTCGGGGTAGATCAGTGTAGATTTGTTTTC
CTGATCTTGAAACAATTGAAGAGTAGATCGATTTTTTTCCCTCAGAGTTGGGCCCCATTGTCCGAAACGATGGCGGCTGGGATACCGAAGCGGCATATGATATGCTCGAAGAGGAATTTGA
AGACATCGTGATCCCTGATGTGTATTAGGTCCGAGGCTTCTACCCACTTTATGAAGTAATCTGTGGCCACTATTAAATACTTTCTTTGTCTGGTTCCAGCTCTTAAGGGTCCGACAATATC
GACTCCCCACTTGGAGAATGGCCAAGGGCTTAGGACAGATTTTATATCCATTAAAGGTGCATGCCTTCTTTTGCCAAATTTTTGGCACTCAATGCAAAACTTTACCATTTTTTTTGAGTCC
TCGTTCATGCGAGGCCAAAAGTATCATTCTAACCTTGCTTTGTGGGCTAGGAATCGCCCACCACTGTGGTTTCCCGTATAACCGTAGTGTATTTCTTTTAATATGTTTGGCCTTCAGTGA
AGGAAAGGCATCTGAGAAGAGGTCATAGGTAAGATTGTTTGTATAGAATTCCATCTCGGAGTTTATAGTTCCCCGATCTTGATATTATCTTGTGTGCAACTTGTTGGACGCTTGGGAGGTC
CCCGTTAGCTAGATATTTTGTAATTTTCATCCTCCAATCATCCACAGGGAGACTCTCTTCGTGGGTTGTTGAGTCTAAAGTAAGGATGTCCATATTTCCCTCTATGGGAATGGATGGGAGC
ATGACCCTTTCAATTCTGACATTAGTAGCCTCGGGGTCTGTTATCATGGAGGCTAGAATTGCCAGGGCGTCAGCGTATCTGTTATTTATCCGATTTAAGTTTCGGAATGTGATTTTTTGTA
TCTGCTCGATGTAGTGCTTGGCGAGTTGCAGGTATCGCTGCAAGCATGGGTCTGAGGCTTGATACATTTCATCTATCTGCGGAACAATGAGTTGCGAGTCACTTGTGAGACGTACTTCACT
GATGCCTAACGCGATTATCACTCGGAGGGCGTGTATGACTTCCTCGTACTCGATTATATTATTATTGGCAGGGAACTCGAGTCTCAGGATGTAGACGATCCTTCGCCCGGTCGGCGTTGTA
AAAACCATTATGATGCCAGAGCCCTTCGAATTGCATGACCCGTCCACGAAGACTTCCCATCGCCTTAGATTTTCTGAGCGAAGTACCTCGGGAGGATCTTCTAATTCGTCTTCCATTCCTG
GGATTTGTTCGATCTCATCAGTGTCATCTAATGGGAAATCGTCCAGGAAATCCGCAACCGCTTGGGTTTTAATTGTCGTTTGGGCTTGATATTTGACTCCAAACTGTTTGATCTGTGCACC
CCATTTGGAGATTCTGCCAGCTCTGTCATCATGATCGAGTACAGATTCTATTGGAGCTTTCGTGATGACTGTTATCGTGTGTGCATGAAAATAAGTACGCAATTTCTGCGTTGCATAGACC
AGAGATAACACCATCTTCTCAATCTTTGAATAGTTCTTCTCTGCCGAAGTCATAGTCTAACTTATAAAATAAATGGGTTGTTCCACTTCCTTGTCTATGCAAGTTAGCACCGCACTGATTG
AGGTATTAGTTTTCGAAAGGTAGATGTAAAGGTTTTCACCTGGGTCGGGCTTCTGAAGGATCGGCAAGGCTGCCAGATGTTTTTTGATCTTTTGAAAAGCTTCTTCGTAGTCGATACTCCA
TGAAAATTTTTCACCATTTTTCAGGGTATTGAAGAAGTCTTTACATTTATCGGACGATCTAGCAATGAATATTCCCAAGGATCCTAGTCGACCATTGACTCTCTGCACGTCTTTGACTGTC
GCAGGGGAAGGCATTTCTAGTATAGCACGCACCTTTTCGGGATCTACCTCGATTCCATGCTTGGTGACCAAGTATCCCAGGAATGTTTCTCGAGGTTACTCCAAAGGCGCATTTTGCAGGAT
TGCTTTTCCATGTTAAATTTCCTCATTTGGTAGAAGACTCCCTGAGATCGTTTGCATGATCTTGTGGAAGCTTTCTTTTGACGAGCATATCATCTACATATACTTCTAAAGTTTTATGTAT
CCAGCCATCAAACATTTTTTCCACCATTCTTTGGTATGTGGCGTCTGCGTTTTTTAATCCAAATGGCATTTTGGTATAGCAGTATAGACCTCGGGGGGTGAAGAAGGAGGTATGCTCCTGG
TCCTCCTCAGCGAGGGGGGTCGAGTTATAACCCGAGTATCCGTCCATTAGGGATATAAATCTGTATCCGGAGGTTTCCTCCACGAGATGATCAATGTTAGGGAGTGGAAGCTGTCTTTTGG
ACAAACTTTATTAAGGTCGCTGAAATCAATGCAGATCCTTACCCCTCCATTTTCTTAGGTACAATCACCATGTTAGATATCCATACTGGATACTGTGCTTCTCGTATCACCCCTGATGCC
TCCATTTTTTTTAGTTCCTTTTGGATAGCTTCGTGGTATTCTGGTGCTACCTGTCGCATCTTTTGTCGTATGGGCTTGTGCTTCGGGTTGATATTTAGGTGGTGACAGCAGACTGCCGGGT
CTATTCCGGGCATGTCCTCCATGCGCAACGCGAAGACATCAGAGTATTCTCTTAATACTTGTATTAATTTTTCTTCTTCCTCCTTATCGAGGATATTTCTTATTCTCACCATGCTCGGATT
TTCGGGTGTTCTGATGTTTATTTCTTTTGTCGGCTCCGAGGCCGAGAAATTGGATTTGAGTCCCCCGATATCGGAGAGATTTTGTATGGTCTCCATGGGCGTGTCTGAGGTAGGGACGTTG
TCCTGAACTGGCTCACTGTTTGATCGTCTTGCCTCATTCGTCATGGATATATAAGCATCGATATCGGCCTGATTTTTGGCATTCTTGGCTCGCTTCTTTTTCGCTCTTGCAGAATTTACTT
TGTTTTCGTCATTCCGTACGTCCATCTAGTAGCAGTGCTTTGCTTCTCCAGAGTCTCCAATAATTTCTGCGACTCTAGATGGCGTGGGGAATCTTAGTCTTTGATGGTAGGTAGGGTGATA
ATACATCTACCACACAGAAATTTGTTTGGGTGAGGATTGTTCCCACCCGGATCTCTAATGTCACCTCGCCTTTAGGGCGAGTTCGAGGAGCCGTTAAAGCCAAATATGTTGTAGGTGGAGGG
TATTAGGCAATAGTCTCTAAGAACCATCTGCTTGAATGTTTCGTAAAACAAGATCTCGACAGAACTTCCTCCGTCGATCAGGATTTTTGGCATCGCCCAAGGCAATGTCTTTGGTGTTTCC
TCCTTGCCTTCGCATTCGGGAAGTGCGATAGCCATCGTGACCACTAGGGGGTCGTTGTGGTTTCTCCCTCCATTTGGTGCCTCCGAGGCTGGGAAACTGATAGGTAATTTTATCCACTCCT
CTATCAGGGGCTCTCTAGCGACACTGTAGACTTCATCACCCGCGAAGTTGCGTTTGTGAATTCATCCTGTAATATTTCCTTCAGGTGTGGGGATCACGATTGACGAGTGTGATATTGTATT
GCACCCTAGATATTGCACCTCACGAGGGATTTGCACCCGAGGTACGGGTGCCCCGGATGTGAGTAGCATCACTGGTTCCGCGATGTACTCTTGAAGCTTTCCTTGGTCGATCATATATTGC
ACCATCTTCTTCAAGTCCCTAAACGAGTCTGTCGTGTGGCTATGCAATTGATGGAACTCGCAGAAGTCAGTTCTGTTTTTAGCCCGCTCTGGATGCTGACCTCTGTTCCATGAATAAATGA
TTTTATGATTTCTTCGATCTGTTTGAGGATCTCGAATATGACAGTGTTTAGCTTCGTGCACACTGGATCTACGAAATTTTTCTTCTTTTGATGGGCCTGGTCTCGCCCTTTCCATCCCCA
GTCCGATGTTTATCGAACTGATTTGACCTATTGTTGGATCGCCCCGCTCCTTGGGGTGGCACCTGAAGGTTCTGTGGTTCTACCACATTTTCCCCTCTACTAGTTCCTTTCAACAACTTAT
CGTAAGTACCGTATTGAAGTTCTTCCAGGGATATTTAATCCTCTTGGATCTCTCTTAACTTCCCAAAATTCTTTGGCATGGTTTCATACATGCGTACGAATATAGGGTATGTCTTCCTAAG
ACTATTTTTGAAGCCCAAGATTGCGTAAGCTTCAGGGACTTTTCCAATCTCGGCACATAGTTTTCTCCATCGTTTTACCAAGGAGCGGAGTGACTCATTTTTCCCTCCGGGATATCTGGAAC
AACGCGTCAACTTCGGGTCGAACTCTGCTATTCTGAATATATATGTTTCTAGAAACATCTCAGATAAGTGAGCATATGACTTAACTGAGCCTTTTGGAAAGTTCTTGAACCACATCAGGGC
TTCGTCTCTTAGGCTAGCCGGAAAACACTTACACAACACTACATCGTAATGATCCCATTGGGTTAGCGTCATGCGGTAAGTCCTTAGGTGCTCGATCGCATCTCCAGTTTCGGTGAACGGA
GTTGTGAACGTTGGTGGGCCCGGAAAGGTGTGTAAAAATTGAAGTGAAATGAAACGGGCCTACAGTAATACCGCAAGTGCACGGTTGTCAGTTATAGCTCGTGCAAGTACGGGTCGATCCA
CAGAGACTGGGTGTGTTTGGAGTTTCTAGCTATTTTGGGCTCTAAAATTGCTATTGTTGGGCTTTGGGTACCAATGGATTTTGGTAACCAATGGGTTATGATTGTGATTTGGGCCTTTGAA
TACTCTTGGAATGAATTGGGGTGTTGTCTTAATGAAATGAACTGAGCTTGGGCTCAGTTGGTTTTGTAGTGCACTAGGCTTTGGTTTTGAATTTGGGCTTGGGCTTATTGTAGTGAACTGA
GCCTTTGGTGGACTGAATTAGACTGGGCTTGTGATGTGATTTGAGACTTGGGTTGAGCTGGGCTTTCACAGTGAACTGTGGCTGGGCCTTAACCTGAACTGTGGTCTTAGCTGAGCCAAGC
TCAGTTGCAGCAGCAGCAGTAACAGCAAGTAGTAGCAGCAGGGGGAGAACAAGGCAGCAGAAGCAAAGGAAGACAGGAAACAAAATCACACAGGGCCAAGGATGGCATTACAATGACAAT
GAAGATGGTAGTGAAACAAACTAACAATGATCATGGGAAAGAAGCAAAAATCAGTGGCAATGGAGGCACAAAGGCAGTTAACCTAAGGCAAGGAGCAGCAATGAAACAAATGGTAACAGTG
GCAGTTAACAGGAAACAAAACCAAATAAACCAAGGCTGTTAGCCAAGGGAAGTGGGAGAACAAGGCACAAAACAGTTATAAAAAAAACAAATTGAGAGCCTAGGCATACAACTAGAGTGGG
AAGAGAACCAGTTTGCTCACAGTCACTAGTGAGCACTAGTTTCTCCCCACTGCTCAATCAATCCAATGCTTCAAAGGCTTAGCCTAACATTCACACAAAACAGTAAACATCACAGAGATA
ACATAGTATCATACACAAATCAGGAGCATTTACAGTGAAATTAAACAGTGAAACATTCTAATTAAACTATCCTAACATAGATTAAACAATGAAACATTCACAAAATTTACAAAATAAAATT
AACAGTAACACAGACAATTAACAGCAAACACTAAGAATAAAATTGAAAATTGAAATTAACTGAATTTGAAATTAAAACTAACTTAAACATGGAATTAAACAATAAAGTATGGAATTGAAAT
TTGGACAGTGGCAGCAAATGAAGAACCCTAAATTAACAGAACCCTAACTGTTTATGAAATTAGAATTTTACAAACTAAAATTAAACCTAATTGGTAACTTGGATAGTCCAAGAAC
ACCCTTCTCTCTACACATCAGTTGCCTTTTATAGGTTTACAAGATATCCCCAAATTTCGAAACCCTAATTTTTGAATTGAAAAGTCCACTCACCTAATCCCTGAATTGATGTCGACCCATG
CCTATTCTCTCTTCTCTGCTCCTCTCCATGTCTTCAATTGCGTCTTATAGCCTCACCTAATTTATTGATTTATCACCTAGGGTTTCAGTGGTGAAAAATCAATGAGTTAGATGGCTATAGA
GGTAGGGGAGGTAGCTACGGCGTGTAGGTGGTTTGGTGACAGTGGAGGAAGTGGCGATGGCAGGGTGGTGGTGCGGAAGGGTATGGTGGTGGTGGCAGTGAGAGCAGGCGGAGTGGTAG
TTGCAGACAGATAGGTTGAGAGGGAGGAGAAGACTCGAGTGTTTTGGGATGTAGGGAGTGTTTGGTTGGGTATAGGTGTTCGATACTTGGGTGTTAGGCGGGATATCAAGATTCGATGTTG
GCGAATCTGAGTCACTGGATGCGAAGCTGGTAGGTGAATCGGATGGCTACCCCTGAAGAATCTGGTAGCGACCGTCGGATGTCTTGATACAACGAAGTTAACGGCGAAGATCGAGGTTAGG
TGCTGTAGTGTTTAGCGGAAGTATCGAGATTTGATGCGCAGGCAAAGAAGCGACCGTAGGATCTAAATGTGATCTAATCTGAAGGCTTGGAATTTAGGTACTATGGTGTTTTGCAGAGACT
TCAGATTTTGATGCTCTACGAAGGATCGACCATTGGATGATGAGATGGATCCGATCTAACGGTTGAGAATGGAGGCGGGTATGGATATAGAAAATGGGTTCGGTAAGGGTTTGGGCCTT
GGGTATGCCAAGCCCATATCTTCTTTAAGAACAATTCTTCCTTCTTGAGCCCATTCCTAGCTTTTTGGACGTGCGCTCCATTCTTTGCGGCTTCCTTGCGTAATTTCTTCCGGCTTTTCAC
CACTCTTCAGCTCTTTTCCGCTCCGCAAGTTATCCAGACTTTATTTATTACCTAAAAATGCAAAATTAAGTAAGAAAAATATTTATTCTTGAAAACAATGAAAATACAGAATATGGGATAA
AATGTAGAATTACTGCACAAAAGATGAGTTAAATGCCAACAAAAAGGGATAAATATATACAATATTTGGAACTCATCAAATACCCCCAAACCTGAATTTTACTTGTCCTCAAGTAAAACAA

AACTAAGGAAATCCTACCTATACCACTGTCGCTGGTCTCTCGAATGCATTTAGCGTATGCACTAAGCCTTTTAAACCACTAAGTGTCCCTAGTGGACGAGTTGAAGTCTCGTGAAGGTTTG
CTTAGAACGTACCTACAAAGTTCTAGGTCAAAATATAAGCTCAGATTCCATCAAATGTGACATGTGCAAGTCAGTTTAAGCTCACAGCAAAATGGAGATGTCAATCTAGCTATCAAAGGCA
CAATCCTAGCACTGATAACAAATAAAGACATGTGATAAGAGTGTAAAGTGTATCTACACATGTGTAAAGAAAGATCGGATGTTATGACTACTAATCACCAAGAGATAGTTTCTCAGGCTAA
GAACCAAGGTCGAAATCTAGCTAGCTGTCCGGACTTTACGAGAATTGTGAATGAGTTGGAGGTATTTCACAATTACTCGCGTTGTACATCAATGGCATACACCCTCCTTGCTTATTACAAT
GAAACAACAAAATGACTATTTACATGACTCTTATTTACATTGACTACTCTCTTTTATTTTTGGAACAAGAGATGATGGAATTGATAAATACTTGACTTTTTTGTATTTTTCTGAAATTTTT
TTTTTTTTTTGAATAAGGAAACAGTTTTGATACATATACAAAAGGAAACAAAAATTACATGACACTTTGCAAGAGGTAGCCCTTTTTGATGCACCCAGTTAAATTCGATGGTTGTCTTTCT
TAATGTAACCTCCACCTTCTATCCCTACCAACCAAAGAACAAGCTAGTCAAGTTTCGTTCAGTATTCTAAAGTGATTGGCAATCGTAACTTCCTATCAAACACCTTGAAGATCGATGCCAT
ACATGTATTGGTAGATCGTGCGCGTGCAAATTTCTTATCACTATGTGAATTGTGCTAGAATCAGGGTGCCTAAATATCTAGACTAAGACTCCTAATAATTACATATTTGCACAAGAGTCAA
CATTTCAAGGTAAATGAGCTCCTTTTTTATGATTTTTCATTTTTTAATTTTTTTGAATTTTTTCGATTTTTTCAAAAGAAGAAGGAGTTCGTTTTCAATTATGGCATATTATCAAAGTATC
TACTTTTCACCCCCAAACCTAAACTAAACATTGTCCTCAATGTTTAAAAATATGAACAAAATTATAATACAACATATGAAGAGGATCATGTTGAGTAGAGAAAAAGGAAAGAGAATACCCG
ATTTCGGCGAAAGCAAGATTAAACTTCGTTATTCAAGGAAAAAAAAATCCAACATATTTCAGCCGAGATCATATTGGATTAGAAAAACATATACAAAATGAACAAAAGGGGTTTTAAGAAAT
TTTATCTACTGGATTATATACAAAAATTCACCATACACTAACAATCGAAAGAGTTGAGGATCAACCCAAAAGACAAAGTGTAGAGGTTTCAACAGCTTCACACAATAATAATATGGTAGGC
ATGCAAGTGAAGCTGTGAAACAAAATGAGCTACCCCCAAACCTGGATTTTACAGAAGATATAATTTTGAAAACAAAATCGCGT

Attempt to amplify CODMa/c and CODMb from CODM-containing BACs



Appendix K- Oligonucleotide primers used in the study

CODM primers used	
BAC library probe design & sequencing	
CODM_F4	TATCCTCCTTGTCCTCGACC
CODM_R3	TGTTCACACTTTCTCACATCC
CODM_F5	GGTTTGTCAATACCAAGTGTTC
CODM_R4	TTGTCCATCAGTAAAGCGTC
CODM_F6	TACGGACAGCAAGATGGAG
CODM_R3	TGTTCACACTTTCTCACATCC
CODM_F7	AAAATCCGCCCTCCATGC
CODM_R5	CCGACTTTGGCCCACTTGT
Cloning & sequencing	
CODM_F1a	CAGTGGTTGCTCAGACCCTTC
CODM_F1b	GAGAGAAAGAAACACAAGTACTTCCC
CODM_F1c	GAAACACAAGTACTTCCCATTCTG
CODM_F1d	GAAAGAAACACAAGTACTTCCCATTA
CODM_F1f	GAGACCTGAAAGATTGATATATGATCTG
CODM_F2	GAATCGGAGGACCAAAGACTT
CODM_F3	CGGGCAGTAGTAAACTCAACA
CODM_R1	CCCAAACAGAGTCTTACATATAGTGA
CODM_R2	CGCTAAAATCCGAGTGTGACGTA
Potential primer pairs for TILLING (Figure 16) & sequencing	
CODM_F8	CGGAGGCAGTATCAAGCAAT
CODM_F9	CGACGCTTTACTGATGGACA
CODM_R5	CATCCACGAAATTTAAGCACAA
CODM_F10	ACGCTTGCAGAAATTCCATC
CODM_R6	ACCCGGATTGCTTGATACTG
CODM_F11	TTTTCTCAATGCCGCACGTG
CODM_F12	AACCATGGAGTCGACGCTTT
CODM_R7	TCTCTGGTGTGACCAAGCTC
CODM_R8	TCTCTGGTGTGACCAAGCTA
LongAmp experiment (Figure 30**)**	
CODM_F	CATATTGTGCTTAAATTTCGTGGATGACTGAGAG
CODM_R	CCATCTTGCTGTCCGTATTTAGTTTTCTCATTCAT
Verification of duplications in CODM promoter sequence (Figure 15**)**	
CODM_F14	TAATTTGATTTTGATGCTAGATCCTG
CODM_R10	AAGGGTCTGAGCAACCACTG
QPCR (Wijekoon and Facchini, 2012)	
CODM_fwd	TTGTGCTTAAATTTCGTGGATGAC
CODM_rev	TGATTACATCACTTGACCCAAACAG

Table 29. CODM primers used in the study

T6ODM primers used	
BAC library probe design	
T6ODM_Bac_F1	CCGAGATTAAGGGTATGTCAGAGG
T6ODM_Bac_R1	CACAAGATCCCCATATGTAGATCCAC
Cloning & sequencing	
T6ODM_F1	GTTCTTAATTCATTAATTAATTTAGAAAAATC
T6ODM_F2	CTTATGAAGCTAGGTAATGGTATGGA
T6ODM_F10	CGTAAATACAAGCATTGTGATTGCAT
T6ODM_F15	CATGAAACCATTCCTGTCATC
T6ODM_F17	GAAAATTTATTATCTCCAGAACCAA
T6ODM_F19	CAAGCATTGTGATTGCATCC
T6ODM_R2	CATCCTCATTGCTTGTGTCC
T6ODM_R4	TCAATTTTCTCATTTGTGCCTCAATAC
T6ODM_R8	ATTAAAGAGAACCATGGATAAC
T6ODM_R9	GTGGAAGAGTGAACATCATAAATATA
T6ODM_LongAmp_F1	ATCTCGATACGTATGCGCCAATGAAAACCT
T6ODM_LongAmp_F2	CTTATGAAGCTAGGTAATGGTATGGA
T6ODM_LongAmp_F4	TGGAGTCGACGCTTCATTGGTGGATAG
T6DOM_LongAmp_F6	ACAAGCAGCCGAGATTAAGGGTATG
T6ODM_LongAmp_F7	GCCATCGGTCTTACGTCGCACTCG
T6ODM_LongAmp_R1	CCGAGTGCGACGTAAGACCGATGGCGAGAT
T6ODM_LongAmp_R2	ATGATCAAAACAATTAAGGACTCGAGCACCAAAAA
T6ODM_LongAmp_R3	ATGATCAAAACAATTAAGGACTCGAGCACCA
Potential primer pairs for TILLING (Figure 31 **&** Figure 32**) & sequencing**	
T6ODM_F13	CTGGTTCCCCCATTGACCAA
T6ODM_R16	ACAATTAAGGACTCGAGCACCA
T6ODM_F14	CACGCTTGCCGAAATTCCAT
T6ODM_R18	AGGTGTACTCCGTTCTAGCTCT
T6ODM_R27	GTGCACAGAAACCGACTAC
LongAmp experiment (Figure 30**)**	
T6ODM_F	TTTTTGGTGCTCGAGTCCTTAATTGTTTTGAT
T6ODM_R	CATCTCCATCTTCCTGTTCATATTTAGTTTTCTCATCC
QPCR (Wijekoon and Facchini, 2012)	
T6ODM_fwd	TTGAGGCACAAATGAGAAAATTGA
T6ODM_rev	CACAACGCACTTTCGAGAAATTAC

Table 30. T6ODM primers used in the study

Other primers used	
Cloning & sequencing	
DIOX2_F2	GAAGCTGGGTAATGGCATGTC
DIOX2_F4	AATCGAATCCTACTCATCAGAAATG
DIOX2_R2	ATCTTTAAACACCTCTGATATTTC
DIOX2_R3	CAATCACGTAATATTATTGTGTGTTC
M13_F	GGAAACAGCTATGACCATG
M13_R	GTAAAACGACGGCCAGT
LongAmp experiment (Figure 30**)**	
DIOX2_F	GTCTGAGTCCTTATTTGTTTTCATCATTGCTTTTTG
DIOX2_R	GAGATGGAATTTCTGCAAGCGTGAGTTTAGC
Validation of TILLING on the Fragment Analyser (Figure 10**)**	
BBE_OF	TCCTAACGAAACTGGAATGG
BBE_0R	GCATCATTTCCGTTTACTCC
BBE_IF	TGTAAAACGACGGCCAGTGTTTACTCCCACGTGCATC
BBE_IR	AGGAAACAGCTATGACCATTCCTCGTCTAATTCATCTGC
QPCR (Wijekoon and Facchini, 2012)	
B-Actin_fwd	TCTCAACCCAAAGGCTAATCG
B-Actin_rev	CCCCAGAATCCAAGACAATAC
GAPDH_fwd	CTCATTTGAAGGGTGGAGC
GAPDH_rev	GTCATTGCGACAGTGG
COR_fwd	CAAACTGAAGAGTGTCTTGGTGAAGCT
COR_rev	GGAGGACAAGATCAGCGTGAGCA

Table 31. Miscellaneous primers used in the study

Table of Abbreviations

2-ODD	2-oxoglutarate/Fe(II) dependent dioxygenase
BAC	Bacterial artificial chromosome
BBE	Berberine bridge enzyme
BIAs	Benzylisoquinolines alkaloids
bp/kb/Mb	Base pair/ kilo base pairs (1,000bp)/ mega base pairs (1,000,000bp)
CODM	Codeine-*O*-demethylase
COR	Codeinone reductase
DIOX	Dioxygenase
DNA/gDNA/cDNA/dsDNA	Deoxyribonucleic acid/ genomic DNA/ complementary DNA/ double stranded DNA
DW	Dry weight
EMS	Ethylmethane sulfonate
EST	Expressed sequence tag
FAD	Flavin adenine dinucleotide
FNM	Fast neutron mutagenesis
LTR	Long Terminal Repeat (Retrotransposons)
MAS	Marker assisted selection
NADPH	Nicotinamide adenine dinucleotide phosphate
ORF	Open reading frame
PARSESNP	Project Aligned Related Sequences and Evaluate SNPs
PODA	Protopine O-dealkylase
PCR/AS-PCR/RT-QPCR	Polymerase chain reaction/ Allele specific PCR/ Real time quantitative PCR
PDA	Personal Digital Assistant
QTL	Quantitative trait loci
RFU	Relative Fluorescence Units
RNA/mRNA	Ribonucleic acid/ messenger RNA
rpm	revolutions per minute
RT	Room temperature
SIFT	Sorting intolerant from tolerant
SNAP	Screening for non-acceptable polymorphisms
SNP	Single nucleotide polymorphism
STORR	(S)- to (R)-reticuline fusion protein
T6ODM	Thebaine-6-*O*-demethylase
TE	Tris-EDTA (Ethylenediaminetetraacetic acid)
TILLING	Targeting Induced Local Lesions in Genomes
Tint	T6ODMa intron (polymorphism)
UPLC	Ultra Performance Liquid Chromatography
VIGS	Virus Induced Gene Silencing
WT	Wild type

Bibliography

Abbott, A., 2015. Europe's genetically edited plants stuck in legal limbo. Nature 528, 319–320. doi:10.1038/528319a

Alcantara, J., Bird, D.A., Franceschi, V.R., Facchini, P.J., 2005. Sanguinarine biosynthesis is associated with the endoplasmic reticulum in cultured opium poppy cells after elicitor treatment. Plant Physiol. 138, 173.

Allen, R.S., Millgate, A.G., Chitty, J.A., Thisleton, J., Miller, J.A.C., Fist, A.J., Gerlach, W.L., Larkin, P.J., 2004. RNAi-mediated replacement of morphine with the nonnarcotic alkaloid reticuline in opium poppy. Nat. Biotechnol. 22, 1559–1566.

American Society of Health-System Pharmacists, 2015. Codeine Monograph for Professionals - Drugs.com [WWW Document]. URL http://www.drugs.com/monograph/codeine.html (accessed 2.9.15).

Aragon-Poce, F., Martınez-Fernández, E., Marquez-Espinos, C., Perez, A., Mora, R., Torres, L.M., 2002. History of opium, in: International Congress Series. Elsevier, pp. 19–21.

Askitopoulou, H., Ramoutsaki, I.A., Konsolaki, E., 2002. Archaeological evidence on the use of opium in the Minoan world, in: International Congress Series. Elsevier, pp. 23–29.

Barken, I., Geller, J., Rogosnitzky, M., 2008. Noscapine Inhibits Human Prostate Cancer Progression and Metastasis in a Mouse Model. Anticancer Res. 28, 3701–3704.

Beaudoin, G.A., Facchini, P.J., 2014. Benzylisoquinoline alkaloid biosynthesis in opium poppy. Planta 1–14.

Beaudoin, G.A.W., Facchini, P.J., 2013. Isolation and characterization of a cDNA encoding (S)-cis-N-methylstylopine 14-hydroxylase from opium poppy, a key enzyme in sanguinarine biosynthesis. Biochem. Biophys. Res. Commun. 431, 597–603. doi:10.1016/j.bbrc.2012.12.129

Belny, M., Hérouart, D., Thomasset, B., David, H., Jacquin-Dubreuil, A., David, A., 1997. Transformation of Papaver somniferum cell suspension cultures with sam1 from A. thaliana results in cell lines of different S-adenosyl-L-methionine synthetase activity. Physiol. Plant. 99, 233–240. doi:10.1111/j.1399-3054.1997.tb05407.x

Bennett, M.D., Smith, J.B., 1976. Nuclear DNA Amounts in Angiosperms. Philos. Trans. R. Soc. Lond. B Biol. Sci. 274, 227–274. doi:10.1098/rstb.1976.0044

Berenyi, S., Csutoras, C., Sipos, A., 2009. Recent Developments in the Chemistry of Thebaine and its Transformation Products as Pharmacological Targets. Curr. Med. Chem. 16, 3215–3242. doi:10.2174/092986709788803295

Bessi, I., Bazzicalupi, C., Richter, C., Jonker, H.R.A., Saxena, K., Sissi, C., Chioccioli, M., Bianco, S., Bilia, A.R., Schwalbe, H., Gratteri, P., 2012. Spectroscopic, Molecular Modeling, and NMR-Spectroscopic Investigation of the Binding Mode of the Natural Alkaloids Berberine and Sanguinarine to Human Telomeric G-Quadruplex DNA. ACS Chem. Biol. 7, 1109–1119. doi:10.1021/cb300096g

Bhandari, M.M., 1990. Out-crossing in opium poppy Papaver somniferum L. Euphytica 48, 167–169.

Bielecka, M., Kaminski, F., Adams, I., Poulson, H., Sloan, R., Li, Y., Larson, T.R., Winzer, T., Graham, I.A., 2014. Targeted mutation of Δ12 and Δ15 desaturase genes in hemp produce major alterations in seed fatty acid composition including a high oleic hemp oil. Plant Biotechnol. J. 12, 613–623. doi:10.1111/pbi.12167

Bird, D.A., Franceschi, V.R., Facchini, P.J., 2003. A tale of three cell types: alkaloid biosynthesis is localized to sieve elements in opium poppy. Plant Cell Online 15, 2626.

Bortesi, L., Fischer, R., 2015. The CRISPR/Cas9 system for plant genome editing and beyond. Biotechnol. Adv. 33, 41–52. doi:10.1016/j.biotechadv.2014.12.006

Botticella, E., Sestili, F., Hernandez-Lopez, A., Phillips, A., Lafiandra, D., 2011. High resolution melting analysis for the detection of EMS induced mutations in wheat SbeIIa genes. BMC Plant Biol. 11, 1–14. doi:10.1186/1471-2229-11-156

Boycheva, S., Daviet, L., Wolfender, J.-L., Fitzpatrick, T.B., 2014. The rise of operon-like gene clusters in plants. Trends Plant Sci. 19, 447–459. doi:10.1016/j.tplants.2014.01.013
Brochmann-Hanssen, E., 1984. A Second Pathway for the Terminal Steps in the Biosynthesis of Morphine. Planta Med. 50, 343–345. doi:10.1055/s-2007-969727
Bromberg, Y., Rost, B., 2007. SNAP: predict effect of non-synonymous polymorphisms on function. Nucleic Acids Res. 35, 3823–3835. doi:10.1093/nar/gkm238
Brownstein, M.J., 1993. A brief history of opiates, opioid peptides, and opioid receptors. Proc. Natl. Acad. Sci. U. S. A. 90, 5391–5393.
Chang, S., Puryear, J., Cairney, J., 1993. A simple and efficient method for isolating RNA from pine trees. Plant Mol. Biol. Report. 11, 113–116. doi:10.1007/BF02670468
Cha, R.S., Zarbl, H., Keohavong, P., Thilly, W.G., 1992. Mismatch amplification mutation assay (MAMA): application to the c-H-ras gene. Genome Res. 2, 14–20. doi:10.1101/gr.2.1.14
Chaterjee, A., Shukla, S., Mishra, P., Rastogi, A., Singh, S.P., 2010. Prospects of in vitro production of thebaine in opium poppy (Papaver somniferum L.). Ind. Crops Prod. 32, 668–670. doi:10.1016/j.indcrop.2010.04.007
Chauhan, S.P., Patra, N.K., 1993. Mutagenic Effects of Combined and Single Doses of Gamma Rays and EMS in Opium Poppy*. Plant Breed. 110, 342–345. doi:10.1111/j.1439-0523.1993.tb00600.x
Cheng, A.-X., Han, X.-J., Wu, Y.-F., Lou, H.-X., 2014. The Function and Catalysis of 2-Oxoglutarate-Dependent Oxygenases Involved in Plant Flavonoid Biosynthesis. Int. J. Mol. Sci. 15, 1080–1095. doi:10.3390/ijms15011080
Chen, X., Facchini, P.J., 2014. Short-chain dehydrogenase/reductase catalyzing the final step of noscapine biosynthesis is localized to laticifers in opium poppy. Plant J. 77, 173–184.
Chitty, J.A., Allen, R.S., Larkin, P.J., Fist, A.J., 2003. Genetic transformation in commercial Tasmanian cultivars of opium poppy, Papaver somniferum, and movement of transgenicpollen in the field. Funct. Plant Biol. 30, 1045–1058.
Cirak, C., Radusiene, J., Camas, N., Caliskan, O., Odabas, M.S., 2013. Changes in the contents of main secondary metabolites in two Turkish Hypericum species during plant development. Pharm. Biol. 51, 391–399. doi:10.3109/13880209.2012.733012
Clifton, I.J., McDonough, M.A., Ehrismann, D., Kershaw, N.J., Granatino, N., Schofield, C.J., 2006. Structural studies on 2-oxoglutarate oxygenases and related double-stranded β-helix fold proteins. J. Inorg. Biochem. 100, 644–669. doi:10.1016/j.jinorgbio.2006.01.024
Cline, S.D., Coscia, C.J., 1988. Stimulation of Sanguinarine Production by Combined Fungal Elicitation and Hormonal Deprivation in Cell Suspension Cultures of Papaver bracteatum. Plant Physiol. 86, 161–165. doi:10.1104/pp.86.1.161
Colbert, T., Till, B.J., Tompa, R., Reynolds, S., Steine, M.N., Yeung, A.T., McCallum, C.M., Comai, L., Henikoff, S., 2001. High-Throughput Screening for Induced Point Mutations. Plant Physiol. 126, 480–484.
Comai, L., Henikoff, S., 2006. TILLING: practical single-nucleotide mutation discovery. Plant J. 45, 684–694. doi:10.1111/j.1365-313X.2006.02670.x
Cooper, J.L., Till, B.J., Laport, R.G., Darlow, M.C., Kleffner, J.M., Jamai, A., El-Mellouki, T., Liu, S., Ritchie, R., Nielsen, N., Bilyeu, K.D., Meksem, K., Comai, L., Henikoff, S., 2008. TILLING to detect induced mutations in soybean. BMC Plant Biol. 8, 9. doi:10.1186/1471-2229-8-9
Cotterill, P., 2010. METHOD OF ALTERING THE ALKALOID COMPOSITION IN POPPY PLANTS.
Dean, G.J., 2011. Modifying poppy growth and alkaloid yield with plant growth regulators (cmaster). University of Tasmania.
DeLoache, W.C., Russ, Z.N., Narcross, L., Gonzales, A.M., Martin, V.J.J., Dueber, J.E., 2015. An enzyme-coupled biosensor enables (S)-reticuline production in yeast from glucose. Nat. Chem. Biol. doi:10.1038/nchembio.1816
Desgagne-Penix, I., Hagel, J.M., Facchini, P.J., 2009. Mutagenesis as a Functional Genomics Platform for Pharmaceutical Alkaloid Biosynthetic Gene Discovery in Opium Poppy.
Desgagné-Penix, I., Khan, M.F., Schriemer, D.C., Cram, D., Nowak, J., Facchini, P.J., 2010. Integration of deep transcriptome and proteome analyses reveals the components of alkaloid metabolism in opium poppy cell cultures. BMC Plant Biol. 10, 1–17.

Díaz Chávez, M.L., Rolf, M., Gesell, A., Kutchan, T.M., 2011. Characterization of two methylenedioxy bridge-forming cytochrome P450-dependent enzymes of alkaloid formation in the Mexican prickly poppy Argemone mexicana. Arch. Biochem. Biophys., P450 Catalysis Mechanisms 507, 186–193. doi:10.1016/j.abb.2010.11.016

Dicker, n.d. The Poppy Industry in Tasmania [WWW Document]. URL http://www.launc.tased.edu.au/online/sciences/agsci/alkalo/popindus.htm (accessed 2.19.15).

Dinsmore, W.W., 2005. Available and future treatments for erectile dysfunction. Clin. Cornerstone, Erectile Dysfunction 7, 37–44. doi:10.1016/S1098-3597(05)80047-X

Dong, C., Vincent, K., Sharp, P., 2009. Simultaneous mutation detection of three homoeologous genes in wheat by High Resolution Melting analysis and Mutation Surveyor®. BMC Plant Biol. 9, 143. doi:10.1186/1471-2229-9-143

Drenkard, E., Richter, B.G., Rozen, S., Stutius, L.M., Angell, N.A., Mindrinos, M., Cho, R.J., Oefner, P.J., Davis, R.W., Ausubel, F.M., 2000. A Simple Procedure for the Analysis of Single Nucleotide Polymorphisms Facilitates Map-Based Cloning in Arabidopsis. Plant Physiol. 124, 1483–1492. doi:10.1104/pp.124.4.1483

Dubey, M.K., Dhawan, O.P., Khanuja, S.P.S., 2008. Downy mildew resistance in opium poppy: resistance sources, inheritance pattern, genetic variability and strategies for crop improvement. Euphytica 165, 177–188. doi:10.1007/s10681-008-9804-4

Escobar, N.M., Haupt, S., Thow, G., Boevink, P., Chapman, S., Oparka, K., 2003. High-Throughput Viral Expression of cDNA–Green Fluorescent Protein Fusions Reveals Novel Subcellular Addresses and Identifies Unique Proteins That Interact with Plasmodesmata. Plant Cell 15, 1507–1523. doi:10.1105/tpc.013284

Facchini, P., Hagel, J., Liscombe, D., Loukanina, N., MacLeod, B., Samanani, N., Zulak, K., 2007. Opium poppy: blueprint for an alkaloid factory. Phytochem. Rev. 6, 97–124. doi:10.1007/s11101-006-9042-0

Facchini, P.J., Bird, D.A., 1998. Developmental regulation of benzylisoquinoline alkaloid biosynthesis in opium poppy plants and tissue cultures. Vitro Cell. Dev. Biol. - Plant 34, 69–79. doi:10.1007/BF02823126

Facchini, P.J., De Luca, V., 2008. Opium poppy and Madagascar periwinkle: model non-model systems to investigate alkaloid biosynthesis in plants. Plant J. 54, 763–784.

Facchini, P.J., De Luca, V., 1995. Phloem-specific expression of tyrosine/dopa decarboxylase genes and the biosynthesis of isoquinoline alkaloids in opium poppy. Plant Cell Online 7, 1811.

Facchini, P.J., Johnson, A.G., Poupart, J., De Luca, V., 1996a. Uncoupled defense gene expression and antimicrobial alkaloid accumulation in elicited opium poppy cell cultures. Plant Physiol. 111, 687.

Facchini, P.J., Loukanina, N., Blanche, V., 2008. Genetic transformation via somatic embryogenesis to establish herbicide-resistant opium poppy. Plant Cell Rep. 27, 719–727. doi:10.1007/s00299-007-0483-8

Facchini, P.J., Park, S.U., 2003. Developmental and inducible accumulation of gene transcripts involved in alkaloid biosynthesis in opium poppy. Phytochemistry 64, 177–186.

Facchini, P.J., Penzes, C., Johnson, A.G., Bull, D., 1996b. Molecular characterization of berberine bridge enzyme genes from opium poppy. Plant Physiol. 112, 1669.

FAO-IAEA, 2015. Mutation Enhanced Technologies for Agriculture (META) [WWW Document]. URL http://mvgs.iaea.org/Default.aspx (accessed 5.26.15).

Farrow, S.C., Facchini, P.J., 2015. Papaverine 7-O-demethylase, a novel 2-oxoglutarate/Fe2+-dependent dioxygenase from opium poppy. FEBS Lett. 589, 2701–2706. doi:10.1016/j.febslet.2015.07.042

Farrow, S.C., Facchini, P.J., 2013. Dioxygenases catalyze O-demethylation and O, O-demethylenation with widespread roles in benzylisoquinoline alkaloid metabolism in opium poppy. J. Biol. Chem. 288, 28997–29012.

Farrow, S.C., Hagel, J.M., Beaudoin, G.A.W., Burns, D.C., Facchini, P.J., 2015. Stereochemical inversion of (S)-reticuline by a cytochrome P450 fusion in opium poppy. Nat. Chem. Biol. advance online publication. doi:10.1038/nchembio.1879

Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D.-L., Wei, P., Cao, F., Zhu, S., Zhang, F., Mao, Y., Zhu, J.-K., 2013. Efficient genome editing in plants using a CRISPR/Cas system. Cell Res. 23, 1229–1232. doi:10.1038/cr.2013.114

Field, B., Fiston-Lavier, A.-S., Kemen, A., Geisler, K., Quesneville, H., Osbourn, A.E., 2011. Formation of plant metabolic gene clusters within dynamic chromosomal regions. Proc. Natl. Acad. Sci. 108, 16116–16121. doi:10.1073/pnas.1109273108

Fisinger, U., Grobe, N., Zenk, M.H., 2007. Thebaine synthase: a new enzyme in the morphine pathway in Papaver somniferum. Nat. Prod. Commun. 2, 249–253.

Fist, A., Byme, C., Gerlach, W., Sayer, C., Bailey, T., 2005. Production of reticuline. US20050257295 A1.

Fist, A.J., 2001. The Tasmanian Poppy Industry: A Case Study of the Application of Science and Technology.

Fist, A.J., Byrne, C.J., Gerlach, W.L., 2000. Papaver somniferum strain with high concentration of thebaine and oripavine. US6067749 A.

Fist, A.J., Miller, J.A.C., Gregory, D., 2013. Papaver somniferum with high concentration of codeine. US8541647 B2.

Fist, A.J., Miller, J.A.C., Gregory, D., 2010. Papaver somniferum with high concentration of codeine. US20100234600 A1.

Flem-Bonhomme, V.L., Laurain-Mattar, D., Fliniaux, M.A., 2004. Hairy root induction of Papaver somniferum var. album, a difficult-to-transform plant, by A. rhizogenes LBA 9402. Planta 218, 890–893. doi:10.1007/s00425-003-1196-z

Floria, F.G., Ghiorghita, G.I., 1980. The influence of the treatment with alkylating agents on Papaver somniferum L., in M1. Rev. Roum. Biol. Biol. Veg. 25, 151–155.

Fossati, E., Ekins, A., Narcross, L., Zhu, Y., Falgueyret, J.-P., Beaudoin, G.A., Facchini, P.J., Martin, V.J., 2014. Reconstitution of a 10-gene pathway for synthesis of the plant alkaloid dihydrosanguinarine in Saccharomyces cerevisiae. Nat. Commun. 5.

Fossati, E., Narcross, L., Ekins, A., Falgueyret, J.-P., Martin, V.J.J., 2015. Synthesis of Morphinan Alkaloids in Saccharomyces cerevisiae. PLoS ONE 10, e0124459. doi:10.1371/journal.pone.0124459

Francia, E., Tacconi, G., Crosatti, C., Barabaschi, D., Bulgarelli, D., Dall'Aglio, E., Valè, G., 2005. Marker assisted selection in crop plants. Plant Cell Tissue Organ Cult. 82, 317–342. doi:10.1007/s11240-005-2387-z

Frenzel, T., Zenk, M.H., 1990. S-adenosyl-l-methionine: 3′-hydroxy-N-methyl-(S)-coclaurine-4′-O-methyl transferase, a regio- and stereoselective enzyme of the (S)-reticuline pathway. Phytochemistry 29, 3505–3511. doi:10.1016/0031-9422(90)85265-H

Frick, S., Chitty, J.A., Kramell, R., Schmidt, J., Allen, R.S., Larkin, P.J., Kutchan, T.M., 2004. Transformation of opium poppy (Papaver somniferumL.) with antisense berberine bridge enzyme gene (anti-bbe) via somatic embryogenesis results in an altered ratio of alkaloids in latex but not in roots. Transgenic Res. 13, 607–613.

Frick, S., Kramell, R., Schmidt, J., Fist, A.J., Kutchan, T.M., 2005. Comparative Qualitative and Quantitative Determination of Alkaloids in Narcotic and Condiment Papaver somniferum Cultivars. J. Nat. Prod. 68, 666–673. doi:10.1021/np0496643

Gady, A.L., Hermans, F.W., Van de Wal, M.H., van Loo, E.N., Visser, R.G., Bachem, C.W., 2009. Implementation of two high through-put techniques in a novel application: detecting point mutations in large EMS mutated plant populations. Plant Methods 5, 13. doi:10.1186/1746-4811-5-13

Galanie, S., Thodey, K., Trenchard, I.J., Interrante, M.F., Smolke, C.D., 2015. Complete biosynthesis of opioids in yeast. Science 349, 1095–1100. doi:10.1126/science.aac9373

Gates, M., Tschudi, G., 1952. The Synthesis of Morphine. J. Am. Chem. Soc. 74, 1109–1110.

Gebhardt, Y., Witte, S., Forkmann, G., Lukačin, R., Matern, U., Martens, S., 2005. Molecular evolution of flavonoid dioxygenases in the family Apiaceae. Phytochemistry, Evolution of Metabolic Diversity 66, 1273–1284. doi:10.1016/j.phytochem.2005.03.030

Gesell, A., Rolf, M., Ziegler, J., Díaz Chávez, M.L., Huang, F.-C., Kutchan, T.M., 2009. CYP719B1 Is Salutaridine Synthase, the C-C Phenol-coupling Enzyme of Morphine Biosynthesis in Opium Poppy. J. Biol. Chem. 284, 24432 –24442. doi:10.1074/jbc.M109.033373
Glenn, W.S., Runguphan, W., O'Connor, S.E., 2013. Recent progress in the metabolic engineering of alkaloids in plant systems. Curr. Opin. Biotechnol. 24, 354–365. doi:10.1016/j.copbio.2012.08.003
GMO Review Submission No. 158 - GSK Australia - Department of Primary Industries, Parks, Water and Environment [WWW Document], 2013. URL http://dpipwe.tas.gov.au/Pages/document.aspx?path=/Documents/GMO-Review-Submission-158-GSK-Submission.pdf (accessed 2.10.15).
Gorres, K.L., Raines, R.T., 2010. Prolyl 4-hydroxylase. Crit. Rev. Biochem. Mol. Biol. 45, 106–124. doi:10.3109/10409231003627991
Gözler, B., Lantz, M.S., Shamma, M., 1983. The Pavine and Isopavine Alkaloids. J. Nat. Prod. 46, 293–309. doi:10.1021/np50027a001
Graham, I.A., Besser, K., Blumer, S., Branigan, C.A., Czechowski, T., Elias, L., Guterman, I., Harvey, D., Isaac, P.G., Khan, A.M., Larson, T.R., Li, Y., Pawson, T., Penfield, T., Rae, A.M., Rathbone, D.A., Reid, S., Ross, J., Smallwood, M.F., Segura, V., Townsend, T., Vyas, D., Winzer, T., Bowles, D., 2010. The Genetic Map of Artemisia annua L. Identifies Loci Affecting Yield of the Antimalarial Drug Artemisinin. Science 327, 328–331. doi:10.1126/science.1182612
Greene, B., Walko, R., Hake, S., 1994. Mutator insertions in an intron of the maize knotted1 gene result in dominant suppressible mutations. Genetics 138, 1275–1285.
Greene, E.A., Codomo, C.A., Taylor, N.E., Henikoff, J.G., Till, B.J., Reynolds, S.H., Enns, L.C., Burtner, C., Johnson, J.E., Odden, A.R., Comai, L., Henikoff, S., 2003. Spectrum of Chemically Induced Mutations From a Large-Scale Reverse-Genetic Screen in Arabidopsis. Genetics 164, 731 –740.
Grothe, T., Lenz, R., Kutchan, T.M., 2001. Molecular Characterization of the Salutaridinol 7-O-Acetyltransferase Involved in Morphine Biosynthesis in Opium Poppy Papaver somniferum. J. Biol. Chem. 276, 30717–30723. doi:10.1074/jbc.M102688200
Hagel, J.M., Beaudoin, G.A., Fossati, E., Ekins, A., Martin, V.J., Facchini, P.J., 2012. Characterization of a flavoprotein oxidase from opium poppy catalyzing the final steps in sanguinarine and papaverine biosynthesis. J. Biol. Chem. 287, 42972–42983.
Hagel, J.M., Facchini, P.J., 2013. Benzylisoquinoline alkaloid metabolism – a century of discovery and a brave new world. Plant Cell Physiol. doi:10.1093/pcp/pct020
Hagel, J.M., Facchini, P.J., 2010a. Dioxygenases catalyze the O-demethylation steps of morphine biosynthesis in opium poppy. Nat. Chem. Biol. 6, 273–275.
Hagel, J.M., Facchini, P.J., 2010b. Biochemistry and occurrence of O-demethylation in plant metabolism. Front. Physiol. 1.
Hagel, J.M., Weljie, A.M., Vogel, H.J., Facchini, P.J., 2008a. Quantitative 1H nuclear magnetic resonance metabolite profiling as a functional genomics platform to investigate alkaloid biosynthesis in opium poppy. Plant Physiol. 147, 1805.
Hagel, J.M., Yeung, E.C., Facchini, P.J., 2008b. Got milk? The secret life of laticifers. Trends Plant Sci. 13, 631–639.
Hansen, B.G., Kerwin, R.E., Ober, J.A., Lambrix, V.M., Mitchell-Olds, T., Gershenzon, J., Halkier, B.A., Kliebenstein, D.J., 2008. A Novel 2-Oxoacid-Dependent Dioxygenase Involved in the Formation of the Goiterogenic 2-Hydroxybut-3-enyl Glucosinolate and Generalist Insect Resistance in Arabidopsis. Plant Physiol. 148, 2096–2108. doi:10.1104/pp.108.129981
Hartmann, T., 1999. Chemical ecology of pyrrolizidine alkaloids. Planta 207, 483–495. doi:10.1007/s004250050508
Hartmann, T., Ober, D., 2000. Biosynthesis and Metabolism of Pyrrolizidine Alkaloids in Plants and Specialized Insect Herbivores, in: Leeper, F., Vederas, J. (Eds.), Biosynthesis, Topics in Current Chemistry. Springer Berlin / Heidelberg, pp. 207–243.
Hartung, F., Schiemann, J., 2014. Precise plant breeding using new genome editing techniques: opportunities, safety and regulation in the EU. Plant J. 78, 742–752. doi:10.1111/tpj.12413

Hashimoto, T., Yamada, Y., 1986. Hyoscyamine 6β-Hydroxylase, a 2-Oxoglutarate-Dependent Dioxygenase, in Alkaloid-Producing Root Cultures. Plant Physiol. 81, 619–625. doi:10.1104/pp.81.2.619
Hawkins, K.M., Smolke, C.D., 2008. Production of benzylisoquinoline alkaloids in Saccharomyces cerevisiae. Nat. Chem. Biol. 4, 564–573. doi:10.1038/nchembio.105
Hebsgaard, S.M., Korning, P.G., Tolstrup, N., Engelbrecht, J., Rouzé, P., Brunak, S., 1996. Splice site prediction in Arabidopsis thaliana pre-mRNA by combining local and global sequence information. Nucleic Acids Res. 24, 3439–3452.
Henikoff, S., Comai, L., 2003. Single-Nucleotide Mutations for Plant Functional Genomics. Annu. Rev. Plant Biol. 54, 375–401. doi:10.1146/annurev.arplant.54.031902.135009
Huang, F.C., Kutchan, T.M., 2000. Distribution of morphinan and benzo [c] phenanthridine alkaloid gene transcript accumulation in Papaver somniferum. Phytochemistry 53, 555–564.
Ikezawa, N., Iwasa, K., Sato, F., 2007. Molecular cloning and characterization of methylenedioxy bridge-forming enzymes involved in stylopine biosynthesis in Eschscholzia californica. FEBS J. 274, 1019–1035. doi:10.1111/j.1742-4658.2007.05652.x
Ikezawa, N., Tanaka, M., Nagayoshi, M., Shinkyo, R., Sakaki, T., Inouye, K., Sato, F., 2003. Molecular Cloning and Characterization of CYP719, a Methylenedioxy Bridge-forming Enzyme That Belongs to a Novel P450 Family, from cultured Coptis japonica Cells. J. Biol. Chem. 278, 38557–38565. doi:10.1074/jbc.M302470200
INCB, 2015. Narcotic Drugs Report 2014-Estimated World Requirements for 2015.
INCB, 2010. INCB Annual Report 2010.
Itkin, M., Heinig, U., Tzfadia, O., Bhide, A.J., Shinde, B., Cardenas, P.D., Bocobza, S.E., Unger, T., Malitsky, S., Finkers, R., Tikunov, Y., Bovy, A., Chikate, Y., Singh, P., Rogachev, I., Beekwilder, J., Giri, A.P., Aharoni, A., 2013. Biosynthesis of Antinutritional Alkaloids in Solanaceous Crops Is Mediated by Clustered Genes. Science 341, 175–179. doi:10.1126/science.1240230
Jain, S.M., Gupta, S.D., 2013. Biotechnology of Neglected and Underutilized Crops. Springer Science & Business Media.
Jiang, G.-L., Dong, Y., Shi, J., Ward, R.W., 2007. QTL analysis of resistance to Fusarium head blight in the novel wheat germplasm CJ 9306. II. Resistance to deoxynivalenol accumulation and grain yield loss. Theor. Appl. Genet. 115, 1043–1052. doi:10.1007/s00122-007-0630-1
Jonczyk, R., Schmidt, H., Osterrieder, A., Fiesselmann, A., Schullehner, K., Haslbeck, M., Sicker, D., Hofmann, D., Yalpani, N., Simmons, C., Frey, M., Gierl, A., 2008. Elucidation of the Final Reactions of DIMBOA-Glucoside Biosynthesis in Maize: Characterization of Bx6 and Bx7. Plant Physiol. 146, 1053–1063. doi:10.1104/pp.107.111237
Jones, M.O., Piron-Prunier, F., Marcel, F., Piednoir-Barbeau, E., Alsadon, A.A., Wahb-Allah, M.A., Al-Doss, A.A., Bowler, C., Bramley, P.M., Fraser, P.D., Bendahmane, A., 2012. Characterisation of alleles of tomato light signalling genes generated by TILLING. Phytochemistry 79, 78–86. doi:10.1016/j.phytochem.2012.04.005
Kai, K., Mizutani, M., Kawamura, N., Yamamoto, R., Tamai, M., Yamaguchi, H., Sakata, K., Shimizu, B., 2008. Scopoletin is biosynthesized via ortho-hydroxylation of feruloyl CoA by a 2-oxoglutarate-dependent dioxygenase in Arabidopsis thaliana. Plant J. 55, 989–999. doi:10.1111/j.1365-313X.2008.03568.x
Kassell, N.F., Helm, G., Simmons, N., Phillips, C.D., Cail, W.S., 1992. Treatment of cerebral vasospasm with intra-arterial papaverine. J. Neurosurg. 77, 848–852. doi:10.3171/jns.1992.77.6.0848
Kato, N., Dubouzet, E., Kokabu, Y., Yoshida, S., Taniguchi, Y., Dubouzet, J.G., Yazaki, K., Sato, F., 2007. Identification of a WRKY Protein as a Transcriptional Regulator of Benzylisoquinoline Alkaloid Biosynthesis in Coptis japonica. Plant Cell Physiol. 48, 8–18. doi:10.1093/pcp/pcl041
Kaul, B.L., Tandon, V., Choudhary, D.K., 1979. Cytogenetic studies inPapaver somniferum L. Proc. Indian Acad. Sci. - Sect. B Part 2 Plant Sci. 88, 321–325. doi:10.1007/BF03046197

Kawai, Y., Ono, E., Mizutani, M., 2014. Evolution and diversity of the 2–oxoglutarate-dependent dioxygenase superfamily in plants. Plant J. 78, 328–343. doi:10.1111/tpj.12479
Khan, R., Khan, M.M.A., Singh, M., Nasir, S., Naeem, M., Siddiqui, M.H., Mohammad, F., 2007. Gibberellic acid and triacontanol can ameliorate the opium yield and morphine production in opium poppy (Papaver somniferum L.). Acta Agric. Scand. Sect. B — Soil Plant Sci. 57, 307–312. doi:10.1080/09064710600982811
Kliebenstein, D.J., Lambrix, V.M., Reichelt, M., Gershenzon, J., Mitchell-Olds, T., 2001. Gene Duplication in the Diversification of Secondary Metabolism: Tandem 2-Oxoglutarate–Dependent Dioxygenases Control Glucosinolate Biosynthesis in Arabidopsis. Plant Cell Online 13, 681–693. doi:10.1105/tpc.13.3.681
Kohany, O., Gentles, A.J., Hankus, L., Jurka, J., 2006. Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. BMC Bioinformatics 7, 474. doi:10.1186/1471-2105-7-474
Kramlinger, V.M., Rojas, M.A., Kanamori, T., Guengerich, F.P., 2015. Cytochrome P450 3A enzymes catalyze the O6-demethylation of thebaine, a key step in endogenous mammalian morphine biosynthesis. J. Biol. Chem. jbc.M115.665331. doi:10.1074/jbc.M115.665331
Krist, S., Stuebiger, G., Unterweger, H., Bandion, F., Buchbauer, G., 2005. Analysis of Volatile Compounds and Triglycerides of Seed Oils Extracted from Different Poppy Varieties (Papaver somniferum L.). J. Agric. Food Chem. 53, 8310–8316. doi:10.1021/jf0580869
Krokida, A., Delis, C., Geisler, K., Garagounis, C., Tsikou, D., Peña-Rodríguez, L.M., Katsarou, D., Field, B., Osbourn, A.E., Papadopoulou, K.K., 2013. A metabolic gene cluster in Lotus japonicus discloses novel enzyme functions and products in triterpene biosynthesis. New Phytol. 200, 675–690. doi:10.1111/nph.12414
Krug, E., Proksch, P., 1993. Influence of dietary alkaloids on survival and growth of Spodoptera littoralis. Biochem. Syst. Ecol. 21, 749–756. doi:10.1016/0305-1978(93)90087-8
Kumar, B., Patra, N.K., 2010. Gene frequency-based estimation of natural outcrossing in opium poppy (Papaver somniferum L.). Mol. Breed. 26, 619–626.
Kurowska, M., Daszkowska-Golec, A., Gruszka, D., Marzec, M., Szurman, M., Szarejko, I., Maluszynski, M., 2011. TILLING - a shortcut in functional genomics. J. Appl. Genet. 52, 371–390. doi:10.1007/s13353-011-0061-1
Kutchan, T., 1995. Alkaloid Biosynthesis-]The Basis for Metabolic Engineering of Medicinal Plants. Plant Cell 7, 1059–1070.
Lam, K.C., Ibrahim, R.K., Behdad, B., Dayanandan, S., 2007. Structure, function, and evolution of plant O -methyltransferases. Genome 50, 1001–1013. doi:10.1139/G07-077
Lavania, U., Srivastava, S., 1999. Quantitative delineation of karyotype variation in Papaver as a measure of phylogenetic differentiation and origin. Curr. Sci. 77, 429–435.
Lee, E.-J., Facchini, P., 2010. Norcoclaurine Synthase Is a Member of the Pathogenesis-Related 10/Bet V1 Protein Family. Plant Cell Online 22, 3489–3503. doi:10.1105/tpc.110.077958
Le Signor, C., Savois, V., Aubert, G., Verdier, J., Nicolas, M., Pagny, G., Moussy, F., Sanchez, M., Baker, D., Clarke, J., Thompson, R., 2009. Optimizing TILLING populations for reverse genetics in Medicago truncatula. Plant Biotechnol. J. 7, 430–441. doi:10.1111/j.1467-7652.2009.00410.x
Li, J.-F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M., Sheen, J., 2013. Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. Nat. Biotechnol. 31, 688–691. doi:10.1038/nbt.2654
Li, M., Kuivenhoven, J.A., Ayyobi, A.F., Pritchard, P.H., 1998. T→G or T→A mutation introduced in the branchpoint consensus sequence of intron 4 of lecithin:cholesterol acyltransferase (LCAT) gene: intron retention causing LCAT deficiency. Biochim.

Biophys. Acta BBA - Lipids Lipid Metab. 1391, 256–264. doi:10.1016/S0005-2760(97)00198-7
Linné, C. von, Salvius, L., 1753. Caroli Linnaei ... Species plantarum :exhibentes plantas rite cognitas, ad genera relatas, cum differentiis specificis, nominibus trivialibus, synonymis selectis, locis natalibus, secundum systema sexuale digestas... Impensis Laurentii Salvii, Holmiae :
Lisch, D., 2013. How important are transposons for plant evolution? Nat. Rev. Genet. 14, 49–61. doi:10.1038/nrg3374
Liscombe, D.K., Facchini, P.J., 2008. Evolutionary and cellular webs in benzylisoquinoline alkaloid biosynthesis. Curr. Opin. Biotechnol., Food biotechnology / Plant biotechnology 19, 173–180. doi:10.1016/j.copbio.2008.02.012
Liscombe, D.K., Facchini, P.J., 2007. Molecular cloning and characterization of tetrahydroprotoberberine cis-N-methyltransferase, an enzyme involved in alkaloid biosynthesis in opium poppy. J. Biol. Chem. 282, 14741.
Liscombe, D.K., MacLeod, B.P., Loukanina, N., Nandi, O.I., Facchini, P.J., 2005. Evidence for the monophyletic evolution of benzylisoquinoline alkaloid biosynthesis in angiosperms. Phytochemistry 66, 1374–1393.
Liu, H., Wang, J., Zhao, J., Lu, S., Wang, J., Jiang, W., Ma, Z., Zhou, L., 2009. Isoquinoline alkaloids from Macleaya cordata active against plant microbial pathogens. Nat. Prod. Commun. 4, 1557–1560.
Livak, K.J., Schmittgen, T.D., 2001. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. Methods 25, 402–408. doi:10.1006/meth.2001.1262
Li, X., Han, Y., Teng, W., Zhang, S., Yu, K., Poysa, V., Anderson, T., Ding, J., Li, W., 2010. Pyramided QTL underlying tolerance to Phytophthora root rot in mega-environments from soybean cultivars “Conrad” and “Hefeng 25.” Theor. Appl. Genet. 121, 651–658. doi:10.1007/s00122-010-1337-2
Li, X., Lassner, M., Zhang, Y., 2002. Deleteagene: A Fast Neutron Deletion Mutagenesis-Based Gene Knockout System for Plants. Comp. Funct. Genomics 3, 158–160. doi:10.1002/cfg.148
Li, X., Zhang, Y., 2002. Reverse genetics by fast neutron mutagenesis in higher plants. Funct. Integr. Genomics 2, 254–258. doi:10.1007/s10142-002-0076-0
Li, Y.-D., Chu, Z.-Z., Liu, X.-G., Jing, H.-C., Liu, Y.-G., Hao, D.-Y., 2010. A Cost-effective High-resolution Melting Approach using the EvaGreen Dye for DNA Polymorphism Detection and Genotyping in Plants. J. Integr. Plant Biol. 52, 1036–1042. doi:10.1111/j.1744-7909.2010.01001.x
Loenarz, C., Schofield, C.J., 2008. Expanding chemical biology of 2-oxoglutarate oxygenases. Nat. Chem. Biol. 4, 152–156. doi:10.1038/nchembio0308-152
Luch, A., 2009. Molecular, Clinical and Environmental Toxicology: Volume 1: Molecular Toxicology. Springer Science & Business Media.
Mahlberg, P., 1993. Laticifers: An historical perspective. Bot. Rev. 59, 1–23. doi:10.1007/BF02856611
Mahmoudian, M., Rahimi-Moghaddam, P., 2009. The Anti-Cancer Activity of Noscapine: A Review. Recent Patents Anticancer Drug Discov. 4, 92–97. doi:10.2174/157489209787002524
Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C., Bryant, S.H., 2011. CDD: a Conserved Domain Database for the functional annotation of proteins. Nucleic Acids Res. 39, D225–D229. doi:10.1093/nar/gkq1189
Martín, B., Ramiro, M., Martínez-Zapater, J.M., Alonso-Blanco, C., 2009. A high-density collection of EMS-induced mutations for TILLING in Landsberg erecta genetic background of Arabidopsis. BMC Plant Biol. 9, 147. doi:10.1186/1471-2229-9-147
Matsumoto, S., Mizutani, M., Sakata, K., Shimizu, B.-I., 2012. Molecular cloning and functional analysis of the ortho-hydroxylases of p-coumaroyl coenzyme A/feruloyl coenzyme A

involved in formation of umbelliferone and scopoletin in sweet potato, Ipomoea batatas (L.) Lam. Phytochemistry 74, 49–57. doi:10.1016/j.phytochem.2011.11.009
McCallum, C.M., Comai, L., Greene, E.A., Henikoff, S., 2000. Targeted screening for induced mutations. Nat. Biotechnol. 18, 455–457. doi:10.1038/74542
Meißner, C., 1819. Über Pflanzenalkalien: II. Über ein neues Pflanzenalkali (Alkaloid). J. Für Chem. Phys. 25, 379–381.
Merck, G., 1848. Annalen der Chemie und Pharmacie: vereinigte Zeitschrift des Neuen Journals der Pharmacie für Ärzte, Apotheker und Chemiker u. des Magazins für Pharmacie und Experimentalkritik. Winter.
Meza, T.J., Moen, M.N., Vågbø, C.B., Krokan, H.E., Klungland, A., Grini, P.E., Falnes, P.Ø., 2012. The DNA dioxygenase ALKBH2 protects Arabidopsis thaliana against methylation damage. Nucleic Acids Res. 40, 6620–6631. doi:10.1093/nar/gks327
Millgate, A.G., Pogson, B.J., Wilson, I.W., Kutchan, T.M., Zenk, M.H., Gerlach, W.L., Fist, A.J., Larkin, P.J., 2004. Analgesia: Morphine-pathway block in top1 poppies. Nature 431, 413–414. doi:10.1038/431413a
Minami, H., Dubouzet, E., Iwasa, K., Sato, F., 2007. Functional Analysis of Norcoclaurine Synthase in Coptis japonica. J. Biol. Chem. 282, 6274–6282. doi:10.1074/jbc.M608933200
Mishra, S., Triptahi, V., Singh, S., Phukan, U.J., Gupta, M.M., Shanker, K., Shukla, R.K., 2013. Wound Induced Tanscriptional Regulation of Benzylisoquinoline Pathway and Characterization of Wound Inducible PsWRKY Transcription Factor from Papaver somniferum. PLoS ONE 8, e52784. doi:10.1371/journal.pone.0052784
Missirian, V., Comai, L., Filkov, V., 2011. Statistical Mutation Calling from Sequenced Overlapping DNA Pools in TILLING Experiments. BMC Bioinformatics 12, 287. doi:10.1186/1471-2105-12-287
Montgomery, C.T., Cassels, B.K., Shamma, M., 1983. The Rhoeadine Alkaloids. J. Nat. Prod. 46, 441–453. doi:10.1021/np50028a001
Morimoto, S., Suemori, K., Moriwaki, J., Taura, F., Tanaka, H., Aso, M., Tanaka, M., Suemune, H., Shimohigashi, Y., Shoyama, Y., 2001. Morphine metabolism in the opium poppy and its possible physiological function - Biochemical characterization of the morphine metabolite, bismorphine. J. Biol. Chem. 276, 38179–38184.
Morimoto, S., Suemori, K., Taura, F., Shoyama, Y., 2003. New dimeric morphine from opium poppy (Papaver somuniferum) and its physiological function. J. Nat. Prod. 66, 987–989. doi:10.1021/np0205831
Morrell, P.L., Buckler, E.S., Ross-Ibarra, J., 2012. Crop genomics: advances and applications. Nat. Rev. Genet. 13, 85–96. doi:10.1038/nrg3097
Mosammaparast, N., Shi, Y., 2010. Reversal of Histone Methylation: Biochemical and Molecular Mechanisms of Histone Demethylases. Annu. Rev. Biochem. 79, 155–179. doi:10.1146/annurev.biochem.78.070907.103946
Munsey, C., 2005. HEROIN® and ASPIRIN® The Connection! & The Collection!-Part I.
Muth, J., Hartje, S., Twyman, R.M., Hofferbert, H.-R., Tacke, E., Prüfer, D., 2008. Precision breeding for novel starch variants in potato. Plant Biotechnol. J. 6, 576–584. doi:10.1111/j.1467-7652.2008.00340.x
Naito, K., Zhang, F., Tsukiyama, T., Saito, H., Hancock, C.N., Richardson, A.O., Okumoto, Y., Tanisaka, T., Wessler, S.R., 2009. Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. Nature 461, 1130–1134. doi:10.1038/nature08479
Nakajima, N., Mori, H., Yamazaki, K., Imaseki, H., 1990. Molecular Cloning and Sequence of a Complementary DNA Encoding 1-Aminocyclopropane-l-carboxylate Synthase Induced by Tissue Wounding. Plant Cell Physiol. 31, 1021–1029.
Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J.D.G., Kamoun, S., 2013. Targeted mutagenesis in the model plant Nicotiana benthamiana using Cas9 RNA-guided endonuclease. Nat. Biotechnol. 31, 691–693. doi:10.1038/nbt.2655
Nessler, C.L., Allen, R.D., Galewsky, S., 1985. Identification and Characterization of Latex-Specific Proteins in Opium Poppy 1. Plant Physiol. 79, 499–504.

Nessler, C.L., Mahlberg, P.G., 1979. Ultrastructure of laticifers in redifferentiated organs on callus from Papaver somniferum (Papaveraceae). Can. J. Bot. 57, 675–685. doi:10.1139/b79-086
Nessler, C.L., Mahlberg, P.G., 1977. Ontogeny and Cytochemistry of Alkaloidal Vesicles in Laticifers of Papaver somniferum L. (Papaveraceae). Am. J. Bot. 64, 541–551. doi:10.2307/2442002
Ng, P.C., Henikoff, S., 2003. SIFT: predicting amino acid changes that affect protein function. Nucleic Acids Res. 31, 3812–3814. doi:10.1093/nar/gkg509
Nielson, B., Röe, J., Brochmann-Hanssen, E., 1983. Oripavine - a new opium alkaloid. Planta Med. 48, 205–206.
Ober, D., 2005. Seeing double: gene duplication and diversification in plant secondary metabolism. Trends Plant Sci. 10, 444–449. doi:10.1016/j.tplants.2005.07.007
Ober, D., Hartmann, T., 2000. Phylogenetic origin of a secondary pathway: the case of pyrrolizidine alkaloids. Plant Mol. Biol. 44, 445–450.
Okabe, Y., Asamizu, E., Saito, T., Matsukura, C., Ariizumi, T., Brès, C., Rothan, C., Mizoguchi, T., Ezura, H., 2011. Tomato TILLING Technology: Development of a Reverse Genetics Tool for the Efficient Isolation of Mutants from Micro-Tom Mutant Libraries. Plant Cell Physiol. 52, 1994–2005. doi:10.1093/pcp/pcr134
Onoyovwe, A., Hagel, J.M., Chen, X., Khan, M.F., Schriemer, D.C., Facchini, P.J., 2013. Morphine biosynthesis in opium poppy involves two cell types: sieve elements and laticifers. Plant Cell Online 25, 4110–4122.
Paddon, C.J., Westfall, P.J., Pitera, D.J., Benjamin, K., Fisher, K., McPhee, D., Leavell, M.D., Tai, A., Main, A., Eng, D., Polichuk, D.R., Teoh, K.H., Reed, D.W., Treynor, T., Lenihan, J., Jiang, H., Fleck, M., Bajad, S., Dang, G., Dengrove, D., Diola, D., Dorin, G., Ellens, K.W., Fickes, S., Galazzo, J., Gaucher, S.P., Geistlinger, T., Henry, R., Hepp, M., Horning, T., Iqbal, T., Kizer, L., Lieu, B., Melis, D., Moss, N., Regentin, R., Secrest, S., Tsuruta, H., Vazquez, R., Westblade, L.F., Xu, L., Yu, M., Zhang, Y., Zhao, L., Lievense, J., Covello, P.S., Keasling, J.D., Reiling, K.K., Renninger, N.S., Newman, J.D., 2013. High-level semi-synthetic production of the potent antimalarial artemisinin. Nature 496, 528–532. doi:10.1038/nature12051
Panicker, S., Wojnoa, H.L., Ziska, L.H., 2007. Quantitation of the Major Alkaloids in Opium from Papaver Setigerum DC. Microgram 13.
Park, I.-K., Lee, H.-S., Lee, S.-G., Park, J.-D., Ahn, Y.-J., 2000. Antifeeding Activity of Isoquinoline Alkaloids Identified in Coptis japonica Roots Against Hyphantria cunea (Lepidoptera: Arctiidae) and Agelastica coerulea (Coleoptera: Galerucinae). J. Econ. Entomol. 93, 331–335. doi:10.1603/0022-0493-93.2.331
Park, S.U., Facchini, P., 2000a. Agrobacterium-mediated transformation of opium poppy, Papaver somniferum, via shoot organogenesis. J. Plant Physiol. 157, 207–214.
Park, S.U., Facchini, P.J., 2000b. Agrobacterium rhizogenes-mediated transformation of opium poppy, Papaver somniferum L., and California poppy, Eschscholzia californica Cham., root cultures. J. Exp. Bot. 51, 1005.
Pathak, S., Mishra, B., Misra, P., Misra, P., Joshi, V., Shukla, S., Trivedi, P., 2012. High frequency somatic embryogenesis, regeneration and correlation of alkaloid biosynthesis with gene expression in <i>Papaver somniferum</i> Plant Growth Regul. 68, 17–25. doi:10.1007/s10725-012-9689-z
Patra, N.K., Ram, R.S., Chauhan, S.P., Singh, A.K., 1992. Quantitative studies on the mating system of opium poppy (Papaver somniferum L.). Theor. Appl. Genet. 84, 299–302.
Pauli, H.H., Kutchan, T.M., 1998. Molecular cloning and functional heterologous expression of two alleles encoding (S)-N-methylcoclaurine 3′-hydroxylase (CYP80B1), a new methyl jasmonate-inducible cytochrome P-450-dependent mono-oxygenase of benzylisoquinoline alkaloid biosynthesis. Plant J. 13, 793–801. doi:10.1046/j.1365-313X.1998.00085.x
Phillips, A.L., 1998. Gibberellins in Arabidopsis. Plant Physiol. Biochem., Arabidopsis thaliana 36, 115–124. doi:10.1016/S0981-9428(98)80096-X

Podevin, N., Davies, H.V., Hartung, F., Nogué, F., Casacuberta, J.M., 2013. Site-directed nucleases: a paradigm shift in predictable, knowledge-based plant breeding. Trends Biotechnol. 31, 375–383. doi:10.1016/j.tibtech.2013.03.004
Qian, J.-Q., 2002. Cardiovascular pharmacological effects of bisbenzylisoquinoline alkaloid derivatives. Acta Pharmacol Sin 23, 1086–1092.
Qi, X., Bakht, S., Leggett, M., Maxwell, C., Melton, R., Osbourn, A., 2004. A gene cluster for secondary metabolism in oat: Implications for the evolution of metabolic diversity in plants. Proc. Natl. Acad. Sci. U. S. A. 101, 8233–8238. doi:10.1073/pnas.0401301101
Raghavan, C., Naredo, M., Wang, H., Atienza, G., Liu, B., Qiu, F., McNally, K., Leung, H., 2007. Rapid method for detecting SNPs on agarose gels and its application in candidate gene mapping. Mol. Breed. 19, 87–101. doi:10.1007/s11032-006-9046-x
Raymond, M.J., 2004. Isolation and characterization of latex-specific promoters from Papaver somniferum L. [WWW Document]. URL http://scholar.lib.vt.edu/theses/available/etd-08312004-174505/ (accessed 7.12.15).
Ribaut, J.-M., de Vicente, M., Delannay, X., 2010. Molecular breeding in developing countries: challenges and perspectives. Curr. Opin. Plant Biol. 13, 213–218. doi:10.1016/j.pbi.2009.12.011
Richardson, K.L., Vales, M.I., Kling, J.G., Mundt, C.C., Hayes, P.M., 2006. Pyramiding and dissecting disease resistance QTL to barley stripe rust. Theor. Appl. Genet. 113, 485–495. doi:10.1007/s00122-006-0314-2
Ririe, K.M., Rasmussen, R.P., Wittwer, C.T., 1997. Product Differentiation by Analysis of DNA Melting Curves during the Polymerase Chain Reaction. Anal. Biochem. 245, 154–160. doi:10.1006/abio.1996.9916
Robiquet, P., 1832. Nouvelles observations sur les principaux produits de l'opium. Ann. Chim. Phys. 51, 225–267.
Robiquet, P., 1817. Observations sur le memoire de M. Sertuerner relatif à l'analyse de l'opium. Ann. Chim. Phys. 5, 275–278.
Rogers, C., Wen, J., Chen, R., Oldroyd, G., 2009. Deletion-Based Reverse Genetics in Medicago truncatula. Plant Physiol. 151, 1077–1086. doi:10.1104/pp.109.142919
Roulin, A., Auer, P.L., Libault, M., Schlueter, J., Farmer, A., May, G., Stacey, G., Doerge, R.W., Jackson, S.A., 2013. The fate of duplicated genes in a polyploid plant genome. Plant J. 73, 143–153. doi:10.1111/tpj.12026
Ruff, B.M., Bräse, S., O'Connor, S.E., 2012. Biocatalytic production of tetrahydroisoquinolines. Tetrahedron Lett. 53, 1071–1074. doi:10.1016/j.tetlet.2011.12.089
Runguphan, W., Glenn, W.S., O'Connor, S.E., 2012. Redesign of a Dioxygenase in Morphine Biosynthesis. Chem. Biol. 19, 674–678. doi:10.1016/j.chembiol.2012.04.017
Safi, M., 2014. Opium poppies to be legalised in Victoria as demand for painkillers soars [WWW Document]. the Guardian. URL http://www.theguardian.com/world/2014/mar/18/opium-poppies-to-be-legalised-in-victoria-as-demand-for-painkillers-soars (accessed 4.16.15).
Samanani, N., Alcantara, J., Bourgault, R., Zulak, K.G., Facchini, P.J., 2006. The role of phloem sieve elements and laticifers in the biosynthesis and accumulation of alkaloids in opium poppy†. Plant J. 47, 547–563.
Samanani, N., Liscombe, D.K., Facchini, P.J., 2004. Molecular cloning and characterization of norcoclaurine synthase, an enzyme catalyzing the first committed step in benzylisoquinoline alkaloid biosynthesis. Plant J. 40, 302–313. doi:10.1111/j.1365-313X.2004.02210.x
Samanani, N., Park, S.U., Facchini, P.J., 2005. Cell Type–Specific Localization of Transcripts Encoding Nine Consecutive Enzymes Involved in Protoberberine Alkaloid Biosynthesis. Plant Cell Online 17, 915.
Sariyar, G., 2002. Biodiversity in the alkaloids of Turkish Papaver species. Pure Appl. Chem. 74, 557–574.
Sato, F., Takeshita, N., Fitchen, J.H., Fujiwara, H., Yamada, Y., 1993. S-adenosyl-l-methionine: Scoulerine-9-O-methyltransferase from cultured Coptis japonica cells.

Phytochemistry, The International Journal of Plant Biochemistry 32, 659–664. doi:10.1016/S0031-9422(00)95151-3
Sato, Y., Shirasawa, K., Takahashi, Y., Nishimura, M., Nishio, T., 2006. Mutant Selection from Progeny of Gamma-ray-irradiated Rice by DNA Heteroduplex Cleavage using Brassica Petiole Extract. Jpn. J. Breed. 56, 179–183. doi:10.1270/jsbbs.56.179
Schaefer, B., 2015. Natural Products in the Chemical Industry. Springer.
Schmeller, T., Latz-Brüning, B., Wink, M., 1997. Biochemical activities of berberine, palmatine and sanguinarine mediating chemical defence against microorganisms and herbivores. Phytochemistry 44, 257–266. doi:10.1016/S0031-9422(96)00545-6
Schumacher, H.-M., Gundlach, H., Fiedler, F., Zenk, M.H., 1987. Elicitation of benzophenanthridine alkaloid synthesis in Eschscholtzia cell cultures. Plant Cell Rep. 6, 410–413. doi:10.1007/BF00272770
Scott, J.B., Hay, F.S., Wilson, C.R., 2004. Phylogenetic analysis of the downy mildew pathogen of oilseed poppy in Tasmania, and its detection by PCR. Mycol. Res. 108, 198–205. doi:10.1017/S095375620300916X
Sellier, M.-J., Reeb, P., Marion-Poll, F., 2011. Consumption of Bitter Alkaloids in Drosophila melanogaster in Multiple-Choice Test Conditions. Chem. Senses 36, 323–334. doi:10.1093/chemse/bjq133
Sertürner, F., 1806. . J. Pharm. Fuer Aerzte Apoth. 14, 47–93.
Shan, Q., Wang, Y., Li, J., Zhang, Y., Chen, K., Liang, Z., Zhang, K., Liu, J., Xi, J.J., Qiu, J.-L., Gao, C., 2013. Targeted genome modification of crop plants using a CRISPR-Cas system. Nat. Biotechnol. 31, 686–688. doi:10.1038/nbt.2650
Sharma, J.R., Lal, R.K., Gupta, A.P., Misra, H.O., Pant, V., Singh, N.K., Pandey, V., 1999. Development of non-narcotic (opiumless and alkaloid-free) opium poppy, Papaver somniferum. Plant Breed. 118, 449–452. doi:10.1046/j.1439-0523.1999.00419.x
Shields, V.D.C., Smith, K.P., Arnold, N.S., Gordon, I.M., Shaw, T.E., Waranch, D., 2008. The effect of varying alkaloid concentrations on the feeding behavior of gypsy moth larvae, Lymantria dispar (L.) (Lepidoptera: Lymantriidae). Arthropod-Plant Interact. 2, 101–107. doi:10.1007/s11829-008-9035-6
Shitan, N., Bazin, I., Dan, K., Obata, K., Kigawa, K., Ueda, K., Sato, F., Forestier, C., Yazaki, K., 2003. Involvement of CjMDR1, a plant multidrug-resistance-type ATP-binding cassette protein, in alkaloid transport in Coptis japonica. Proc. Natl. Acad. Sci. 100, 751–756. doi:10.1073/pnas.0134257100
Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539. doi:10.1038/msb.2011.75
Simas, N.K., Ferrari, S.F., Pereira, S.N., Leitão, G.G., 2001. Chemical ecological characteristics of herbivory of Siparuna guianensis seeds by buffy-headed marmosets (Callithrix flaviceps) in the Atlantic Forest of southeastern Brazil. J. Chem. Ecol. 27, 93–107.
Singh, S., Jain, L., Pandey, M.B., Singh, U.P., Pandey, V.B., 2009. Antifungal activity of the alkaloids from Eschscholtzia californica. Folia Microbiol. (Praha) 54, 204–206. doi:10.1007/s12223-009-0032-7
Slade, A.J., McGuire, C., Loeffler, D., Mullenberg, J., Skinner, W., Fazio, G., Holm, A., Brandt, K.M., Steine, M.N., Goodstal, J.F., Knauf, V.C., 2012. Development of high amylose wheat through TILLING. BMC Plant Biol. 12, 69. doi:10.1186/1471-2229-12-69
Solovyev, V., Kosarev, P., Seledsov, I., Vorobyev, D., 2006. Automatic annotation of eukaryotic genes, pseudogenes and promoters. Genome Biol. 7, S10. doi:10.1186/gb-2006-7-s1-s10
Stadler, R., Kutchan, T.M., Loeffler, S., Nagakura, N., Cassels, B., Zenk, M.H., 1987. Revision of the early steps of reticuline biosynthesis. Tetrahedron Lett. 28, 1251–1254. doi:10.1016/S0040-4039(00)95338-3
Stadler, R., Kutchan, T.M., Zenk, M.H., 1989. (S)-norcoclaurine is the central intermediate in benzylisoquinoline alkaloid biosynthesis. Phytochemistry 28, 1083–1086. doi:10.1016/0031-9422(89)80187-6

Stephenson, P., Baker, D., Girin, T., Perez, A., Amoah, S., King, G.J., Østergaard, L., 2010. A rich TILLING resource for studying gene function in Brassica rapa. BMC Plant Biol. 10, 62. doi:10.1186/1471-2229-10-62
Straeten, D.V. der, Wiemeersch, L.V., Goodman, H.M., Montagu, M.V., 1990. Cloning and sequence of two different cDNAs encoding 1-aminocyclopropane-1-carboxylate synthase in tomato. Proc. Natl. Acad. Sci. 87, 4859–4863. doi:10.1073/pnas.87.12.4859
Sue, M., Nakamura, C., Nomura, T., 2011. Dispersed Benzoxazinone Gene Cluster: Molecular Characterization and Chromosomal Localization of Glucosyltransferase and Glucosidase Genes in Wheat and Rye. Plant Physiol. 157, 985–997. doi:10.1104/pp.111.182378
Suzuki, T., Eiguchi, M., Kumamaru, T., Satoh, H., Matsusaka, H., Moriguchi, K., Nagato, Y., Kurata, N., 2007. MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. Mol. Genet. Genomics 279, 213–223. doi:10.1007/s00438-007-0293-2
Takemura, T., Ikezawa, N., Iwasa, K., Sato, F., 2013. Molecular cloning and characterization of a cytochrome P450 in sanguinarine biosynthesis from Eschscholzia californica cells. Phytochemistry, Meinhart H. Zenk Memorial Issue 91, 100–108. doi:10.1016/j.phytochem.2012.02.013
Takeshita, N., Fujiwara, H., Mimura, H., Fitchen, J.H., Yamada, Y., Sato, F., 1995. Molecular Cloning and Characterization of S-Adenosyl-L-Methionine:Scoulerine-9-O-Methyltransferase from Cultured Cells of Coptis japonica. Plant Cell Physiol. 36, 29–36.
Takeuchi, T., Watanabe, Y., Takano-Shimizu, T., Kondo, S., 2006. Roles of jumonji and jumonji family genes in chromatin regulation and development. Dev. Dyn. 235, 2449–2459. doi:10.1002/dvdy.20851
Tasmanian Government, 2012. Legislative Council Select Committee Inquiry: Tasmanian Poppy Industry.
Taylor, N.E., Greene, E.A., 2003. PARSESNP: a tool for the analysis of nucleotide polymorphisms. Nucleic Acids Res. 31, 3808–3811. doi:10.1093/nar/gkg574
Thodey, K., Galanie, S., Smolke, C.D., 2014. A microbial biomanufacturing platform for natural and semisynthetic opioids. Nat. Chem. Biol. 10, 837–844. doi:10.1038/nchembio.1613
Tibi, S., 2003. Opium in ninth century Baghdad. Pharm. J. 271, 855–858.
Till, B.J., Burtner, C., Comai, L., Henikoff, S., 2004. Mismatch cleavage by single-strand specific nucleases. Nucleic Acids Res. 32, 2632–2641. doi:10.1093/nar/gkh599
Till, B.J., Cooper, J., Tai, T.H., Colowit, P., Greene, E.A., Henikoff, S., Comai, L., 2007. Discovery of chemically induced mutations in rice by TILLING. BMC Plant Biol. 7, 19. doi:10.1186/1471-2229-7-19
Till, B.J., Reynolds, S.H., Greene, E.A., Codomo, C.A., Enns, L.C., Johnson, J.E., Burtner, C., Odden, A.R., Young, K., Taylor, N.E., Henikoff, J.G., Comai, L., Henikoff, S., 2003. Large-Scale Discovery of Induced Point Mutations With High-Throughput TILLING. Genome Res. 13, 524–530. doi:10.1101/gr.977903
Till, B.J., Zerr, T., Comai, L., Henikoff, S., 2006. A protocol for TILLING and Ecotilling in plants and animals. Nat. Protoc. 1, 2465–2477. doi:10.1038/nprot.2006.329
Tohge, T., Alseekh, S., Fernie, A.R., 2014. On the regulation and function of secondary metabolism during fruit development and ripening. J. Exp. Bot. ert443. doi:10.1093/jxb/ert443
Townsend, T., Segura, V., Chigeza, G., Penfield, T., Rae, A., Harvey, D., Bowles, D., Graham, I.A., 2013. The Use of Combining Ability Analysis to Identify Elite Parents for Artemisia annua F1 Hybrid Production. PLoS ONE 8, e61989. doi:10.1371/journal.pone.0061989
Tsai, H., Howell, T., Nitcher, R., Missirian, V., Watson, B., Ngo, K.J., Lieberman, M., Fass, J., Uauy, C., Tran, R.K., Khan, A.A., Filkov, V., Tai, T.H., Dubcovsky, J., Comai, L., 2011. Discovery of Rare Mutations in Populations: TILLING by Sequencing. Plant Physiol. 156, 1257–1268. doi:10.1104/pp.110.169748

Uauy, C., Paraiso, F., Colasuonno, P., Tran, R.K., Tsai, H., Berardi, S., Comai, L., Dubcovsky, J., 2009. A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. BMC Plant Biol. 9, 115. doi:10.1186/1471-2229-9-115
Ugozzoli, L., Wallace, R.B., 1991. Allele-specific polymerase chain reaction. Methods 2, 42–48. doi:10.1016/S1046-2023(05)80124-0
Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., Leunissen, J.A.M., 2007. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Res. 35, W71–W74. doi:10.1093/nar/gkm306
Unterlinner, B., Lenz, R., Kutchan, T.M., 1999. Molecular cloning and functional expression of codeinone reductase: the penultimate enzyme in morphine biosynthesis in the opium poppy Papaver somniferum. Plant J. Cell Mol. Biol. 18, 465–475.
Van Hout, M., Bergin, M., Foley, M., Rich, E., Rapca, A., Harris, R., Norman, I., 2014. A Scoping Review of Codeine Use, Misuse and Dependence, final report.
van Ree, J.M., Gerrits, M.A.F.M., Vanderschuren, L.J.M.J., 1999. Opioids, Reward and Addiction: An Encounter of Biology, Psychology, and Medicine. Pharmacol. Rev. 51, 341–396.
Vazquez-Flota, F., Carolis, E.D., Alarco, A.-M., Luca, V.D., 1997. Molecular cloning and characterization of desacetoxyvindoline-4-hydroxylase, a 2-oxoglutarate dependent-dioxygenase involved in the biosynthesis of vindoline in Catharanthus roseus (L.) G. Don. Plant Mol. Biol. 34, 935–948. doi:10.1023/A:1005894001516
Vialart, G., Hehn, A., Olry, A., Ito, K., Krieger, C., Larbat, R., Paris, C., Shimizu, B., Sugimoto, Y., Mizutani, M., Bourgaud, F., 2012. A 2-oxoglutarate-dependent dioxygenase from Ruta graveolens L. exhibits p-coumaroyl CoA 2′-hydroxylase activity (C2′H): a missing step in the synthesis of umbelliferone in plants. Plant J. 70, 460–470. doi:10.1111/j.1365-313X.2011.04879.x
Vree, T.B., van Dongen, R.T., Koopman-Kimenai, P.M., 2000. Codeine analgesia is due to codeine-6-glucuronide, not morphine. Int. J. Clin. Pract. 54, 395–398.
Wakhlu, A.K., Bajwa, P.S., 1987. Cytological Analysis in Embryogenic Callus Cultures and Regenerated Plants of *Papaver somniferum* L. (Opium Poppy). Cytologia (Tokyo) 52, 631–638. doi:10.1508/cytologia.52.631
Walker, E.L., Robbins, T.P., Bureau, T.E., Kermicle, J., Dellaporta, S.L., 1995. Transposon-mediated chromosomal rearrangements and gene duplications in the formation of the maize R-r complex. EMBO J. 14, 2350–2363.
Wangkumhang, P., Chaichoompu, K., Ngamphiw, C., Ruangrit, U., Chanprasert, J., Assawamakin, A., Tongsima, S., 2007. WASP: a Web-based Allele-Specific PCR assay designing tool for detecting SNPs and mutations. BMC Genomics 8, 275. doi:10.1186/1471-2164-8-275
Wang, X., Jiang, G.-L., Green, M., Scott, R.A., Hyten, D.L., Cregan, P.B., 2012. Quantitative trait locus analysis of saturated fatty acids in a population of recombinant inbred lines of soybean. Mol. Breed. 30, 1163–1179. doi:10.1007/s11032-012-9704-0
Watts, S.M., Dodson, C.D., Reichman, O.J., 2011. The Roots of Defense: Plant Resistance and Tolerance to Belowground Herbivory. PLoS ONE 6, e18463. doi:10.1371/journal.pone.0018463
Weid, M., Ziegler, J., Kutchan, T.M., 2004. The roles of latex and the vascular bundle in morphine biosynthesis in the opium poppy, Papaver somniferum. Proc. Natl. Acad. Sci. U. S. A. 101, 13957.
Weiss, D., Baumert, A., Vogel, M., Roos, W., 2006. Sanguinarine reductase, a key enzyme of benzophenanthridine detoxification. Plant Cell Environ. 29, 291–302. doi:10.1111/j.1365-3040.2005.01421.x
Wells, R., Trick, M., Fraser, F., Soumpourou, E., Clissold, L., Morgan, C., Pauquet, J., Bancroft, I., 2013. Sequencing-based variant detection in the polyploid crop oilseed rape. BMC Plant Biol. 13, 111. doi:10.1186/1471-2229-13-111
Weng, J.-K., Philippe, R.N., Noel, J.P., 2012. The rise of chemodiversity in plants. Science 336, 1667–1670.
Wenningmann, I., Dilger, J.P., 2001. The Kinetics of Inhibition of Nicotinic Acetylcholine Receptors by (+)-Tubocurarine and Pancuronium. Mol. Pharmacol. 60, 790–796.
WHO, 2013. WHO Model List of Essential Medicines.

Wijekoon, C.P., Facchini, P.J., 2012. Systematic knockdown of morphine pathway enzymes in opium poppy using virus-induced gene silencing. Plant J. 69, 1052–1063. doi:10.1111/j.1365-313X.2011.04855.x
Williams, G., 2013. A Submission to the Review of the moratorium on genetically modified organisms (GMOs).
Wilmouth, R.C., Turnbull, J.J., Welford, R.W.., Clifton, I.J., Prescott, A.G., Schofield, C.J., 2002. Structure and Mechanism of Anthocyanidin Synthase from Arabidopsis thaliana. Structure 10, 93–103. doi:10.1016/S0969-2126(01)00695-5
Winzer, T., Gazda, V., He, Z., Kaminski, F., Kern, M., Larson, T.R., Li, Y., Meade, F., Teodor, R., Vaistij, F.E., Walker, C., Bowser, T.A., Graham, I.A., 2012. A Papaver somniferum 10-Gene Cluster for Synthesis of the Anticancer Alkaloid Noscapine. Science 336, 1704–1708. doi:10.1126/science.1220757
Winzer, T., Kern, M., King, A.J., Larson, T.R., Teodor, R., Donninger, S., Li, Y., Dowle, A.A., Cartwright, J., Bates, R., Ashford, D., Thomas, J., Walker, C., Bowser, T.A., Graham, I.A., 2015. Morphinan biosynthesis in opium poppy requires a P450-oxidoreductase fusion protein. Science aab1852. doi:10.1126/science.aab1852
Xie, K., Yang, Y., 2013. RNA-Guided Genome Editing in Plants Using a CRISPR–Cas System. Mol. Plant 6, 1975–1983. doi:10.1093/mp/sst119
Xu, Q., Lin, M., 1999. Benzylisoquinoline Alkaloids from Gnetum parvifolium. J. Nat. Prod. 62, 1025–1027. doi:10.1021/np980472f
Yamada, Y., Kokabu, Y., Chaki, K., Yoshimoto, T., Ohgaki, M., Yoshida, S., Kato, N., Koyama, T., Sato, F., 2011. Isoquinoline Alkaloid Biosynthesis is Regulated by a Unique bHLH-Type Transcription Factor in Coptis japonica. Plant Cell Physiol. 52, 1131–1141. doi:10.1093/pcp/pcr062
Yamaguchi, S., 2008. Gibberellin Metabolism and its Regulation. Annu. Rev. Plant Biol. 59, 225–251. doi:10.1146/annurev.arplant.59.032607.092804
Yazaki, K., 2006. ABC transporters involved in the transport of plant secondary metabolites. FEBS Lett. 580, 1183–1191. doi:10.1016/j.febslet.2005.12.009
Yazaki, K., 2005. Transporters of secondary metabolites. Curr. Opin. Plant Biol. 8, 301–307. doi:10.1016/j.pbi.2005.03.011
Yi, C., Yang, C.-G., He, C., 2009. A Non-Heme Iron-Mediated Chemical Demethylation in DNA and RNA. Acc. Chem. Res. 42, 519–529. doi:10.1021/ar800178j
Zerr, T., Henikoff, S., 2005. Automated band mapping in electrophoretic gel images using background information. Nucleic Acids Res. 33, 2806–2812. doi:10.1093/nar/gki580
Zhang, K., Halitschke, R., Yin, C., Liu, C.-J., Gan, S.-S., 2013. Salicylic acid 3-hydroxylase regulates Arabidopsis leaf longevity by mediating salicylic acid catabolism. Proc. Natl. Acad. Sci. 110, 14807–14812. doi:10.1073/pnas.1302702110
Zhang, W., Hu, J.-F., Lv, W.-W., Zhao, Q.-C., Shi, G.-B., 2012. Antibacterial, Antifungal and Cytotoxic Isoquinoline Alkaloids from Litsea cubeba. Molecules 17, 12950–12960. doi:10.3390/molecules171112950
Zhang, Z., Ren, J.-S., Clifton, I.J., Schofield, C.J., 2004. Crystal Structure and Mechanistic Implications of 1-Aminocyclopropane-1-Carboxylic Acid Oxidase—The Ethylene-Forming Enzyme. Chem. Biol. 11, 1383–1394. doi:10.1016/j.chembiol.2004.08.012
Zhou, L., Wang, L., Palais, R., Pryor, R., Wittwer, C.T., 2005. High-Resolution DNA Melting Analysis for Simultaneous Mutation Scanning and Genotyping in Solution. Clin. Chem. 51, 1770–1777. doi:10.1373/clinchem.2005.054924
Zhu, W., 2008. CYP2D6: a key enzyme in morphine synthesis in animals. Med. Sci. Monit. 14, SC15–SC18.
Ziegler, J., Diaz-Chávez, M.L., Kramell, R., Ammer, C., Kutchan, T.M., 2005. Comparative macroarray analysis of morphine containing Papaver somniferum and eight morphine free Papaver species identifies an O-methyltransferase involved in benzylisoquinoline biosynthesis. Planta 222, 458–471. doi:10.1007/s00425-005-1550-4
Ziegler, J., Facchini, P.J., 2008. Alkaloid biosynthesis: metabolism and trafficking. Annu Rev Plant Biol 59, 735–769.

Ziegler, J., Voigtländer, S., Schmidt, J., Kramell, R., Miersch, O., Ammer, C., Gesell, A., Kutchan, T.M., 2006. Comparative transcript and alkaloid profiling in Papaver species identifies a short chain dehydrogenase/reductase involved in morphine biosynthesis. Plant J. 48, 177–192. doi:10.1111/j.1365-313X.2006.02860.x

Zulak, K.G., Cornish, A., Daskalchuk, T.E., Deyholos, M.K., Goodenowe, D.B., Gordon, P.M.K., Klassen, D., Pelcher, L.E., Sensen, C.W., Facchini, P.J., 2006. Gene transcript and metabolite profiling of elicitor-induced opium poppy cell cultures reveals the coordinate regulation of primary and secondary metabolism. Planta 225, 1085–1106. doi:10.1007/s00425-006-0419-5