The University Of Sheffield.

Investigating PLAG1's Role in Pleomorphic Adenoma: A Study of Transcriptomics, Biotin-Based Proximity Labelling, and Functional Analysis

Fatima Elalawy

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

The University of Sheffield
Faculty of Medicine, Dentistry and Health
Department of Oral & Maxillofacial Pathology

School of Clinical Dentistry

September 2024

Acknowledgements

I would like to express my deepest gratitude to Dr Lynne Bingle. Her unwavering support, insightful guidance, and invaluable advice have been instrumental in the completion of this thesis. Her dedication and commitment to my research have been truly inspiring, and I am incredibly grateful for the opportunity to have worked under her mentorship. I would also like to sincerely thank Professor Ali Khurram and Professor Colin Bingle, for their continued support and valuable insights throughout this journey. Their expertise and encouragement have been crucial to my development as a researcher, and I deeply appreciate all the time and effort they have invested in helping me succeed.

A special thanks to Brenka McCabe, Matthew Worsely, Lisa Chang and Jason Heath for their exceptional assistance in the lab. Without their technical expertise, guidance, and patience, I would not have been able to complete the practical aspects of this research. Their contributions have been indispensable, and I am truly thankful for their support.

I am also profoundly grateful to the bioinformatician, Dr Phil Lewis, whose expertise in bioinformatic analysis has been invaluable. His thoughtful suggestions, clear explanations, and patience have made the complex aspects of data analysis much more accessible.

I would like to extend my heartfelt thanks to my parents, siblings, nieces, and nephews. Their constant encouragement and belief in me have been a source of strength and motivation throughout this journey. Special thanks to my niece Jood for her help and support- your assistance and concern have truly meant a lot to me.

A special and heartfelt thanks to my husband, Abdelbaset, for his unwavering support and sacrifices. I wouldn't have been able to achieve this without his support and patience, and encouragement.

To my dear son Adam ♥, I want to express my deepest apologies for not always being there while I was working on this thesis. Your understanding and patience mean the world to me. I am

incredibly grateful for your love and support. You have been a constant source of joy and inspiration, and I hope to make you proud as you grow.

Abstract

Background: Pleomorphic adenoma (PA) is the most common benign salivary gland tumour, with PLAG1 proposed to play a critical role in its development. The true involvement of PLAG1 in tumorigenesis is not fully understood, particularly regarding its potential dual role in both promoting and inhibiting tumour progression.

Aims and Objectives: The aims of this project were to investigate the role of PLAG1 in PA by exploring its impact on downstream gene expression, identifying novel PLAG1-interacting proteins, and assessing the effect of PLAG1 on cell behaviour.

Approach: HEK293 cells were transfected with PLAG1 or a mock vector, after which transcriptomic analysis was conducted using RNA sequencing to identify altered gene expression profiles. Gene set enrichment analysis provided insights into the affected pathways. PLAG1-proximal proteins were identified through biotin-based proximity labelling, followed by IPA analysis to elucidate the interaction networks and pathways. Functional assays, including proliferation, colony formation, migration, invasion, and apoptosis assays, were performed to examine the biological impact of PLAG1 on cell behaviour.

Results: Transcriptomic analysis revealed significant changes in pathways related to extracellular matrix remodelling, growth factor binding, and signalling pathways such as Hedgehog and Wnt. IPA analysis indicated that PLAG1 interacts with key upstream regulators such as MYB and JUN, influencing tumorigenic pathways. Functional assays demonstrated increased cell proliferation and colony formation following de novo PLAG1 expression but no effect on migration or invasion was determined. The apoptosis assay yielded inconclusive results, potentially due to the short post-transfection time used for this assay.

Conclusion: The of findings of this study suggest that PLAG1 plays a multifaceted role in PA, influencing various tumorigenic processes. While it supports cell proliferation and colony formation, its effect on migration and apoptosis remains unclear. This study provides a foundation

for further exploration of PLAG1's dual role and its implications in tumour progression and transformation.

Table of Contents

List of Figures

List of Tables

Abbreviations

ARC - Activity-regulated cytoskeleton-associated protein

AML - Acute myeloid leukemia

ACC - Adenoid cystic carcinoma

ANOVA - Analysis of variance

Bp - Base Pair

BCA - Bicinchoninic acid

BMP - Bone morphogenic protein

BP - Biological process

BSA - Bovine serum albumin

CAMs - Cell adhesion molecules

Ca Ex PA - Carcinoma ex pleomorphic adenoma

CC - Cellular components

CRABP2 - Cellular retinoic acid binding protein 2

CREBBP - CREB-binding protein

C° - Celsius

cDNA - Complementary DNA

CML - Chronic myeloid leukaemia

DAPI - 4', 6-diamidino-2-phenylindole

Δ - Delta

DEGs - Differentially expressed genes

DEPs - Differentially expressed proteins

DMSO - Dimethyl sulfoxide

DMEM - Dulbecco's Modified Eagle Medium

EV - Empty vector

ES - Enriched or purified sample

ES - Enrichment score

EGF - Epidermal growth factor

EMT - Epithelial-mesenchymal transitions

EDTA - Ethylenediaminetetraacetic acid

ECM - Extracellular matrix

EP300 - Histone acetyltransferase p300

FDR - False discovery rate

FTA - Fast technology for analysis of nucleic acids

FLNC - Filamin-C

FET - Fisher exact test

FACS - Fluorescence-activated cell sorting

FBS - Foetal bovine serum

GLI - Glioma-associated oncogene

GO - Gene ontology

GSEA - Gene set enrichment analysis

GAPDH - Glyceraldehyde 3-phosphate dehydrogenase

GCO - Global Cancer Observatory

GFP - Green fluorescent protein

HMGA2 - High Mobility Group AT-hook 2

HNSCC - Head and neck squamous cell carcinoma

HOX - Homeobox genes

HRP - Horseradish peroxidase

HEK293- Human embryonic kidney cell 293

HPV - Human papilloma virus

IF - Immunofluorescence

IPA - Ingenuity pathway analysis

IGF1 - Insulin-like growth factor 1

IGF2 - Insulin-like growth factor 2

IRES - Internal ribosomal entry site

JUN - Jun Proto-Oncogene, AP-1 Transcription Factor Subunit

KGM - Keratinocyte growth medium

KDa - Kilo Dalton

KEGG - Kyoto Encyclopedia of Genes and Genome

KRAS - Kirsten rat sarcoma virus

LQF - Label-Free Quantitation intensity

LIFR - Leukemia inhibitory factor receptor

LG - Low glucose

LEF1 - Lymphoid enhancer-binding factor 1

MPA - Metastasising pleomorphic adenoma

µg - Microgram

µl - Microlitre

µm - Micrometre

ml - Millilitre

MF - Molecular functions

MEC - Mucoepidermoid carcinoma

MTS - Methyltetrazolium Assay

MYB - Myeloblastosis oncogene

OSC - Oral squamous cell

P - Parotid

PBS - Phosphate-buffered saline

PBST - Phosphate-buffered saline with Tween

PDGF - Platelet-derived growth factor subunit B

PA - Pleomorphic adenoma

PLAG1 - Pleomorphic adenoma gene 1

PAGE - Polyacrylamide gel electrophoresis

PCR - Polymerase chain reaction

PCA - Principal component analysis

PI - Propidium Iodide

PTCH - Protein patched homolog 1

qPCR - Quantitative polymerase chain reaction

RIPA - Radioimmunoprecipitation assay buffer

RT-PCR - Reverse transcriptase polymerase chain reaction

RT - Reverse transcription

RA - Retinoic acid

RAR - Retinoic acid receptor

RNA - Ribonucleic acid

RUNX1 - Runt-related transcription factor 1

SDC - Salivary duct carcinoma

SGC - Salivary gland carcinoma

SGs - Salivary glands

SHH - Sonic hedgehog

STRING - Search tool for the retrieval of interacting genes

STR - Short tandem repeat

SDS - Sodium dodecyl-sulfate

SL - Sublingual

SM - Submandibular

SMO – Smoothened, Frizzled Class Receptor

TGFβ - Transforming growth factor beta

TGFβR3 - Transforming Growth Factor Beta Receptor 3

TMT - Tandem mass tag

UV - Ultraviolet

UF - Unbound fraction

V-FITC - Fluorescein isothiocyanate

WNT - Wingless-related integration site

WHO - World Health Organisation

1 Chapter 1 Literature Review

1.1 The Normal Salivary Glands

The human salivary glands are exocrine glands that produce and secrete saliva into the oral cavity through a specialised ductal system. Salivary glands (SG) are categorised into two main groups: major and minor. The major salivary glands, including the parotid, submandibular (SM), and sublingual glands (SL), occur in pairs. In contrast, the minor salivary glands are small clusters of glands dispersed throughout the oral mucosa.

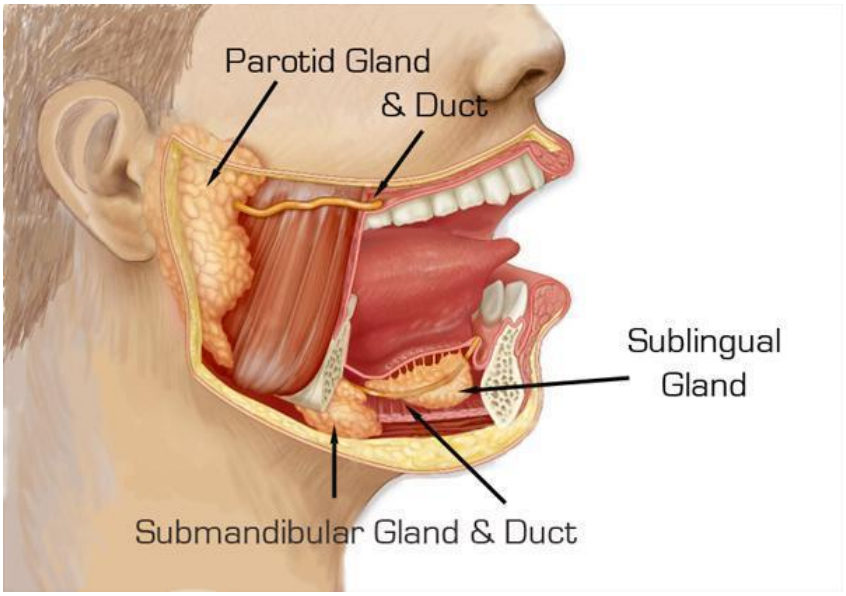


Figure 1.1 Anatomy of the Major Salivary Glands.
The illustration depicts the three major salivary glands and their respective ducts within the human head. The parotid gland, the largest salivary gland, is located near the cheek and secretes saliva through the parotid duct into the oral cavity. The submandibular gland, situated beneath the lower jaw, produces both serous and mucous secretions and drains saliva through the submandibular duct. The sublingual gland, found beneath the tongue, primarily produces mucous secretions and releases saliva through multiple small ducts.. Image taken from www.parotidsurgerymd.com

1.1.1 Development and Gross Anatomy

The salivary glands develop in sequential order during embryogenesis. The parotid glands are the first to form, beginning around 6-7 weeks of gestation, and they originate from ectodermal tissue. The submandibular glands develop shortly after, starting around 7 weeks of gestation, and are of

endodermal origin. The sublingual glands are the last of the major salivary glands to develop, starting their formation between 8-12 weeks of gestation, and they also originate from the endoderm. The minor salivary glands, which are scattered throughout the oral cavity in areas such as the lips, cheeks, tongue, and palate, begin to develop later during fetal life, around the 12th to 13th weeks of gestation, and their development continues postnatally. These glands are primarily of endodermal origin, although some ectodermal contributions may occur depending on their location (Som and Miletich, 2015; Priya et al., 2020).

The parotid gland is by far the largest and it may weigh up to 30 g. It is triangular and its apex lies just below the angle of the mandible with the upper base of the triangle being located near the zygomatic arch (Figure 1.1). The submandibular SM, the second largest, weighs 7 to 15 g and occupies a large part of the triangle that is bordered by the inferior border of the mandible (superiorly), the anterior belly of the digastric muscle (anteriorly) and the posterior belly of the digastric muscle (posteriorly). The almond-shaped SL is the smallest, weighing around 2 to 4 g, and lies just below the mucosa of the floor of the mouth between the tongue and the sublingual fossa of the mandible (Holsinger and Bui, 2007).

1.1.2 Acinar cell types in major and minor salivary glands

Salivary fluid is composed primarily of water, accounting for approximately 99% of its total volume. The remaining 1% consists of a complex mixture of electrolytes, proteins, glucose, nitrogenous substances like urea and ammonia, and other components (Humphrey and Williamson, 2001). This unique fluid is produced and released by the acinar lobules, the fundamental units of the salivary glands. Acinar cells within these lobules are classified into serous or mucous types, each with distinct cellular and intracellular characteristics, based on their secretory function (Table 1.1).

Table 1.1 Acinar cell type in the major and minor salivary glands.

Major Salivary Glands	**Type of Acinar Cells**
Parotid	Pure serous
Submandibular	Mixed serous and mucous
Sublingual	Mixed mucous and serous
Minor Salivary Glands	**Type of Acinar Cells**
Palate	Pure mucous
Tongue	Mixed mucous and serous
Lip	Pure mucous
Buccal mucosa	Pure mucous

1.1.3 Histology

- **Acini**

The secretory portion of the serous acini consists of 8 to 12 pyramidal cells surrounded by a basement membrane. The serous cells are arranged in a small spherical structure surrounding a tiny central lumen, through which the salivary secretion drains into the ducts (Figure 1.2). The nuclei of these cells are round and basally located in the cytoplasm. The most distinctive feature of the serous acini is the presence of many cytoplasmic secretory granules which are found mainly in the apical part of the cell (Berkovitz et al., 1992). The secretion of the serous cells is marked by elevated levels of the digestive enzyme 'amylase', ions and water. Nonspecific antimicrobial lysozyme and lactoferrin have also been found in the cytoplasm of these cells.

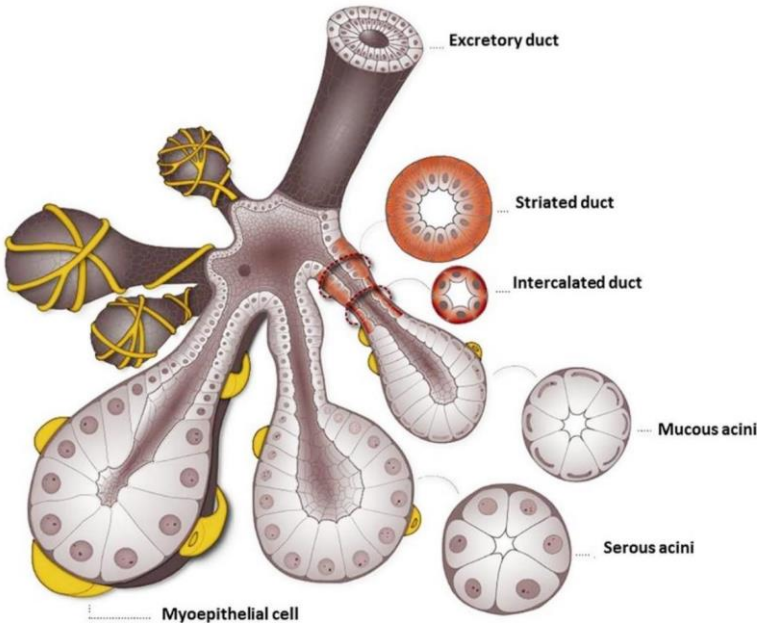


Figure 1.2 Representation of a fully developed mixed salivary gland.
The diagram represents the structural organization of a mixed salivary gland, which contains both serous and mucous acini responsible for producing different types of secretions. Serous acini (depicted with dark-staining cells) secrete a watery, enzyme-rich fluid, while mucous acini (depicted with pale-staining cells) produce a thicker, mucus-rich secretion. The myoepithelial cells (shown in yellow) are contractile cells that surround the acini and help propel secretions into the ductal system. The ductal system consists of:

- Intercalated ducts (small, simple cuboidal epithelium) that collect secretions from the acini.
- Striated ducts (lined with columnar cells with basal striations) that modify saliva by actively reabsorbing sodium and secreting potassium and bicarbonate.
- Excretory ducts (larger ducts lined with stratified epithelium) that transport the final saliva mixture into the oral cavity. Source: www.anatomypubs.onlinelibrary.wiley.com

The cross-section of the mucous acini shows large mucous cells which are arranged in a tubular pattern surrounding a central lumen that is larger than the lumen of the serous acini. The mucous cells have flat nuclei located in the basal part of the cell, and abundant cytoplasm with significant quantities of mucin granules in the apical portion of the cytoplasm. The basic function of the viscous mucous saliva is to form a lubricating and protective covering of the oral mucosa (Nanci, 2013). In mixed serous and mucous glands, serous cells are attached to the periphery of mucous acini in a crescent-shaped structure called the serous demilune. The secretions of the serous demilunes reach the mucous acini lumen via intercellular canaliculi found between the mucous cells.

- **Myoepithelial Cells**

Myoepithelial or “basket” cells are contractile cells that lie between the basement membrane of the acinar cells and the intercalated ducts and in some cases the striated and excretory ducts. They resemble smooth muscle cells but originate from the epithelium. The myoepithelial cells which are attached to the secretory portion are stellate in shape and have long cytoplasmic processes to embrace the secretory end piece. Cells present around the intercalated ducts are fusiform in shape and have fewer processes (Figure 1.3).

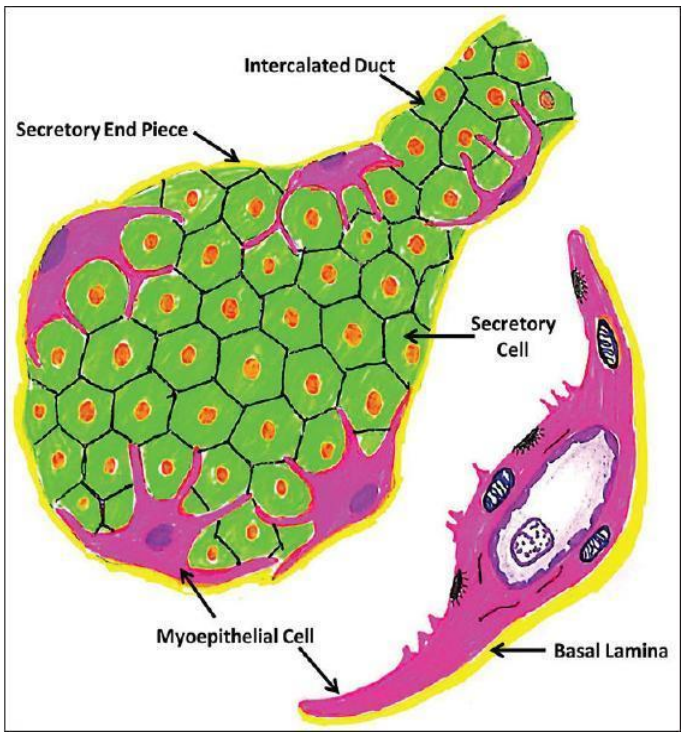


Figure 1.3 Myoepithelial cells.
Histological representation of a secretory unit, showing secretory end pieces (acini) and associated ducts. Secretory cells (green) are surrounded by myoepithelial cells (purple), which have a contractile function aiding in secretion. Myoepithelial cells attached to the secretory portion are stellate in shape, with long cytoplasmic processes that embrace the secretory end piece. Those around the intercalated ducts are fusiform with fewer processes. The basal lamina (yellow) provides structural support. Image taken from Antonio Nanci: TenCate's Oral Histology: Development, Structure, and Function, 8th edition. Elsevier Mosby.

Contraction of these cells is thought to give support to the secretory acini end portion at the time of active secretion of the salivary fluid. Moreover, the cells may promote excretion and movement of the fluid from the acinar end piece towards the ducts. In addition to their initial contractile role, it has been suggested that myoepithelial cells can perform other important functions. They pass signals to the functional excretory cells essential for preserving the secretory acini's structural organisation and cell polarity. It has also been claimed that myoepithelial cells produce numerous proteins such as proteinase inhibitors and antiangiogenesis factors that have tumour suppressor potential (Shah et al.,2016). Furthermore, myoepithelial cells participate in the synthesis of the basal lamina proteins such as laminin, fibronectin, and type III collagen, and thus provide a protective barrier against the invasion of epithelial neoplasms. It has also been suggested that myoepithelial cells are essential during salivary gland development and may act as a progenitor cell source (Ogawa, 2003; Ianez et al., 2010; Chitturi et al., 2015).

In some salivary gland neoplasms, due to their dual properties, altered myoepithelial cells may manifest epithelial, mesenchymal or both characteristics. Variable myoepithelial cell morphologies have been demonstrated in several salivary neoplasms such as pleomorphic adenoma (Dardick, et al., 1983), clear cell carcinoma (Corio, et al., 1982), myoepithelioma (Crissman, et al., 1977), adenocarcinoma (Dardick, et al., 1985), monomorphic adenoma (Dardick, et al., 1984), adenoid cystic carcinoma (Chaudhry, at al., 1986) and mucoepidermoid carcinoma (Dardick, et al., 1984). Altered myoepithelial cells exhibit four different morphological patterns (Table 1.2).

Table 1.2 The different morphological patterns of altered myoepithelial cells.

Cell type	Salivary gland tumour
Stellate or myxoid	Chondromyxoid areas of pleomorphic adenoma
Spindle-shaped or myoid	Pleomorphic adenoma and some types of myoepithelioma
Hyaline or plasmacytoid	Pleomorphic adenoma and may be seen in myoepithelioma
Clear or epithelial cells	Epithelial-myoepithelial carcinoma

- **Ducts**

The duct system of human salivary glands comprises a diverse network made up of small tubules which increase in diameter, starting within the secretory end pieces and ending in the oral cavity. Further to its initial role in transporting saliva, this complex system is actively involved in the synthesis and modification of primary saliva.

There are three types of ducts: intercalated, striated and excretory. The intercalated and striated ducts are “intralobular” lying entirely within the secretory lobules. They are referred to as “secretory ducts” because of their contribution to the production of saliva, whereas the excretory ducts are interlobular.

The saliva secreted from the secretory acini end pieces initially passes via the intercalated ducts that are directly connected with the secretory acini end pieces. The lumen of these ducts is lined by a simple cuboidal epithelium and an outer layer of contractile myoepithelial cells that are situated along the basal surface. Several ducts may join to form a bigger duct before passing primary saliva to the striated ducts. Intercalated ducts are of great importance for the regeneration and replenishment of salivary glands as their undifferentiated cells are thought to proliferate, differentiate and give rise to new acinar, ductal and myoepithelial cells (Ellis & Auclair, 2008).

Striated ducts comprise the largest part of the ductal complex. Their cells are columnar with central nuclei. The diameter of the striated ducts is larger than the diameter of secretory acini end pieces, and their lumen is wider than that of the intercalated ducts. In addition to transporting primary salivary fluid from the intercalating ducts to the excretory ducts, striated ducts help regulate the secretion and reabsorption of electrolytes.

The interlobular excretory ducts are situated within the connective tissue that lies between glandular lobules. These ducts serve an important role in modulating the inflow and outflow of some elements such as potassium and sodium and delivering the final salivary fluid to the mouth. Excretory ducts are composed of pseudostratified epithelium which may change to stratified

while getting closer to the oral opening. Furthermore, varied columnar epithelial cells demonstrate microvilli protruding from the cells towards the duct lumen and they are thought to play a part in the sensitivity, secretion or reabsorption of saliva (Berkovitz et al., 1992; Nanci, 2013).

1.1.4 Function

Saliva serves a crucial function in the preservation and maintenance of oral health and has been utilised as a non-invasive diagnostic and prognostic tool in many oral and systemic diseases. The function of saliva can be classified into five main categories: (1) protection and lubrication (2) buffering effect (3) preserving tooth integrity (4) taste and digestion (5) antimicrobial action (Moss, 1995; Mandel, 1987).

1.2 Head and neck cancer

Head and neck cancer is a broad term encompassing malignancies arising from various anatomical sites, including the oral cavity, pharynx, larynx, nasal cavity, paranasal sinuses, and salivary glands. The most prevalent subtype, accounting for approximately 90% of cases, is oral squamous cell carcinoma (OSCC) (Johnson et al., 2011). Despite advancements in diagnostic and therapeutic modalities, the overall survival rate for head and neck cancer remains unsatisfactory, highlighting the need for a comprehensive understanding of its pathogenesis and the development of novel treatment strategies (Muzaffar et al., 2021; Eberly et al., 2024).

A complex interplay of genetic, environmental, and lifestyle factors contributes to the development of head and neck cancer. Tobacco smoking and excessive alcohol consumption are well-established risk factors, although the etiological landscape is expanding to include human papillomavirus (HPV) infection, particularly in oropharyngeal cancers (Hammond et al., 2009). Clinical manifestations vary depending on the primary tumour site but commonly include persistent sore throat, hoarseness, difficulty swallowing, and oral ulcers. Early detection is crucial for optimal outcomes, as localised disease is often amenable to curative treatment.

Given its predominance, oral squamous cell carcinoma (OSCC) warrants specific attention. As previously mentioned, OSCC accounts for the majority of head and neck cancers, with an estimated 500,000 new cases diagnosed annually worldwide (Matta and Ralhan, 2009). In 2020, there were 377,713 cases of OSCC reported globally. The Global Cancer Observatory (GCO) predicts a 40% increase in OSCC incidence by 2040, along with a rise in mortality (Tan et al., 2023). Despite advances in treatment, the 5-year survival rate has plateaued, underscoring the need for improved therapeutic strategies.

1.3 Salivary Glands Tumours

Salivary gland tumours constitute a heterogeneous group of neoplasms arising from the salivary glands. Although the majority of salivary gland tumours are benign, a significant proportion exhibits malignant potential, emphasising the importance of accurate diagnosis and timely management.

The incidence and prevalence of salivary gland tumours vary across different populations and geographic regions. While precise global estimates are challenging due to variations in diagnostic practices and reporting systems, it is generally accepted that these tumours are relatively uncommon. In the United Kingdom, the annual incidence of salivary gland tumours is estimated to be approximately 2.5 to 3 cases per 100,000 population (Speight and Barrett, 2020). While specific prevalence data is less commonly reported, salivary gland tumours make up 6% to 8% of all head and neck cancers and less than 0.5% of all cancers in the human body; (Barnes, 2005; Lin et al., 2018).

They are more common in major than minor salivary glands with only around 9-23% originating from the minor salivary glands. About 64-80 % arise in the parotid gland and between 7 to 11% in the SM gland. The SL gland, the less commonly affected site, accounts for less than 1% of salivary neoplasms (Auclair, et al., 1991; Eneroth, 1971; Guzzo et al., 2010).

Approximately 20% to 25% of salivary gland tumours are malignant. This means that one in five to one in four salivary gland tumours are cancerous, while the majority are benign. The probability

of malignancy ranges from 15 to 32% in patients with parotid gland tumours, around 41-45% in those with SM gland lesions, 50% with minor salivary tumours and up to 90% in patients who present with SL gland tumours. Salivary gland tumours are particularly challenging to surgeons and pathologists due to their variable clinical and histological characteristics (Leegaard and Lindeman, 1970; Barnes, 2005; Guzzo et al., 2010).

The exact aetiology of most salivary gland tumours remains elusive. However, several risk factors have been identified, including exposure to ionising radiation, certain viral infections, genetic predisposition, diet, history of childhood benign salivary gland tumour, and exposure to chemicals. Chronic inflammation has also been implicated in the development of some salivary gland tumours (Batsakis, 1979.; Auclair et al., 1991; Guzzo et al., 2010).

1.3.1 Histopathology

The World Health Organization (WHO) classification provides a standardised framework for categorising salivary gland tumours based on their histological features (Wenig et al., 2023). The classification distinguishes between benign, malignant, and borderline tumours.

- **Benign tumours**: Pleomorphic adenoma (PA) is the most common benign salivary gland tumour. Other common benign tumours include Warthin's tumour, canalicular adenoma, and myoepithelioma.
- **Malignant tumours**: The spectrum of malignant salivary gland tumours is diverse, with mucoepidermoid carcinoma, adenoid cystic carcinoma, and carcinoma ex-pleomorphic adenoma (Ca Ex PA) being the most frequent subtypes.
- **Borderline tumours**: This category includes tumours with intermediate biological behaviour, such as basal cell adenoma, a benign tumour that can sometimes show features suggestive of malignancy, and polymorphous adenocarcinoma (previously known as Polymorphous Low-Grade Adenocarcinoma), which is considered low-grade malignant but often behaves in a less aggressive manner, making it borderline in terms of malignancy.

1.3.2 Molecular Aspects of Salivary Gland Tumours

Advances in molecular biology have provided insights into the pathogenesis of salivary gland tumours. The molecular underpinnings of salivary gland tumours have been the focus of intensive research in recent years. A comprehensive understanding of the genetic and epigenetic alterations driving tumorigenesis is crucial for developing targeted therapies and improving patient outcomes.

- **Genetic aberrations** constitute a cornerstone in the pathogenesis of salivary gland tumours. A prime example is the overexpression of Pleomorphic Adenoma Gene 1 (PLAG1), a transcription factor implicated in tumour development and growth, commonly observed in PA. This overexpression often arises from chromosomal translocations (Abdel-Hafiz et al., 2012). Chromosomal rearrangements resulting in fusion genes like ETV6-NTRK3 in secretory carcinoma and MYB/MYBL1-NFIB in adenoid cystic carcinoma highlight the intricate genetic landscape of these tumours (Chicoteka et al., 2019; Zhang et al., 2020). The broader spectrum of genetic alterations, encompassing amplifications and deletions, underscores the complex nature of salivary gland tumorigenesis.
- **Epigenetic modifications**, such as DNA methylation and histone modifications, complement the genetic alterations. Aberrant DNA methylation patterns silence tumour suppressor genes while activating oncogenes (Zhang et al., 2015). Histone modifications, including acetylation and methylation, influence chromatin accessibility and subsequent gene expression.
- **Signalling pathway dysregulation**, particularly in the PI3K/AKT/mTOR, EGFR, Wnt/β-catenin, and Notch pathways, is a hallmark of salivary gland tumorigenesis. These pathways orchestrate cellular processes like proliferation, survival, angiogenesis, and differentiation (Carlsson et al., 2009).
- **MicroRNAs (miRNAs)**, as post-transcriptional regulators, contribute to the intricate web of molecular alterations. Their dysregulation influences proliferation, apoptosis, invasion, and metastasis (Liu et al., 2018).

1.3.3 Fusion Genes in Salivary Gland Tumours

Fusion genes, formed by aberrant chromosomal rearrangements, play a critical role in the pathogenesis of several cancer types. In recent years, significant advancements in molecular techniques have unveiled the importance of these genetic alterations in salivary gland tumours.

1.3.3.1 History of fusion genes

Acquired abnormalities of the human chromosomes were first proposed by Boveri in 1914 to be the causative factor in the development of cancer. However, this only remained an appealing hypothesis until the discovery of the Philadelphia chromosome in chronic myeloid leukaemia (CML) in 1960 (Nowell and Hungerford, 1960). With the introduction of more advanced molecular cytogenetic technologies, many specific chromosome aberrations have been identified in multiple tumour types, including some solid tumours. The molecular characterisation of the Philadelphia chromosome translocation in CML revealed the BCR-ABL1 gene fusion, and the MYC gene in Burkitt's Lymphoma improved our perception of the importance of translocations and their end products, gene fusions, in the initiation of cancers, and subsequently increased interest in cytogenetics of cancer as an effective means of identifying genes potentially involved in the initial steps of carcinogenesis (Rowley, 2001). Advanced cytogenetic techniques have led to the detection of more than 337 distinct fusion genes involved in benign and malignant tumours (Futreal et al., 2004). A more recent analysis identified 420 known oncogenic fusions and 25 unclassified gene fusions across 26 different cancer types (Darabi et al., 2023).

1.3.3.2 Definition and mechanism

Fusion proteins are formed by the joining of two different genes, initially coding for separate proteins, resulting in chromosomal structural rearrangements such as translocation (Figure 1.4). They mediate their oncogenicity through a number of mechanisms that can result in (1) the deregulation of the involved genes as a result of promoter swapping (e.g., overexpression of structurally normal proteins), (2) the formation of a new chimeric, novel, proteins when gene breakpoints and fusions take place in the coding region in one or both of the relevant genes (e.g.,

hyperactive kinase) or (3) the premature truncation of protein products that are caused by the presence of a termination codon as a result of gene mutation (Mitelman et al., 2007; Latysheva and Babu, 2019).

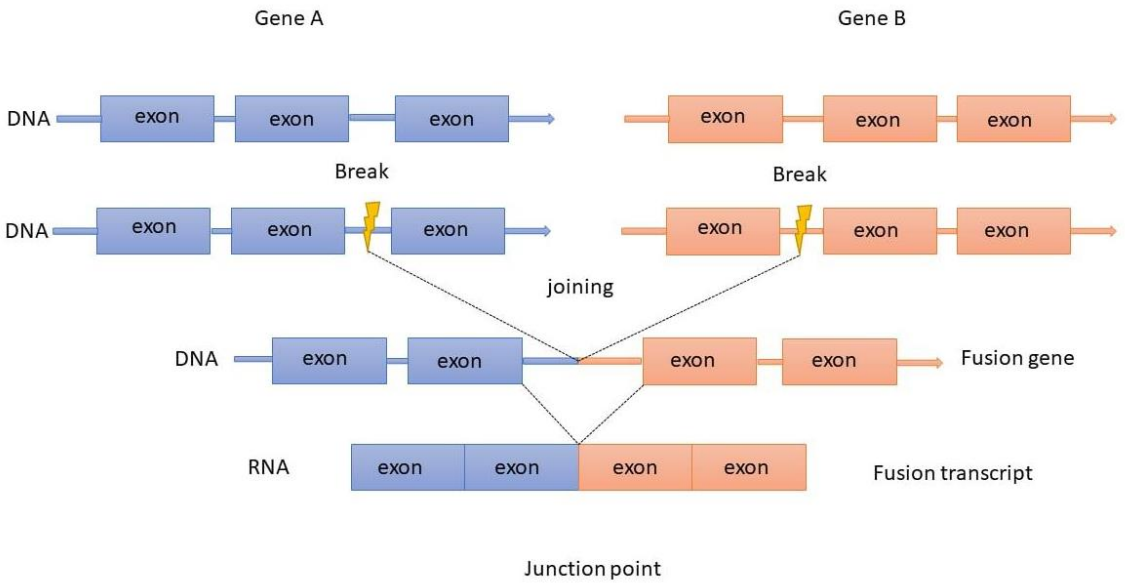


Figure 1.4 Fusion transcripts.
Fusion transcripts are mRNAs formed by the joining of two parts of different genes due to chromosomal rearrangements, such as translocations, inversions, or deletions. In the diagram, Gene A and Gene B, each consisting of multiple exons, undergo DNA breaks at specific points, represented by lightning bolt symbols. These breaks allow fragments from the two genes to join together, forming a fusion gene at the DNA level. This fusion gene is then transcribed into a fusion transcript, which contains exons from both genes. The junction point in the RNA marks the exact location where exons from Gene A and Gene B have been fused.

1.3.3.3 Significance of fusion genes

While fusion genes have been extensively studied in haematological malignancies and solid tumours such as lung and prostate cancer, their role in salivary gland tumours is relatively less explored. However, emerging evidence suggests their potential significance in these neoplasms.

Fusion oncogenes have been recognised for their importance in the development of head and neck tumours and have been demonstrated in malignancies arising from the maxilla, nasopharynx, salivary glands, lacrimal glands, auditory meatus, oesophagus, and thyroid gland

(Persson et al., 2009; Chen et al., 2014; Skálová et al., 2024). They set targets for therapies that can be beneficial to patients with unresectable malignancies while offering valuable contributions to the understanding of tumorigenesis, diagnosis, and prognosis. Additionally, as a result of their tumour-specific expression, fusion genes and their products have proved useful as tumour subtype biomarkers (Latysheva and Babu, 2016). Table 1.3 lists examples of fusion genes that have been identified in salivary gland tumours with their potential role in pathogenesis.

Table 1.3 Examples of fusion genes in salivary gland tumours and their role in pathogenesis (Stenman, 2013)

Fusion Gene	Tumour Type	Potential Role in Pathogenesis
PLAG1	Pleomorphic adenoma	Overexpression of PLAG1 leads to increased cell proliferation and tumour growth
HMGA2	Pleomorphic adenoma	Disrupts normal chromatin architecture, leading to uncontrolled cell growth
MYB–NFIB	Adenoid cystic carcinoma	Drives tumorigenesis by dysregulating MYB transcription factor activity
CRTC1–MAML2	Mucoepidermoid carcinoma	Activates CREB-mediated transcription, promoting cell proliferation and survival
ETV6–NTRK3	Mammary analogue secretory carcinoma	Leads to constitutive activation of the TRK signalling pathway, driving oncogenesis

Salivary gland tumours exhibit significant molecular heterogeneity, necessitating a comprehensive understanding of the underlying genetic and epigenetic mechanisms driving tumorigenesis. This study focuses on elucidating the molecular landscape of PA, specifically investigating the role of PLAG1 overexpression, potentially driven by gene fusions, in tumorigenesis.

1.3.4 Pleomorphic Adenoma

PA is the most common benign salivary gland neoplasm. It accounts for about 60% of all salivary gland tumours and affects both major and minor salivary glands. The parotid gland is the most

common site and accounts for 85% of the cases followed by SM accounting for 8%, while the SL gland is rarely affected. About 7% of PA arise in the minor salivary glands mainly in the palatal glands. PA may affect glands at other sites in the body such as the lacrimal glands and breast. The tumour has female predilection (2:1 ratio) and occurs in people of all ages including young adults and children, although it is more common in the 3rd to 6th decades (Zhan et al., 2016).

PA are typically slow-growing asymptomatic swellings. They are usually mobile and commonly present in the superficial lobe of the parotid gland with swelling just underneath and in front of the ear. According to Riad et al (2011) the mean diameter of PA is 32 mm, but it can grow to more than 4 cm, especially if left unattended.

Microscopically, PA is made up of epithelial (ductal) and myoepithelial cells. These cells have been found to have the potential to undergo mesenchymal metaplasia, thus giving rise to the development of a wide array of stromal tissue ranging from myxoid, chondroid, chondromyxoid or less frequently to osseous or adipose tissue. As a result of the broad spectrum of stromal tissue type and the mixture of epithelial and myoepithelial cells, this tumour is referred to as pleomorphic' or 'mixed tumour' (Enescu et al., 2014).

The origin of PA was first thought to be the multiple germ layers, but this was proven inaccurate. The different stromal tissues are developed due to the mesenchymal metaplasia of the altered epithelial and myoepithelial cells (Noguchi et al., 1996). Various factors, including genetic alterations, epigenetic modifications, and interactions with the tumour microenvironment likely influence this process. Moreover, cytogenetic studies have demonstrated alteration in the long arms of chromosomes 8 and 12 in areas encoding for PLAG1 (pleomorphic adenoma gene) and HMGA2 (High Mobility Group AT-Hook 2) genes, respectively. PLAG1 and HMGA2 gene rearrangements have been demonstrated in PA as well as its malignant counterpart Ca ex PA (Matsuyama et al., 2011).

The ductal or epithelial cells, which are cuboidal in shape, generally proliferate in three different patterns: ducts, tubules or solid sheets. The lumen of these ducts may contain some eosinophilic secretions. The epithelial cells could also undergo squamous, oncocytic or mucous metaplasia.

The oncocytic cells appear as deep eosinophilic cells filled with mitochondria. Spindle-shaped myoepithelial cells can be seen surrounding the ductal or tubular structures. They can also proliferate separately as sheets. Myoepithelial cells may become altered and oval-shaped with eosinophilic cytoplasm and eccentric nuclei and, as they resemble plasma cells, they are referred to as plasmacytoid cells. The stroma of this tumour varies from myxoid (loose stroma), to fibrous or, sometimes, hyalinized. Other mesenchymal elements can also be seen including chondroid, osseous and adipose tissue (Enescu et al., 2014) (Figures 1.5 & 1.6). Most importantly, PA is encapsulated and has surrounding fibrous tissue, although under the microscope it may appear with incomplete capsule, capsular penetration, finger-like projections (pseudopodia) or satellite nodules (Figures 1.7 & 1.8).

The primary treatment for PA is surgical removal. The preferred surgical approach depends on the location and size of the tumour. Local resection with minimal surrounding tissue is often possible, especially in the submandibular and minor salivary glands. For tumours in the parotid gland, the standard treatment is a superficial parotidectomy, which involves removing the affected portion of the gland while preserving the facial nerve. This approach is associated with low recurrence rates of 3-4%. Enucleation, a less extensive procedure that involves removing the tumour capsule, is no longer recommended due to its higher recurrence rate of 40% reported in earlier studies (Witt, 2002; Kanatas et al., 2018).

It has been hypothesised by Dulguerov et al (2017) that the recurrence of pleomorphic adenoma can be categorised into pathology-related factors such as the thickness of the capsule, lack of capsule, pseudopodia, satellite nodules and multicentricity of the tumour (Figure 1.7 & 1.8), or surgery-related factors which include; rupture of the tumour, spillage of tumour contents, inadequate resection of the tumour mass margins, because of the proximity of the branches of the facial nerve, and insufficient excision due to the sort of surgical procedure.

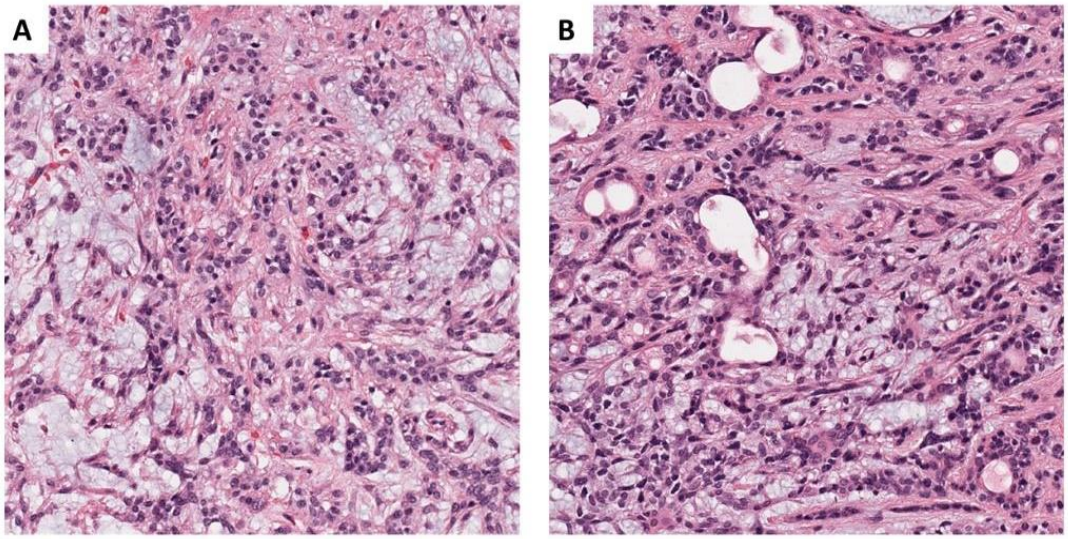


Figure 1.5 Pleomorphic Adenoma.
A) Myoepithelial cells in chondromyxoid stroma. B) Randomly scattered ductal structures in richly epithelial areas.

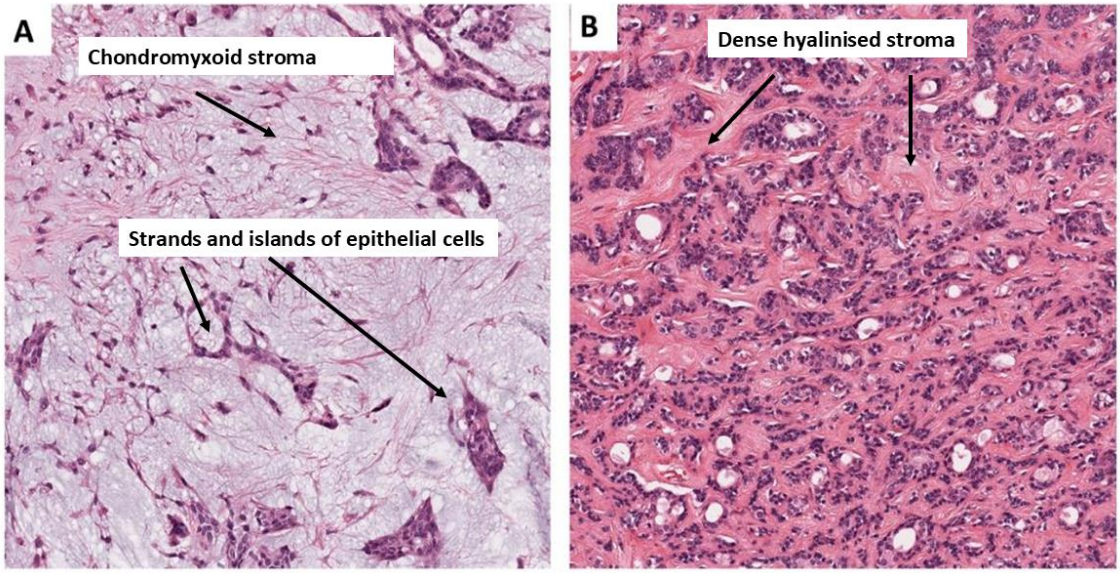


Figure 1.6 PA stroma.
A) Chondromyxoid area contains islands and strands of epithelial cells. B) Myoepithelial cells in dense hyalinised stroma.

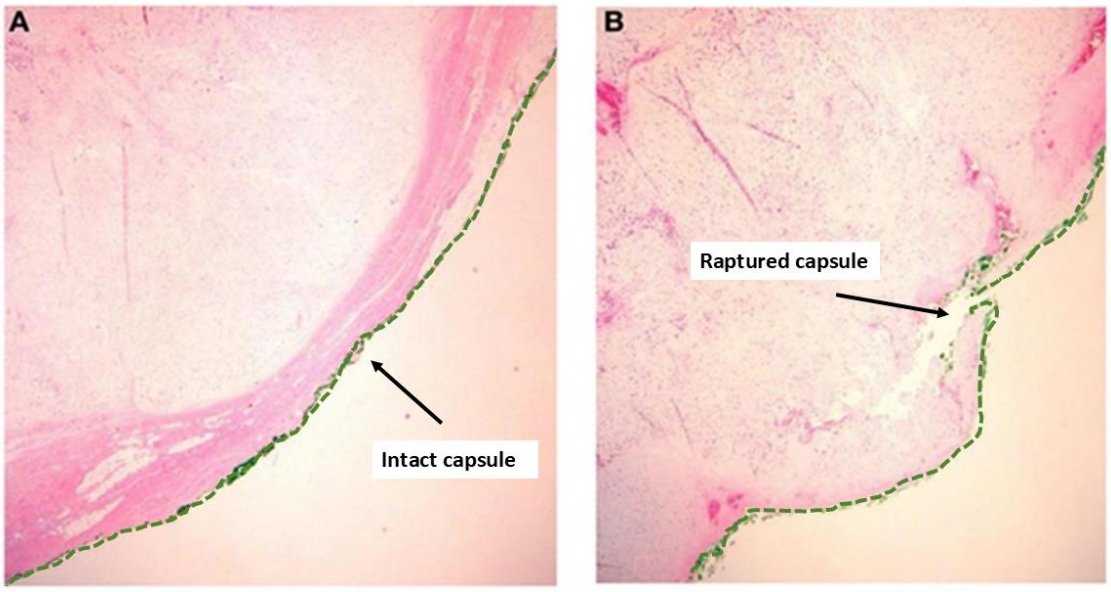


Figure 1.7 Integrity of PA capsule.
A) Intact capsule. B) Ruptured capsule.

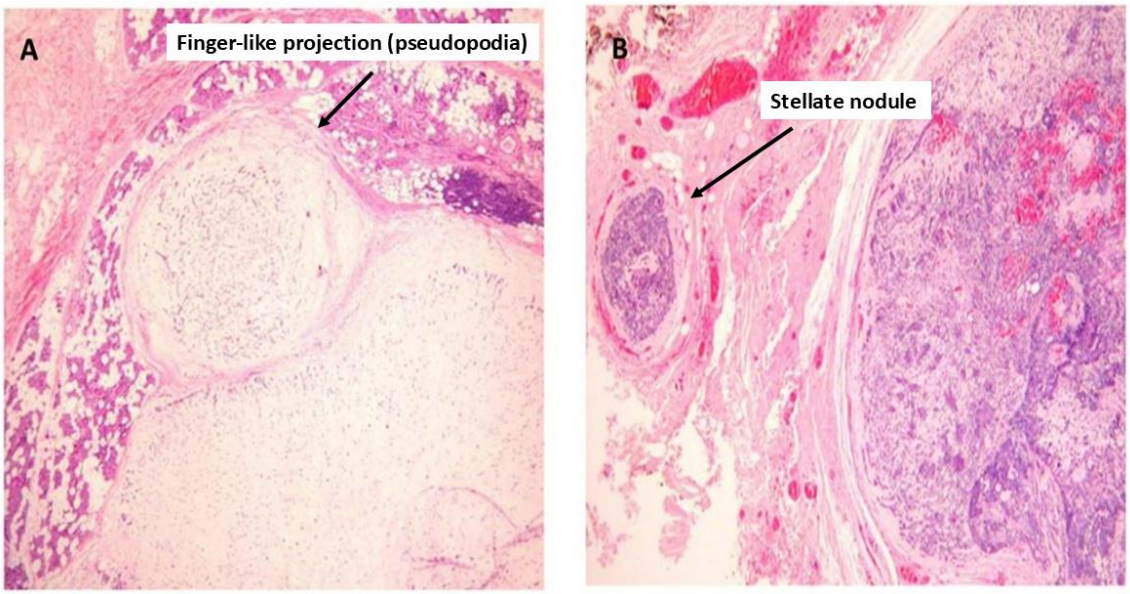


Figure 1.8 Potential causes of PA recurrence.
A) Finger-like projection (pseudopodia). B) Stellate nodule. Distinct tumour nodule in the vicinity of the main tumour mass but separated from it. Images taken from Dulguerov et al., 2017.

- **Metastasizing Pleomorphic Adenoma**

While the majority of PA are benign and well-behaved, a rare subset can exhibit aggressive behaviour, leading to the development of metastasising pleomorphic adenoma (MPA). MPA is a rare occurrence, with reported incidence rates varying widely. Metastasis typically occurs years after the initial diagnosis of PA, often following multiple local recurrences. The most common sites for metastasis include bone, lung, and cervical lymph nodes. However, other sites such as the liver, skin, and central nervous system have also been reported.

Histologically, MPA is indistinguishable from conventional PA, exhibiting a biphasic pattern of epithelial and myoepithelial cells embedded in a myxoid stroma. This paradox of benign histology with malignant behaviour has led to the term "mixed malignant tumour."

The molecular mechanisms underlying the metastatic potential of PA remain incompletely understood. While genetic alterations and epigenetic modifications are likely involved, specific driver mutations associated with MPA have not been consistently identified. Some studies suggest that alterations in cell adhesion molecules and extracellular matrix components may contribute to the invasive and metastatic phenotype (Scarini et al., 2023).

The treatment of MPA is challenging due to its unpredictable behaviour. Surgery remains the primary modality, with complete excision being the goal. However, recurrence and metastasis are common, even after multiple surgeries. Radiotherapy and chemotherapy may be considered in advanced cases, but their effectiveness is limited. Prognosis for patients with MPA is generally poor, with survival rates significantly lower compared to patients with conventional PA (Manucha and Ioffe, 2008; Knight and Ratnasingham, 2015; Fonseca et al., 2022).

1.3.5 Carcinoma ex pleomorphic adenoma (Ca ex PA)

Ca ex PA is a rare, aggressive malignancy arising from a primary (de novo) or recurrent PA. Despite its low incidence, Ca ex PA poses a significant clinical challenge due to its aggressive behaviour and potential for metastasis. Ca ex PA accounts for 5-15% of all malignant salivary gland tumours

and is seen predominantly in the 6th to 8th decades of life and at a slightly higher rate in females than males (Gnepp, 1993; Antony et al., 2012; Tondi-Resta et al., 2023).

Nouraei et al (2005) and Zbären et al (2008) noted that 21% to 25% of Ca ex PA patients had a previously treated benign PA. A delay in the treatment of long-standing PA can significantly increase the risk of malignant transformation into Ca ex PA, with approximately 12% of recurrent PAs potentially progressing to this aggressive malignancy. The diagnosis of Ca ex PA mainly relies on the recognition of a previously existing component of PA, or a history of prior surgical excision of a PA at the site of the lesion.

The parotid and the submandibular glands are the most commonly affected sites, although minor salivary glands of the hard and soft palate may be involved (Damm and Fantasia, 2001) in which case the lesion tends to be smaller in size. Apart from the salivary glands, Ca ex PA has been observed in other locations such as the breast, nasal cavity, lacrimal glands and trachea (Cho et al., 1995; Baredes et al., 2003; Hayes et al., 2005; Hu et al., 2016).

Ca ex PA, by definition, is made up of a combination of PA and carcinoma (Figure 1.9). Lewis et al (2001) found that the malignant (carcinomatous) components of the tumour comprised more than 50% of the tumour tissue in 84% of the 73 Ca ex PA cases involved in their study. The proportion of malignant tissue within a tumour can vary significantly as it can occupy the whole tumour mass leaving no trace of benign PA components. Alternatively, the tumour can be composed predominantly of benign PA components with only a few scattered malignantly transformed foci characterised by nuclear pleomorphism, atypical mitotic figures and areas of haemorrhage and necrosis. This often poses a challenge for the pathologists and may lead to a misdiagnosis. The malignant component of the tumour is most often adenocarcinoma not otherwise specified or undifferentiated carcinoma; however, it can be any other subtype of salivary gland carcinoma (SGC). This may include salivary duct carcinoma (SDC), mucoepidermoid carcinoma (MEC) or adenoid cystic carcinoma (ACC) (El-Naggar et al., 2017).

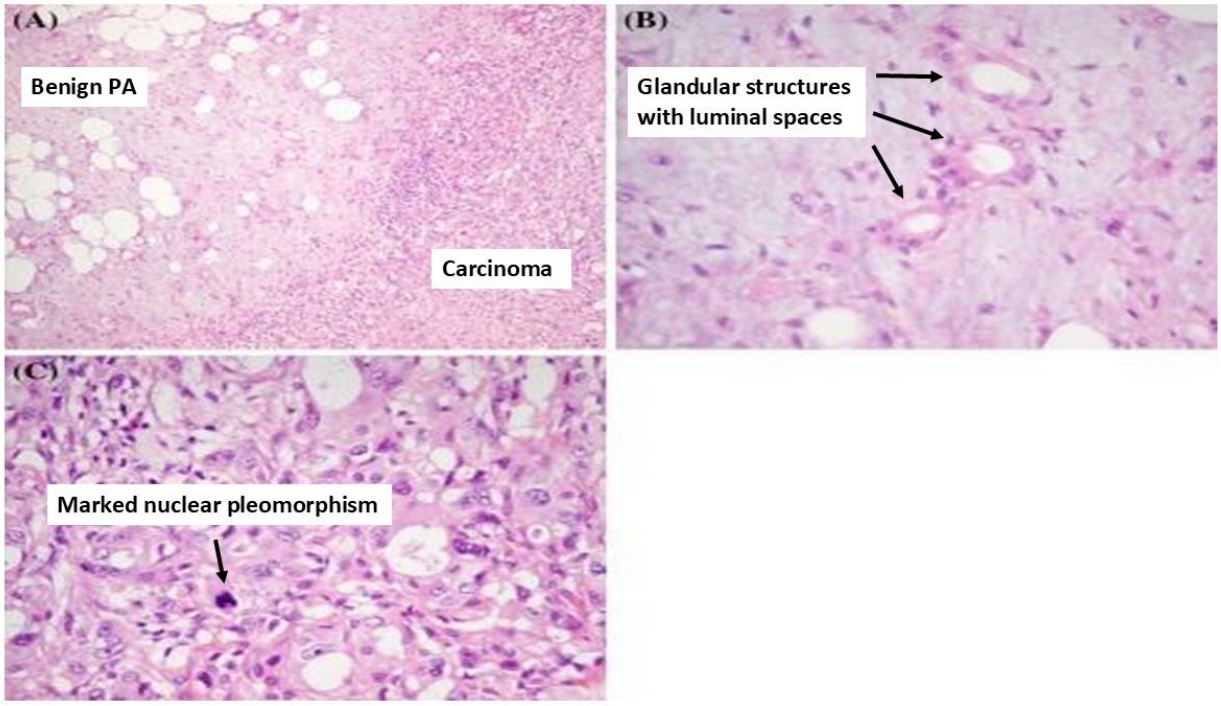


Figure 1.9 Carcinoma ex Pleomorphic Adenoma.
(A) Histological transition from benign PA to carcinoma: The left side shows a well-defined pleomorphic adenoma (PA) composed of myxoid stroma and glandular structures, while the right side displays malignant transformation into carcinoma. This highlights the coexistence of benign and malignant components. (H&E, original magnification ×10). (B) Higher magnification of the PA component: The PA area consists of glandular structures lined by myoepithelial cells, radiating into an abundant myxoid stromal background. Luminal spaces, indicative of secretory activity, are also present. (H&E, original magnification ×40). (C) Higher magnification of the carcinoma component: The malignant region exhibits a poorly differentiated adenocarcinoma with significant nuclear pleomorphism, atypical mitotic figures, and disrupted glandular architecture. Only sparse glandular formations and lumens remain, reflecting the loss of normal glandular differentiation. (H&E, original magnification ×40).. Source: Antony et al., 2012.

Olsen and Lewis (2001) reported that the degree of invasion beyond the capsule varied from 2 to 100 mm and based on this, Ca ex PA is referred to as non-invasive (intracapsular including in situ carcinoma), minimally invasive (< 1.5mm capsular penetration) or widely invasive (> 1.5mm extracapsular invasion) (Barnes et al., 2005). However, the 2017 WHO classification of salivary gland tumours introduced several significant changes regarding Ca ex PA. Most notably, it clearly stated that Ca ex PA should not be considered a standalone diagnosis. Instead, the histological type of the malignant component must be specified in the diagnostic report. Additionally, the

classification proposed a preliminary threshold of 4-6 mm for minimal invasion, recognising the need for further research to validate this cutoff (Speight and Barrett, 2020). Differentiation between the three categories is predictive of the prognosis and survival rate as widely invasive tumours are highly aggressive with a poor prognosis.

Wide resection of the tumour is the treatment of choice, but patients may also be offered a combination of radiotherapy and chemotherapy. There is limited data, however, surrounding the efficiency of chemotherapy in the treatment of Ca ex PA.

1.3.6 PA and Ca ex PA fusion genes

The typical genomic feature or hallmark of PA is translocations affecting the PA gene1 (PLAG1) and high mobility group A2 (HMGA2) oncogenes (Figure 1.10). Cytogenetic studies of PA tumours have demonstrated a high chromosomal abnormality rate (70%). Approximately 40%-50% have a chromosomal translocation involving band 8q12 targeting the developmentally regulated transcription factor PLAG1 while around 8% 15% have translocations in band 12q14-15 targeting the transcriptional regulating factor HMGA2. A small percentage of PA cases have other chromosomal abnormalities, however, around 30% of PA cases have a normal karyotype (Kandasamy et al., 2007; Stenman, G., 2013; Stenman et al., 2022).

As a transcription factor, PLAG1 is involved in regulating gene expression by binding to specific DNA sequences and influencing the transcription of target genes involved in cell growth and differentiation. It belongs to the PLAG family of zinc finger proteins, which are known for their involvement in developmental processes and tumorigenesis. PLAG1 is normally expressed in several tissues during development. Studies of gene expression have revealed that PLAG1 is expressed in high levels in the foetal liver, kidney and lung. PLAG1 is also expressed in developing salivary glands, where it plays a role in glandular development and differentiation. In adult tissue, however, expression levels are below the limits of detection or absent (Kas et al., 1997; Voz et al., 2000).

PLAG1 was first identified through positional cloning while scientists were investigating salivary gland PA (Kas et al., 1997). The role of PLAG1 in different cancers has been meticulously studied since then. PLAG1 has also been found to play a role in types of tumours other than pleomorphic adenoma of the salivary glands. This includes hepatoblastoma, lipoblastoma, and acute myeloid leukaemia (AML) (Åström et al., 2000; Zatkova et al., 2004; Landrette et al., 2005). Additionally, PLAG1 ectopic expression has been detected in uterine leiomyoma, leiomyosarcoma and tumours of smooth muscles (Åström et al., 1999).

One of the main causes of PLAG1-induced human neoplasms is thought to be an event of chromosomal translocation. This leads to the exchange of the PLAG1 promoter region with ubiquitously expressed partner genes in a process known as promoter swapping.

In PA and Ca ex PA, PLAG1 has been found to have at least eight translocation partner genes, as shown in Figure 1.10, including CTNNB1 (encoding β-catenin), LIFR (which encodes the leukaemia inhibitory factor receptor) and the transcription elongation factor SII gene (Kas et al., 1997; Voz et al., 1998; Åström et al., 1999; Dalin et al., 2017). The breakpoints of the chromosomes are localised between the promoter and the coding region of the involved translocation genes, maintaining the coding sequence of both genes. As a result of promoter swapping, PLAG1 expression comes under the control of the ubiquitously expressed partner gene, leading to ectopic overexpression of PLAG1. This in turn leads to the up or down-regulation of PLAG1 target genes and subsequent tumour formation (Voz et al., 2000, 2004).

PLAG1 exerts its oncogenicity via the deregulation of target genes that modulate cell growth, proliferation and apoptosis. It has been demonstrated that PLAG1 can control cell proliferation by regulating the expression of several growth factors such as insulin-like growth factor 2 (IGF2), cytokine-like factor 1 (CLF1), bone-derived growth factor (BDGF), vascular endothelial growth factor (VEGF) and placental growth factor (PGF). IGF-2 is upregulated in several types of neoplasia, stimulating cell proliferation via paracrine or autocrine mechanisms, which has led to the hypothesis that the oncogenicity of PLAG1 is very likely to be mediated through the activation of the IGF-2 mitogenic pathway. Therefore, insulin-like growth factor receptors (IGF1R), crucial for

mediating the biological effects of IGF-1 and IGF2, such as promoting cell growth and inhibiting apoptosis, could be potential therapeutic targets and a subject of research with therapeutic purposes (Dyck et al, 2007; Juma et al, 2016; Stenman et al., 2022).

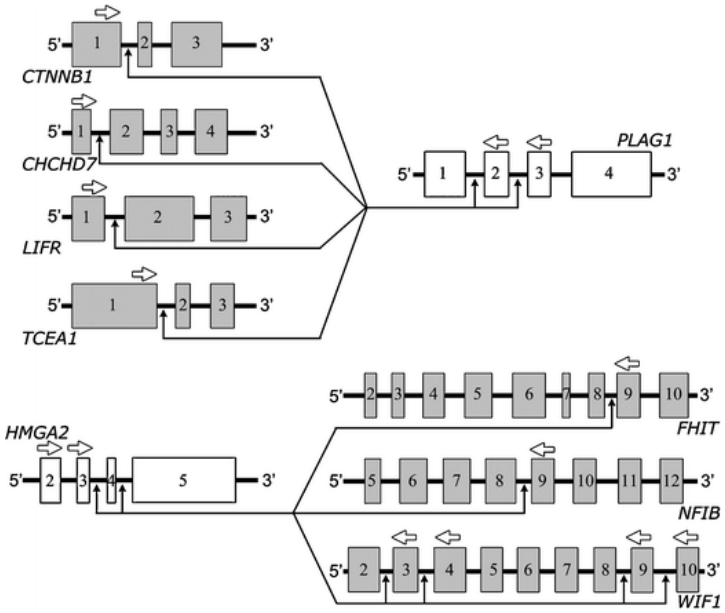


Figure 1.10 Schematic representation of the breakpoints in fusion genes involving PLAG1 and HMGA2. This diagram illustrates fusion gene events resulting from chromosomal rearrangements, where exons from different genes are joined to create hybrid genes with potentially altered functions. Exons are represented as numbered boxes, with arrows indicating transcription direction. Several genes, including CTNNB1, CHCHD7, LIFR, and TCEA1, undergo rearrangement and fuse with PLAG1, a transcription factor frequently implicated in tumorigenesis. These fusion events can lead to the dysregulation of PLAG1, contributing to abnormal cell growth. Similarly, HMGA2, a chromatin-modifying gene, is shown fusing with different partners such as FHIT, NFIB, and WIF1, which may disrupt tumour suppressor functions or activate oncogenic pathways. The fusion of these genes can have significant biological implications, particularly in cancer development, by altering gene expression and functional protein interactions.
. Image source: Matsuyama et al., 2011.

Unsurprisingly, the malignant Ca ex PA shares the same translocation events and fusion genes as PA. This includes PLAG1 and HMGA2 and their gene fusions. However, the high-grade tumours are, in most cases, characterised by gene instability and may suffer copy number alterations. de Brito et al (2016) found that expression of PLAG1 was maintained in the minimally invasive Ca ex PA but lost in the more aggressive forms indicating the activation of other pathways. Thus, the loss or lack of PLAG1 gene expression could be a potential hallmark of Ca ex PA carcinogenesis.

The progression of malignant Ca ex PA is preceded by a carcinogenesis process of multiple steps demonstrating a gradual loss of heterozygosity (LOH, loss of one parental copy of a genome) firstly at chromosomal arm 8q, followed by 12q, and eventually 17p. El-Naggar et al (2000) studied the genetic material of PA and Ca ex PA cases and found that PA and the benign part of Ca ex PA demonstrated greatest loss on chromosomal arm 8q, with less frequent losses on 12q and 17p. Nevertheless, the carcinoma elements in particular displayed exactly the same or insignificantly increased LOH at chromosome 8q and considerably higher degrees of LOH on 12q and 17p. Therefore, LOH on chromosomal arm 8q could be suggestive of early or initial events in PA, while loss on chromosomal arm 12q could identify the category of benign PA with increased likelihood for becoming malignant carcinoma. Chromosome arm 17p alterations were detected in the advanced stages of carcinogenesis. As previously mentioned, alterations of chromosomal arm 8q in PA generally involve PLAG1 (8q12.1) and MYC (8q22.1–q24.1) (Röijer et al., 2002; Martins et al., 2005). However, the process of malignant transformation of a PA to Ca ex PA could be attributed to the 12q genes (HMGIC, HMGA2 and MDM2). These genes may interact and synergize to drive tumorigenesis. For example, PLAG1 and MYC overexpression may cooperate to promote cell proliferation and invasion. Additionally, alterations in MDM2 could contribute to the malignant transformation of PA by inactivating p53, allowing for uncontrolled cell growth (Zhang et al., 2020).

Given the frequent occurrence of chromosomal aberrations involving band 8q12, targeting the transcription factor PLAG1, in approximately 40-50% of PA cases, this study aimed to investigate the role of PLAG1 in the development and progression of salivary gland PA. To elucidate the functional consequences of PLAG1 overexpression, we will conduct in vitro experiments by transfecting primary salivary gland cells or established cell lines with PLAG1 expression vectors.

1.4 Project hypothesis, aims and objectives

1.4.1 Hypothesis

Previous studies have demonstrated that chromosomal translocations frequently lead to PLAG1 overexpression in PA. In this study we hypothesise that PLAG1 overexpression is a key driver of

PA pathogenesis. The specific molecular signature associated with PLAG1 overexpression, including altered gene expression patterns and downstream signalling pathways, contributes to the unique phenotypic characteristics of PA. Investigating the function and mechanism of action of PLAG1 would enable a better understanding of the biological effects of its *de novo* expression, the pathogenesis of PA, and may potentially identify novel diagnostic and therapeutic targets.

1.4.2 Aims

The overall aim of this project is to investigate the role and mechanism of action of PLAG1 overexpression in the tumorigenesis of PA, focusing on the associated molecular signature and downstream signalling pathways and altered functions. This study specifically examines PLAG1 overexpression in human primary salivary gland cells and the HEK293 cell line.

1.4.3 Objectives

1. Investigate the effects of de novo PLAG1 expression on gene expression patterns and downstream pathways in transfected cells.
    - Conduct transcriptomic analysis to identify differentially expressed genes and affected pathways.
2. Identify novel protein interactions with PLAG1 and study their downstream effects on cellular pathways and biological processes.
    - Utilise biotin-based proximity labelling to identify interacting proteins.
    - Analyse the impact of these proteins on signalling pathways and cellular functions.
3. Investigate the effects of PLAG1 overexpression on cell behaviour.
    - Perform functional assays to assess the impact of PLAG1 on cell proliferation, survival, migration, invasion, colony formation and apoptosis.

2 Chapter 2 Materials and methods

2.1 Cell culture work

2.1.1 Ethics

Normal salivary gland tissue was collected from patients undergoing surgery for the excision of salivary gland tumours in the parotid, submandibular and sublingual salivary glands. The request for the isolation and use of primary salivary gland cells was approved by the National Research Ethics Committee NREC (approval number 13/NS/0120).

2.1.2 Cell culture reagents

Table 2.1 details frequently used cell culture supplements, sourced from Sigma-Aldrich, Gillingham, UK unless otherwise specified.

Table 2.1 Cell culture reagents.

Reagent	Catalogue number	Uses
Dulbecco's modified Eagle's medium (DMEM)	D5546	Routinely used for growing
L-glutamine	G7513	Routinely used for growing
Dulbecco's phosphate-buffered saline (PBS)	P4417	Routinely used for washing
Trypsin/EDTA (ethylenediaminetetraacetic acid)	T3924	Routinely used for expanding
Foetal bovine serum (FBS)	10270-106	Routinely used for growing
Nutrient mixture F-12 Ham	N4888	Routinely used for growing
Epidermal growth factor (EGF)	E9644	Routinely used for growing
Human insulin	I9278	Routinely used for growing
Hydrocortisone	H0396	Routinely used for growing
Adenine	A2786	Routinely used for growing
Penicillin/ streptomycin	P0781	Routinely used for growing
Dimethyl sulfoxide (DMSO)	D2650	Routinely used for freezing

2.1.3 Cell lines and culture media

All cell culture work was carried out in class 2 laminar hoods under strict aseptic techniques. Table 2.2 details the cell lines and primary cells used in this study and their corresponding culture media.

Table 2.2 Cell lines and culture media used in this study.

Cells	Type	Culture medium
Human parotid salivary gland (PG)	Primary	Keratinocyte Growth Medium (KGM)
Human sublingual salivary gland (SLG)	Primary	Keratinocyte Growth Medium (KGM)
Human submandibular salivary gland (SMG)	primary	Keratinocyte Growth Medium (KGM)
BMI transduced human foetal parotid salivary gland	Transduced primary cells (extended proliferative potential)	Keratinocyte Growth Medium (KGM)
BMI transduced human foetal sublingual salivary gland	Transduced primary cells (extended proliferative potential)	Keratinocyte Growth Medium (KGM)
HEK293	Immortalised human embryonic epithelial kidney cells.	Low glucose DMEM supplemented with 10% (v/v) FBS

2.1.4 Cell preparation

Primary salivary gland cells

The primary cells were cultured from tissues obtained from normal salivary glands during human surgical biopsies and isolated by Dr. Lynne Bingle. Briefly, the obtained piece of tissue was cut into small pieces using sterile scalpel or scissors. The fragments were then treated with trypsin for 1 hour at 37°C in a water bath, with intermittent shaking of the tube, to release individual cells. The trypsin was inactivated by incubating the isolated cells and remaining tissue in DMEM with

10% FBS for 5 minutes before mechanical dissociation of the tissue by gentle pipetting. The resulting cell suspension was filtered to remove clumps, centrifuged to pellet the cells and finally, the isolated cells were resuspended in culture medium, plated in a culture dish and incubated under optimal conditions for growth and proliferation (37°C with 5% CO_2).

BMI transduced human foetal salivary gland cells

The transduction experiment outlined below was kindly conducted by our collaborators at the Genetic and Genomic Medicine Department, University College London, under the supervision of Prof Stephen Harts.

To enhance their proliferative capacity, human foetal parotid and sublingual salivary gland cells were transfected with a lentivirus containing Bmi-1 integrated into a pLVX-Puro vector (Clontech, 632164) digested with XhoI and BamHI (L-BMI-Puro).

Primary cells at early passage (P4) were seeded into 6 well plates at a density of 100,000 cells per well and incubated overnight at 37°C with 5% CO_2. The following day, 1 ml of LV-BMI-Puro solution (stock: $2x10^8$ TU/ml) was added to the well at varying multiplicities of infection as follows:

1. MOI 1 (0.5 µl LV-BMI-Puro in 200 µl OPTI-MEM+ 800 µl ExPlus)
2. MOI 4 (2 µl LV-BMI-Puro in 200 µl OPTI-MEM+ 800 µl ExPlus)
3. MOI 16 (8 µl LV-BMI-Puro in 200 µl OPTI-MEM+ 800 µl ExPlus)

The plates were then incubated overnight at 37°C with 5% CO_2. The next day, 1 ml of ExPlus media was added to each well, bringing the total volume to 2 ml per well. Once the cells reached confluency in the 6 well plates, they were transferred to T75 flasks.
To select successfully transfected cells, in the T75 flasks, puromycin (Gibco™, Catalogue number A1113803) was introduced at concentrations of 2.5 and 5 µg/ml. After 24 hours, the cells were washed, and fresh medium containing puromycin was added. The cells were then subjected to puromycin treatment for around two weeks.

Keratinocyte growth medium (KGM) was used for the cultivation of normal primary and BMI-transduced salivary glands, including the parotid, submandibular, and sublingual glands. The medium comprises Low glucose Dulbecco's modified Eagle's medium (DMEM) supplemented with 23% (v/v) Ham's F12 (Sigma-Aldrich), 10% (v/v) Foetal bovine serum (FBS), 100 µg/ml penicillin, 100 U/ml streptomycin, 2 mM L-glutamine, 180 µM adenine (Sigma-Aldrich), 0.5 µg/ml hydrocortisone (Sigma-Aldrich), and 10 ng/ml epidermal growth factor (EGF; Sigma-Aldrich). HEK293 cells were cultured in low-glucose Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% (v/v) FBS, 100 µg/ml penicillin, 100 U/ml streptomycin, and 2 mM L-glutamine.

2.1.5 Thawing of cells

Upon retrieval from liquid nitrogen storage, the cryovial containing the cells was immediately thawed in a water bath maintained at 37°C for less than 1 minute. Next, the cells were aseptically transferred to centrifuge tubes and resuspended by the addition of 10 ml of pre-warmed growth medium. Following this, centrifugation was performed at 1000 rpm for 5 minutes. The supernatant was discarded, and the cell pellet was resuspended in the appropriate growth medium. The resuspended cells were then seeded into a T75 flask and placed in an incubator regulated at 37°C with 5% CO2. Cells were regularly monitored, and the medium replaced to eliminate any dead cells or cell debris.

2.1.6 Routine culture and maintenance of cells

All cells were grown as adherent monolayers. Growing cells were monitored on a regular basis and the spent medium was changed every two to three days. Cultured cells were passaged after reaching 70-80% confluent. Exhausted culture medium was discarded, cells washed twice with modified sterile calcium- and magnesium-free Dulbecco's Phosphate buffered saline (PBS, 2x5ml) (catalogue number D8537, Sigma- Aldrich), and then incubated for 3-5 minutes at 37°C with 5% CO_2 in pre-warmed Trypsin/EDTA (0.05% trypsin/ 0.02% EDTA w/v) (catalogue number T3924, Sigma-Aldrich) to completely detach the adherent cells. Pre-warmed fresh medium containing 10% FBS, double the volume of the trypsin, was added to neutralise the effect of the enzyme. The

cell solution was moved to a 50ml conical tube and centrifuged at 1000rpm for 5 minutes. The supernatant was decanted, and the cell pellet was re-suspended by adding 10 ml of fresh media. Cell counting was carried out using a haemocytometer, and the total number of cells estimated as described below. The cells were then split at different ratios based on the seeding density needed. Finally, cells were diluted with pre-warmed growth medium and moved to new vessels and incubated at 37°C with 5% CO_2.

The number of cells/ml = The number of cells X dilution factor X 10^4 /4

2.1.7 Freezing of cells

The cell pellets were collected as described in section 2.1.6. After discarding the supernatant, the pellet was gently mixed with freezing media containing FBS and 10% dimethyl sulfoxide (DMSO) (catalogue number D2650-100ML, Sigma Aldrich). 1 ml of the suspension was then transferred to cryo-vials and placed in a Nalgene freezing container (ThermoFisher, UK) filled with isopropanol to enable slow freezing of the cells in a -80°C freezer for 24 hours. Subsequently, the cryo-vials were transferred to liquid nitrogen for long-term storage.

2.1.8 Cell pellet preparation

Upon reaching 80% confluency, the used media was removed from the cells, which were subsequently subjected to two washes with PBS. Following this, 5 ml of ice-cooled PBS was added, and the cells were gently dislodged from the bottom of the flask and pelleted by centrifugation at 1000 rpm for 5 minutes. After discarding the supernatant, the resulting cell pellet was re-suspended in RNA lysis buffer (Monarch® Total RNA Miniprep Kit, New England BioLabs® Inc) using RNase-free filtered tips for RNA extraction. For basic protein analysis, the cell pellets were lysed directly in RIPA lysis buffer (catalogue number sc-24948A ChemCruz, Santa Cruz Biotechnology) supplemented with protease and phosphatase inhibitors (catalogue number 04693159001, Roche).

2.2 Cloning

2.2.1 Constructs

2.2.1.1 FLAG Tagged PLAG1 pCDNA 3.1+IREs GFP vector

The use of pCDNA 3.1+IREs GFP plasmid allows the generation of an expression construct in a fused RNA that uses an internal ribosomal entry site so that a single RNA transcript generates two different proteins. Essentially this means that all green cells should also co-express the protein of interest**.** A full length human PLAG1 sequence was synthesised by Biomatik Corporation©, Ontario, Canada using the PLAG1 RefSeq sequence with NheI and XhoI restriction sites at the 5' and 3' ends of the sequence. A FLAG Tag was added to the C-terminus immediately before a stop codon. The expression vector, pCDNA 3.1+ IRES GFP was obtained from Addgene, USA (catalogue number 51406). Both plasmids were cut using NheI and XhoI restriction enzymes to produce products with compatible sticky ends. The T4 ligase enzyme (Promega, USA) was used to ligate the two DNA fragments, the expression vector and the PLAG1 insert.

The recombinant plasmids generated in the ligation reaction were transformed into NEB® 5-alpha Competent E. coli (High Efficiency) (New England Biolabs, UK) to produce recombinant plasmid DNA. A volume of 10-100 µl of transformed bacterial suspension was spread on an LB (Luria-Bertani) Agar plate (40 grams of LB Agar powder in 1 litre of distilled water) with ampicillin resistance and then incubated at 37°C overnight. Single bacterial colonies were picked, cultured in autoclaved 2ml LB broth (10 g of LB broth powder in 400 mL of distilled water) (Catalogue number 1289-1650, Fisher ScientificTM, UK) supplemented with 50 µg/ml ampicillin, and incubated overnight at 37°C on a shaking incubator.

To extract the PLAG1 plasmid a mini-prep Plasmid Kit (catalogue number SK-PLPU100; Eurogenetic, UK) was used according to the manufacturer's instructions. The successful construction of the PLAG1 expression plasmid (FLAG-tag PLAG1 pCDNA3.1+ IREs GFP) was confirmed through diagnostic restriction enzyme digestion using an internal restriction enzyme (BamHI) and sequencing.

2.2.1.2 Flag TurboID PLAG1 construct

The Flag PLAG1 TurboID pcDNA3.1(+) fusion construct was custom-designed by Biomatik Corporation© (Ontario, Canada). An empty Flag TurboID pcDNA3.1(+) vector (EV) was used as a negative control. The recombinant construct was verified by complete sequencing by Biomatik and by restriction enzyme digestion using NheI (GCTAGC) - XhoI (CTCGAG) restriction enzymes which flanked the synthesised sequence. The gene synthesis reports and detailed sequences of both constructs are provided in the appendix.

2.2.2 Plasmid isolation 'Midiprep'

In order to obtain a sufficient amount of DNA for the transfection and long-term storage of the FLAG tag PLAG1 pCDNA 3.1+IREs GFP, a 2 ml aliquot of the previously transformed mixture was cultured in sterile LB broth supplemented with 50 mg/ml ampicillin (Sigma-Aldrich, UK) and incubated overnight at 37°C in a rotary shaker at 225 rpm. Subsequently, a glycerol stock of the bacterial culture was prepared by combining 500 µl of autoclaved 50% glycerol solution with 500 µl of the bacterial culture in cryo-vials, which were then stored at -80°C. The bacterial cells were harvested by centrifugation at 3,400× g for 10 minutes at 20°C. Following this, plasmid DNA extraction and purification were conducted using the ZymoPURETM Plasmid Midiprep kit (ZymoPURE, USA). The DNA concentration was determined utilising a Nanodrop spectrophotometer, and the resultant samples were subsequently stored at -20°C.

2.2.3 jetOPTIMUS® Transfection

HEK293 cells were seeded into 6-well plates at a density of 5×10^5 cells per well in antibiotic-free growth media and subsequently incubated overnight at 37°C with 5% CO_2. Twenty-four hours later, the media was replaced prior to the transfection procedure. The cells were either transfected with FLAG-tag PLAG1pCDNA3.1+ IREs or mock-transfected with an empty vector as a negative control. Two µg of DNA were diluted in 200µl of transfection buffer jetOPTIMUS® (Cat# 101000025; Polyplus, UK). A 1:1 ratio of DNA to transfection reagent was previously determined to give the best transfection efficiency and so was used for these experiments. The transfection

mixture was incubated at room temperature for 10-15 minutes, followed by the addition of 200µl of the transfection solution to the designated wells. After 4 hours, the media was replaced with fresh media, and the cells cultured for an additional 24-48 hours. Cell viability and green fluorescence expression were assessed using an inverted fluorescent microscope (Axiovert 200M, Zeiss). Transfection efficiency was evaluated either by fluorescence-activated cell sorting (FACS) or through microscopic observation. Subsequently, the cells were lysed directly in RIPA lysis buffer supplemented with a protease inhibitor for downstream analysis.

2.2.3.1 Evaluation of the timeframe of PLAG1 protein expression level

Before conducting substantial experiments, it was essential to determine the timeframe for PLAG1 protein expression. Transfection was carried out following the protocol described above with two sets of samples, including a negative control, being prepared and monitored at different time points. One set was lysed 24 hours after transfection, while the other was monitored for green fluorescence over 7 days before being lysed. PLAG1 protein expression was assessed using western blotting to evaluate levels over time.

2.2.3.2 Large-scale transfection

When a greater quantity of cells was required for downstream functional analysis, the transfection process was conducted in T75 flasks. The protocol was appropriately adjusted to accommodate the larger scale, as delineated in Table 2.3.

Table 2.3 Quantity required for T75 transfection.

Culture vessel	**flask 75 cm2**
Volume of medium during transfection	10 mL
Volume of jetOPTIMUS® buffer	1000 µL
Amount of DNA added	10 µg
Volume of jetOPTIMUS® reagent	10 µL (1: 1 ratio)

2.2.4 Fluorescence-activated cell sorting (FACS)

To assess transfection efficiency, transfected and control cells were subjected to flow cytometry (FACS) analysis. Cells were initially washed twice with phosphate-buffered saline (PBS), trypsinised, and resuspended in growth medium. Subsequently, cell sorting was performed on a BD FACS Aria Ilu cell sorter (Research Core Facility, University of Sheffield) to discriminate between GFP-positive (transfected) and GFP-negative (control) cell populations. This process was conducted by Sue Clark (Flow Cytometry Core Facility technician). Un-transfected and empty vector-transfected cells served as controls to establish a baseline fluorescence threshold, minimising false positive GFP signals.

2.2.5 Confirmation of cell identity - STR profiling

Two different batches of primary adult parotid salivary gland cells (PA-P6 and PA-P12) were resuspended in KGM media, spotted onto FTA cards labelled with the culture type and cell concentration, and sent to NorthGene™, Newcastle, UK, for analysis.

2.3 Gene expression analysis

2.3.1 RNA extraction, purification and quantification

HEK293 cells were transiently transfected, as described in section 2.2.3. Total RNA extraction was performed using a Monarch total RNA isolation kit (BioLabs, UK).

Prior to RNA isolation, the cell culture media was discarded, and the cells underwent 2-3 washes with ice-cold PBS to minimise sample degradation. Next, 300 µl of RNA lysis buffer was added to each well. The resulting lysate was harvested using cell scrapers and transferred into genomic DNA (gDNA) removal columns, followed by centrifugation at 16,000 x g for 30 seconds. Subsequent to the removal of gDNA columns, an equivalent volume of ethanol (≥95%) was added to each flow-through, mixed via pipetting, and then transferred to RNA purification columns, which were centrifuged at 16,000 x g for 30 seconds. The following step involved washing each

column with 500 µl of wash buffer and centrifuging for 30 seconds at 16,000 x g. A DNase I mixture, comprising 5 µl of DNase I with 75 µl of DNase I reaction buffer for each column, was applied to ensure complete removal of genomic DNA, followed by incubation at room temperature (RT) for 15 minutes.

After the incubation, each column was loaded with 500 µl of RNA priming buffer and centrifuged for 30 seconds at 16,000 x g and then washed with 500 µl of RNA wash buffer. The columns were centrifuged twice at 16,000 x g for 30 seconds and 2 minutes, successively. Each column was then placed in a sterile and RNase-free Eppendorf tube, loaded with 50 µl of nuclease-free water, incubated at RT for at least 1 minute, and centrifuged at 16,000 x g for 1 minute. The columns were then discarded, and the flow-through containing RNA was placed on ice.

The quality and yield of the RNA was assessed using a Nanodrop 1000 spectrophotometer (Thermo Scientific) and samples stored at -80°C for long-term storage.

2.3.1.1 Reverse Transcription (RT) of mRNA to cDNA

Total RNA was reverse transcribed to cDNA using a High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, USA) following the manufacturer's protocol with 500 ng of RNA being used for each reverse transcription reaction (Table 2.4). Two reaction mixtures were prepared: one with reverse transcriptase and the other without reverse transcriptase, which served as a negative control. Both reactions were incubated in a thermal cycler (DNA Engine, DYAD) at 25 °C for 10 minutes, 37 °C for 120 minutes, and 85 °C for 5 minutes, and then cooled to 4 °C. cDNA was used immediately or stored at 80°C for long-term preservation.

Table 2.4 Reverse Transcription (RT) master mix reagents.

Maset mix reagents	Volume per sample
Nuclease free water	4.2 µl
Deoxynucleotides (dNTPs)	0.8 µl
Random primers	2 µl
Reverse transcriptase buffer	2 µl
Reverse Transcriptase (Multiscribe)	1 µl
Total volume	10 µl

2.3.1.2 Quantitative Real-Time Polymerase Chain Reaction (qPCR)

To assess the expression levels of the genes of interest, real-time quantitative polymerase chain reaction (qPCR) was conducted using the cDNAs and TaqMan primers purchased from ThermoFisher Scientific (Table 2.5). The master mix for each target gene was prepared for each sample as per the specifications provided in Table 2.6.

Table 2.5 TaqMan primers.

Primer	Assay ID and catalogue number
PLAG1	Hs00965049_g1, 4351372
FLNC	Hs00155124_m1, 4331182
CRABP2	Hs00275636_m1, 4331182
ARC	Hs01045540_g1, 4331182

Table 2.6 . Real-time qPCR Taqman master mix reagents.

Master mix reagents	Volume per sample
Nuclease free water	3.5 µl
Taqman primer	0.5 µl
GAPDH endogenous control (catalogue number 4326317E)	0.5 µl
qPCRBIO Probe Blue Mix Lo-ROX (catalogue number PB20.25-05, qPCRBIO)	5 µl
Total volume	9.5 µl

In triplicate, 0.5 µl of each sample was added to a PCR tube (Fisherbrand, catalogue number 14230225), followed by the addition of 9.5 µl of master mix to achieve a final volume of 10 µl per tube. Next, the tubes were inserted into the Rotor-Gene Q PCR system (Qiagen, Germany), and a two-step program was initiated, as outlined in Table 2.7.

The quantification of gene expression was performed using the delta-delta CT (2-ΔΔCT) method established by Livak and Schmittgen in 2001. The cycle threshold denotes the cycle number at which the fluorescent signal reaches a predetermined threshold. The ΔCT value represents the difference in CT values between the target genes and the endogenous controls. The calculated 2-ΔΔCT values depict the relative fold change in gene expression between the samples. The data will be presented as the fold-change in target gene expression in comparison to the endogenous controls.

Table 2.7 Thermal cycle settings for qPCR machine

Setting	Step 1	Step 2	Step 3	Cycle
Temperature	95°C	95°C	60°C	40
Time	10 minutes	10 Seconds	45 Seconds	

2.4 Protein Expression Analysis

2.4.1 Protein extraction

Before extracting protein, tissue culture plates were placed on ice, the cell culture media was removed, and the cells were washed 2-3 times with ice-cold PBS to prevent sample degradation. After removing the PBS, RIPA lysis buffer (catalogue number sc-24948A from ChemCruz, Santa Cruz Biotechnology) supplemented with both protease and phosphatase inhibitors (catalogue number 04693159001, Roche) was added to each well. The volume of lysis buffer used differed according to plate size and density of the cells. Cell lysates were collected using cell scrapers and transferred to sterile 1.5 ml microfuge tubes. After 30 minutes of incubation on ice, the samples were centrifuged at maximum speed for 10 minutes at 4°C. The supernatants were then transferred to new sterile Eppendorf tubes and kept on ice for the next steps or stored at -80°C.

2.4.2 Protein quantification

The total protein concentration was determined using the Pierce bicinchoninic acid assay (BCA) (catalogue number 23225, Thermo Fisher Scientific, UK) in accordance with the manufacturer's instructions. 2-10 µl of lysate and 190 µl of the mixed BCA reagents were added to each well of a 96-well plate. Simultaneously, in the same plate, 10 µl of bovine serum albumin (BSA) standards and 190 µl of the mixed BCA reagents were added. The protein concentration of the standards ranged from 0 mg/ml to 2 mg/ml. The reagent mix was created by combining BCA reagents A and B in a 50:1 ratio. The plate was then placed in an incubator at 37°C for 10-15 minutes to allow the colourimetric reaction to develop. The protein concentration was estimated using a plate reader (Infinite M200, TECAN) at a wavelength of 562nm. The absorbance values were assessed in Excel through interpolation, with the absorbance of each sample compared to that of the standards plotted in a standard curve.

2.4.3 Gel preparation

10-12% SDS-polyacrylamide gels were prepared in plastic cassettes (Invitrogen Ltd, UK) to a thickness of 1 mm, as outlined in Table 2.8.

Table 2.8 Recipe of 10-12% acrylamide gels.

Resolving Gel					
Gel%	40% Acrylamide (ml)	Lower TRIS Buffer 1.5M Tris 0.4% SDS TRIS BASE 45.425g SDS 1g (pH 8.8) (ml)	H2O (ml)	TEMED (µl)	APS 10% (µl)
10	2.475	2.5	4.825	5	350
12	1.2	2.5	6.1	5	350
Stacking Gel					
	Acrylamide (ml)	Upper TRIS Buffer 0.5M Tris plus 0.4% SDS TRIS BASE 15.15g SDS 1g pH (6.8) (ml)	H2O (ml)	TEMED (µl)	APS 10% (µl)
	1	2.1	4.7	17	100

2.4.4 Sample preparation

To prepare the protein samples for western blot analysis, the volume of lysate needed to obtain the desired amount of protein was calculated by dividing the required protein amount by the predetermined concentration of protein samples, and then mixed with distilled water and 5X SDS loading dye buffer. The final volume was adjusted based on the capacity and number of wells in the gel used for electrophoresis (e.g. 25-30 µl for a 10-well gel, 15-20 µl for a 15-well gel). After mixing, the samples were vortexed and briefly centrifuged. They were then heated in a thermo-block (JENCONS-PLUS) at 95°C for 10 minutes.

2.4.5 Gel electrophoresis

SDS-polyacrylamide gels were inserted into the XCell SureLock Mini Cell (Invitrogen, Cambridge, UK). The tank and the minicell were filled with 1x SDS-Tris-glycine running buffer (25mM TRIS, 250mM Glycine, 0.1%SDS) prior to loading protein samples and also 5µl of the protein ladder (Prime-Step™ Prestained Broad Range Protein Ladder). An initial gel run at 80 V for 20 minutes was followed by a run at approximately 120 V for approximately 60 minutes.

2.4.6 Electro-transfer

The proteins were then transferred onto Trans-Blot Turbo Mini 0.2 µm Nitrocellulose Transfer Packs (catalogue number 1704158; Bio-Rad, UK) using the Trans-Blot Turbo Transfer System (Bio-Rad, UK) with mixed molecular weight settings. After the transfer was complete, and to assess protein transfer, the nitrocellulose membrane was stained with Ponceau solution (catalogue number 27195, Sigma Aldrich), followed by washing with distilled water.

2.4.7 Immunodetection and development

To prevent non-specific binding, the membrane was incubated with 5% skimmed milk (catalogue number 84615.0500; VWR Prolabo Chemicals, UK) dissolved in 1× 10 mM Tris-buffered saline (1×TBS) for 1 hour at room temperature on a shaker.

The blocking buffer was discarded and the membrane incubated overnight at 4°C with primary antibody diluted in fresh blocking solution as shown in Table 2.9. The membranes then underwent three 10-minute washes with TBS-T (Tris-Buffered Saline with 0.1% Tween-20) before probing with secondary antibody diluted in 5% TBS-T milk for 1 hour at room temperature on a rocking platform (Table 2.9).

Table 2.9 List of primary and secondary antibodies.

Primary Antibody	Dilution	Details	Secondary Antibody
Anti-PLAG1	1:1000	Mouse monoclonal antibody (M02), clone 3B7 Catalogue number H00005324-M02; Abnova, UK.	Anti-mouse IgG, HRP-linked Antibody #7076 – Cell Signalling Technology 1:3000
Anti-GAPDH	1:3000	Mouse monoclonal antibody Catalogue number 60004-1Ig; ProteinTech, USA.	Anti-mouse IgG, HRP-linked Antibody #7076 – Cell Signalling Technology 1:3000

Anti-StrepTactin	1:5000	Precision Protein StrepTactin-HRP Conjugate, 125 µl #1610381 -BIO-RAD, UK.	NA

The membrane was then subjected to a triple wash using TBS-T followed by incubation with enhanced chemiluminescent (ECL) Clarity Western ECL substrates (catalogue number 1705060, BioRad) for 1-2 minutes at room temperature. Activation of this substrate is facilitated by the HRP bound to the secondary antibody. The protein was identified utilising a Li-Cor C-Digit Western Blot Scanner and Image Studio Software. In the case of re-probing with alternative antibodies, the membranes were washed with 1× TBS-Tween prior to stripping with 10 ml of stripping buffer (Thermo Fisher Scientific, Cambridge, UK) while being agitated for 15-30 minutes at room temperature.

2.5 TurboID proximity labelling

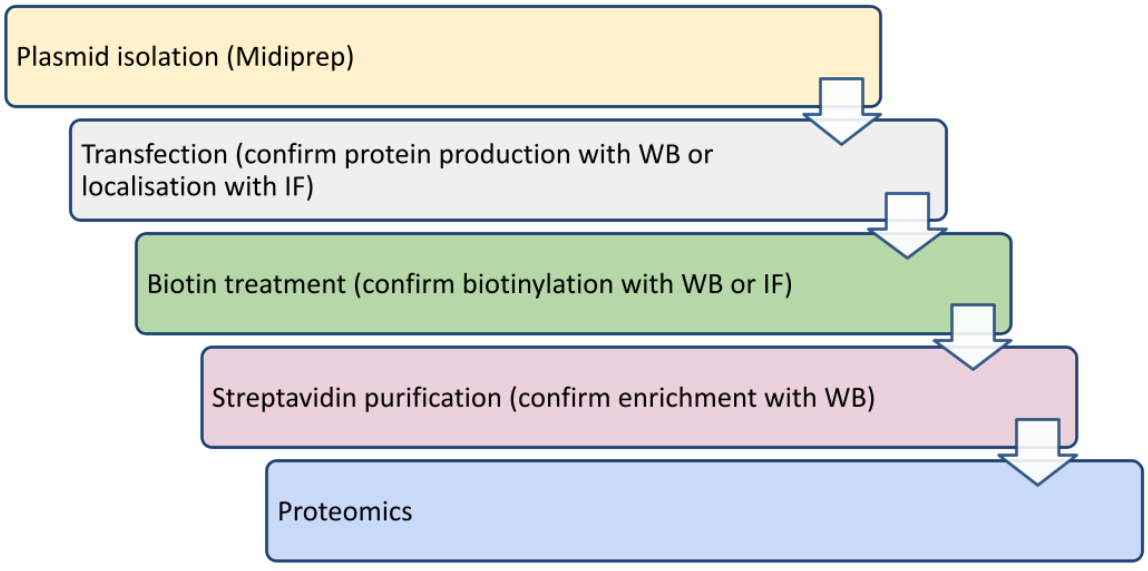


Figure 2.1 Workflow of the proximity labelling technique.
The experiment begins by generating large quantities of plasmid DNA (midiprep). Next, HEK293 cells are transfected with the bait protein (PLAG1) fused to TurboID. After that, the cells are treated with exogenous biotin, and the biotinylated proteins are isolated on streptavidin beads for purification. This process prepares the samples for mass spectrometry analysis.

2.5.1 Optimisation of experimental conditions

A preliminary small-scale experiment was undertaken in a 6-well plate to optimise conditions prior to a large-scale experiment being performed to generate samples for proteomic analysis. The initial designation of this sample was "sample one," and was not included in the current study; the four biological replicates analysed are thus denoted as samples 2, 3, 4, and 5.

2.5.2 Transfection

HEK293 cells were seeded into T75 flasks in DMEM supplemented with 10% FBS and incubated at 37°C with 5% CO_2. When the cells reached 70-80% confluency, the old media was replaced with fresh antibiotic-free media and the transfection carried out as described in section 2.2.3. One flask was transfected with 10 µg of TurboID PLAG1 DNA, while the other flask was transfected with an empty TurboID (EV) construct, serving as a negative control. To visualise green fluorescence and assess transfection efficiency, both flasks were co-transfected with 0.2 µg of a pEGFP-N1 reporter gene (ClonTech Laboratories); transfection efficiency was evaluated after 24hrs using an inverted fluorescent microscope (ZOE Fluorescent Cell Imager – Bio-Rad – UK).

2.5.3 Biotin treatment and Protein extraction

Following successful transfection, the cells were treated with 0.5 mM biotin (Sigma-Aldrich/USA—CAS no. 58-85-5) for 20 minutes and total protein isolated by adding 500 µL of RIPA buffer supplemented with a protease inhibitor, as indicated in section 2.4.1. The total protein concentration of the samples was quantified using a BCA assay as described in section 2.4.2, and the samples were stored at -80°C.

2.5.4 Streptavidin purification (enrichment)

Pierce™ Streptavidin Magnetic Beads were used for labelled protein purification. 100 µl of magnetic beads were washed with 500 µl of RIPA buffer without protease inhibitor, placed into a magnetic stand (EasySepTM Magnet – Stem Cell Technologies, Catalogue number 18000 - UK) to allow the beads to attach to the sides of the tube, and the supernatant collected and discarded.

This process was repeated three times and then500 µg of whole cell lysate was added to the magnetic beads and incubated with end-to-end rotation at 4°C overnight. The remaining material from the whole cell lysate (WL) was saved and stored at -20°C for subsequent western blot analysis. The beads were then pelleted using the magnetic stand, and the supernatant (flow-through) was collected in fresh microcentrifuge tubes for western blot analysis. The beads were then washed with 1 ml of the following solutions:

1. 2% SDS/50 mM Tris pH 7.4
2. Lysis buffer
3. 2M Urea/50 mM ammonium bicarbonate (x2 washes)
4. 50 mM ammonium bicarbonate

Steps 1 and 2 are to remove nonspecific interactions and break up protein complexes, while steps 3 and 4 are to remove detergents. For each wash, the samples were incubated with end-to-end rotation for 5 minutes at room temperature, briefly spun down, and then placed in a magnetic stand to collect and discard the wash buffer. Finally, 100 µl of fresh 50 mM ammonium bicarbonate was added to the beads, and the enriched samples (ES) were then stored at -80°C. The recipe for all buffers is detailed in Table 2.10.

Table 2.10 Streptavidin Beads Washing Buffer Components.

Lysis Buffer – 50 ml	
50 mM Tris-HCL pH8.0	2.5 ml of 1M stock
150 mM NaCl	16.6 ml of 0.5M stock
1% Triton-X100	500 µl of 100% stock
0.1% SDS	250 µl of 20% stock
0.5% sodium deoxycholate	0.25g
Protease Inhibitor	1:200 dilution
Wash buffer 1 – 50 ml	
2%SDS	100 µl of 20% stock
50 mM Tris pH 7.2	50 µl of 1 M stock
Wash buffer 2 – 50 ml	

2% SDS in ddH2O	5 ML of 20% stock
Wash buffer 3 – 50ml	
2 M Urea	25 ml of 4 M stock
50 mM Ambic	5 ml of 0.5 M stock
Wash buffer 4 – 50 ml	
50 mM Ambic	5 ml of 0.5 M stock

2.5.5 Western blot

The efficiency of each experimental step in this protocol was evaluated through Western blot analysis, as detailed in section 2.4. Following transfection, PLAG1 protein expression was detected using an anti-PLAG1 antibody (catalogue number H00005324-M02; Abnova, UK) and protein biotinylation was assessed using an anti-streptactin antibody (StrepTactin-HRP Conjugate; catalogue number 1610381, BIO-RAD, UK). Finally, the efficiency of the protein purification process was evaluated by comparing the expression profiles of the three sample sets: the residual whole lysate (WL) samples, the supernatant (SN) obtained after an overnight bead incubation, and the enriched or purified samples (ES) probed with an anti- StrepTactin-HRP Conjugate antibody.

10% of the final volume of the ES was mixed with 5 µL of 5X SDS lysis buffer, briefly spun, and warmed at 37-45°C for 20 minutes to elute the sample from the beads. 5 µl of 5X SDS lysis buffer was also added to each WL and SN, followed by brief centrifugation and heating at 95°C for 10 minutes.

To ensure that only the eluted proteins were loaded, the tubes of the ES were placed in a magnetic stand to keep the beads attached to the bottom of the Eppendorf tube and then three sets of protein samples, along with 5 µl of the protein ladder (Prime-StepTM Prestained Broad Range Protein Ladder), were loaded onto the gel. The Western blot procedure was continued according to the established protocol outlined in section 2.4.

2.5.6 Immunofluorescence staining

Immunofluorescence staining was conducted to confirm the expression of PLAG1 protein post-transfection using an anti-PLAG1 antibody and to verify protein biotinylation an anti-Streptavidin antibody was used post-biotin treatment.

5 x 10^5 HEK293 cells were plated onto microscope glass coverslips (Chance Propper LTD., England) in 6-well plates and a transfection reaction was carried out as described earlier. The cells were washed three times with PBS for five minutes each, fixed with methanol for 6 minutes at -20°C, washed again three times with PBS for 5 minutes each and stored at 4°C. The plate was sealed with parafilm to avoid contamination.

The cells were incubated with 1 ml of PBST and 3% BSA for 1 hour at room temperature before the primary antibody was added and cells incubated in a humid chamber at 4°C overnight. The conjugated streptavidin antibody (Table 2.11) was incubated with the cells in the dark in a humid chamber at room temperature for 90 minutes.

Coverslips were washed three times with PBS for 5 minutes each, placed on pre-labelled glass slides and mounted with VECTASHIELD® Antifade Mounting Medium with DAPI (©Vector Laboratories, Inc., UK). When completely dry, the coverslips were sealed with nail polish and kept in the dark for 24 hours before visualisation under a fluorescent microscope (Leica Thunder-Imager—LAS X software UK).

Table 2.11 The primary and secondary antibodies used in immunofluorescence staining.

Primary Antibody	Dilution	Details	Secondary Antibody
PLAG1	1:100	Rabbit Polyclonal Antibody Catalogue number 18018-1-AP-150UL; Proteintech, UK.	Cy™3 AffiniPure™ Donkey Anti-Rabbit IgG (H+L); Catalogue number AB_2307443, Jackson ImmunoResearch Europe Ltd., UK.

Streptavidin conjugated with Alexa	1:3000	Streptavidin, Alexa Fluor™ 647 conjugate Catalogue number: S21374, Thermo Fisher Scientific, UK.	NA

2.6 Functional Assays

2.6.1 Proliferation assay

The CellTiter 96® AQueous One Solution Cell Proliferation Assay (MTS; Promega UK (G3580)) was employed to measure the proliferation of viable cells. The MTS assay protocol measures the reduction of the MTS tetrazolium compound by living cells, producing a coloured formazan dye that dissolves in cell culture media. The assay was conducted on both PLAG1 and mock-transfected cells (empty vector). This experimental design allowed us to compare the influence of the PLAG1 ectopic expression on cell proliferation.

Twenty-four hours following transfection, 30,000 HEK293 cells transfected with PLAG1 and EV were seeded in a 96-well plate containing 125µl of growth media and left to settle overnight at 37°C with 5% CO2. 20 µl of CellTiter 96® AQueous One Solution Reagent was added directly to the culture wells, each containing 100 µl of fresh media. Following a one-hour incubation period at 37°C with 5% CO2, the absorbance at 490nm was measured using a microplate reader (Infinite® M Nano, TECAN). A well containing only media and reagent provided a negative control (blank) for the experiment. The quantity of formazan product, determined by measuring the absorbance at 490nm, is directly correlated with the number of viable cells present in the culture. The experiment included three biological replicates.

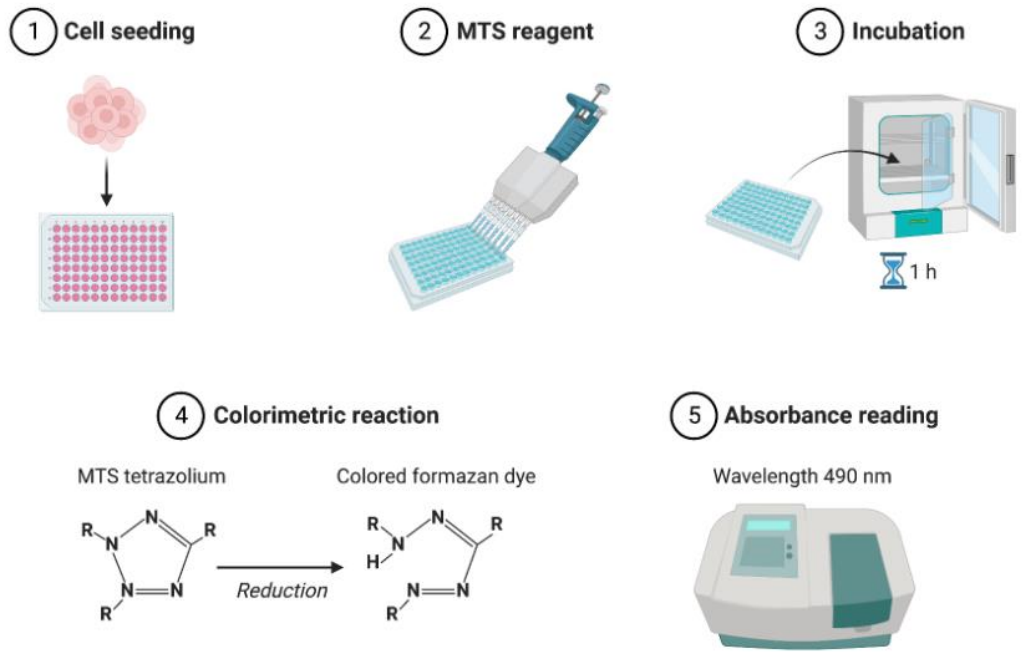


Figure 2.2 Proliferation assay.
Schematic representation of the MTS assay workflow used to assess cell proliferation. (1) HEK293 cells transfected with PLAG1 or an empty vector (EV) were seeded in a 96-well plate. (2) The MTS reagent (CellTiter 96® AQueous One Solution) was added to the culture wells. (3) The plate was incubated at 37°C with 5% CO_2 for one hour to allow the reaction to occur. (4) Living cells reduced the MTS tetrazolium compound to a coloured formazan dye. (5) Absorbance was measured at 490 nm using a microplate reader, with absorbance levels corresponding to cell viability. The negative control consisted of media and reagent without cells. Illustration generated using BioRender.

2.6.2 Transwell migration assay

The transwell migration assay was employed to investigate the migratory behaviour of cells transfected with PLAG1. This assay involved the use of ThinCertMT cell culture inserts, with 0.8 µm pores and designed for 24-well plates (Cat# 662 838; Greiner Bio-one, UK). Twenty-four hours after transfection, 10,000 HEK293 cells transfected with either PLAG1 or EV were seeded onto the ThinCertMT tissue culture inserts in 400 µl of normal growth media. 700 µl of the same media was placed in the bottom chamber and the plate incubated at 37°C with 5% CO2 for 4 hours. The media in the transwell chamber was then replaced with growth media supplemented with 1% FBS, while the media in the lower chamber was retained to serve as a chemoattractant for the

cells. The cells were then incubated at 37°C and 5% CO2 for 48 hours. On the second day, non-migrating cells were carefully removed with a cotton bud, and the remaining cells were fixed with 500 µl of 100% cool methanol for 6 minutes at -20°C. Subsequently, the cells were stained with Harris Hematoxylin (Thermo-Scientific) and Shandon Eosin Y (Thermo-Scientific) as outlined in Table 2.12. Following staining, the insert membranes were excised using a scalpel blade and positioned on SuperFrost Plus slides (Cat # 406/0179/00; VWR International, UK) with migrating cells facing upward. The prepared slides were then mounted with EcoMount mounting media and sealed with coverslips. The experiment was performed three times in triplicate.

The membranes were scanned using a Panoramic 250 Slide Scanner and QuPath software was employed to accurately count and quantify the number of migrating cells. The average score was computed for each membrane, and these values were then subjected to statistical analysis. An Unpaired two-tailed t-test was used to compare the average between the two groups using Graph Pad prism.

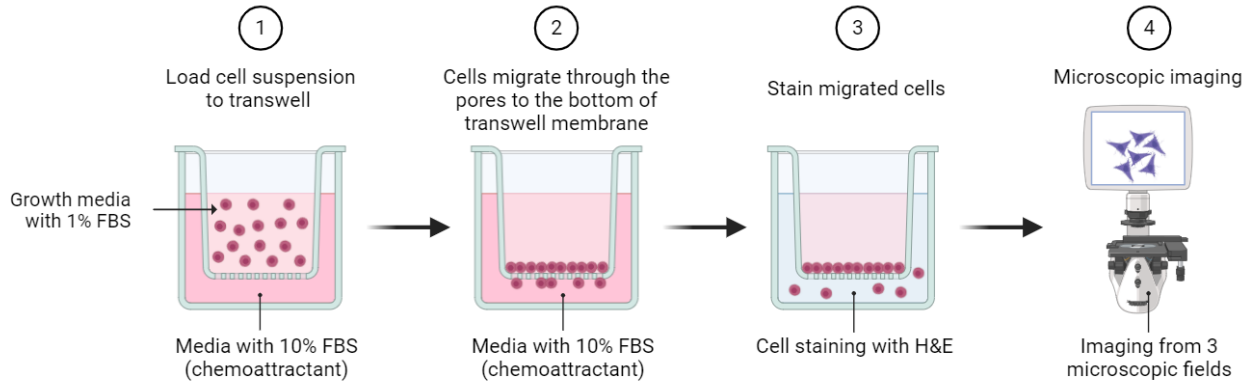


Figure 2.3 Transwell migration assay.
Schematic representation of the transwell migration assay used to assess the migratory behaviour of HEK293 cells transfected with PLAG1 or an empty vector (EV). (1) Cells were seeded in the upper chamber of a transwell insert with growth media containing 1% FBS, while the lower chamber contained media with 10% FBS as a chemoattractant. (2) Cells migrated through the transwell membrane pores toward the chemoattractant over 48 hours. (3) Migrated cells were fixed and stained with Harris Hematoxylin and Eosin (H&E). (4) Microscopic imaging was performed, and three microscopic fields were analysed for quantification. Illustration generated using BioRender.

Table 2.12 Haematoxylin and Eosin staining steps for migrating cells.

Step	Time
PBS	5 min
PBS	5 min
Haematoxylin	2 min
Distilled water	5 min
Distilled water	5 min
Eosin	2 min
Distilled water	5 min
Distilled water	5 min

2.6.3 Transwell invasion assay

The assay was conducted to assess the influence of PLAG1 expression on cell invasiveness with the experimental method employed being similar to the transwell migration assay but uses specialised invasion assay inserts. These inserts are coated with Matrigel to mimic the extracellular matrix (ECM) which the cells must degrade to invade the surrounding structure.

Following effective transfection, $1x10^4$ PLAG1 and EV-expressing HEK293 cells were seeded onto precoated Corning® BioCoat™ Matrigel® Invasion Chambers (Cat# 354480, 354481) in 400 µL of growth media with 10% FBS. 700 µL of the same media was then added to the bottom chamber and the plate incubated at 37°C and 5% CO2 for 4 hours. The media in the transwell chamber was then replaced with growth media supplemented with 1% FBS, while the media in the lower chamber was retained as a chemoattractant. The plate was subsequently incubated at 37°C with 5% CO2 for a period of 3 to 5 days to facilitate cellular invasion. The membranes were fixed, stained, and analysed using the same techniques as described in the migration assay (section 2.6.2).

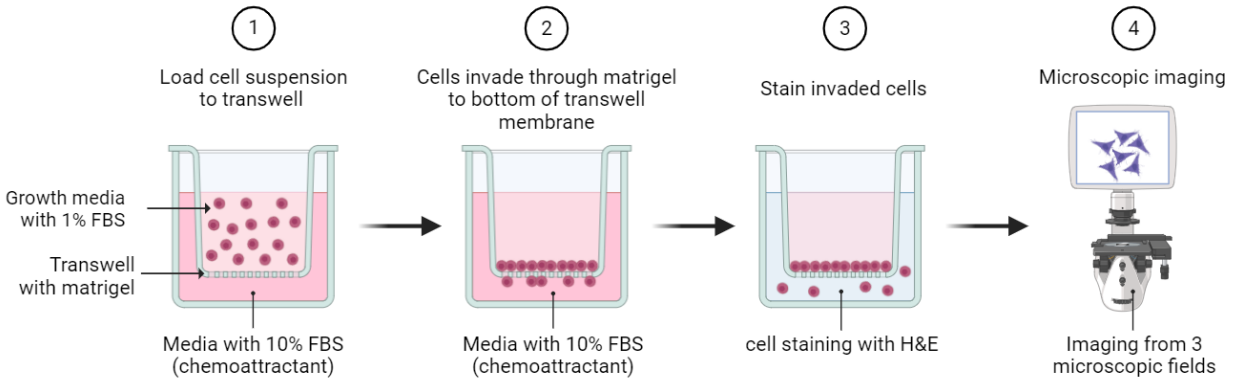


Figure 2.4 Transwell invasion assay.
Schematic representation of the transwell invasion assay used to assess the invasive behaviour of HEK293 cells transfected with PLAG1 or an empty vector (EV). (1) Cells were seeded in the upper chamber of a Matrigel-coated transwell insert with growth media containing 1% FBS, while the lower chamber contained media with 10% FBS as a chemoattractant. (2) Cells invaded through the Matrigel matrix and migrated to the bottom of the transwell membrane over 3 to 5 days. (3) Invaded cells were fixed and stained with Harris Hematoxylin and Eosin (H&E). (4) Microscopic imaging was performed, and three microscopic fields were analysed for quantification. Illustration generated using BioRender.

2.6.4 Colony formation assay

This assay, also referred to as clonogenic assay, is an in vitro cell survival assay used here to evaluate the proliferation potential and survival ability of cells harbouring PLAG1 through the assessment of the capacity of a single cell to undergo unlimited division and the formation of a visible colony.

Cells previously transfected with PLAG1 and empty vector were seeded at a density of 500 cells per well in standard growth medium in a 6-well plate for ten days. Following incubation, the cells underwent two 5-minute washes with PBS and then fixed with 100% Methanol and 100% acetic acid in a 7:1 ratio for 5 minutes at room temperature. The fixative was then aspirated, and a 0.2% Crystal Violet staining solution was added to each well for 2-minutes. Post-staining, the plate was rinsed under running tap water for 3 minutes and allowed to air dry. Colony quantification was

performed using the ImageJ software colony counter, and values were then used for statistical analysis. The experiment was conducted three times in duplicates.

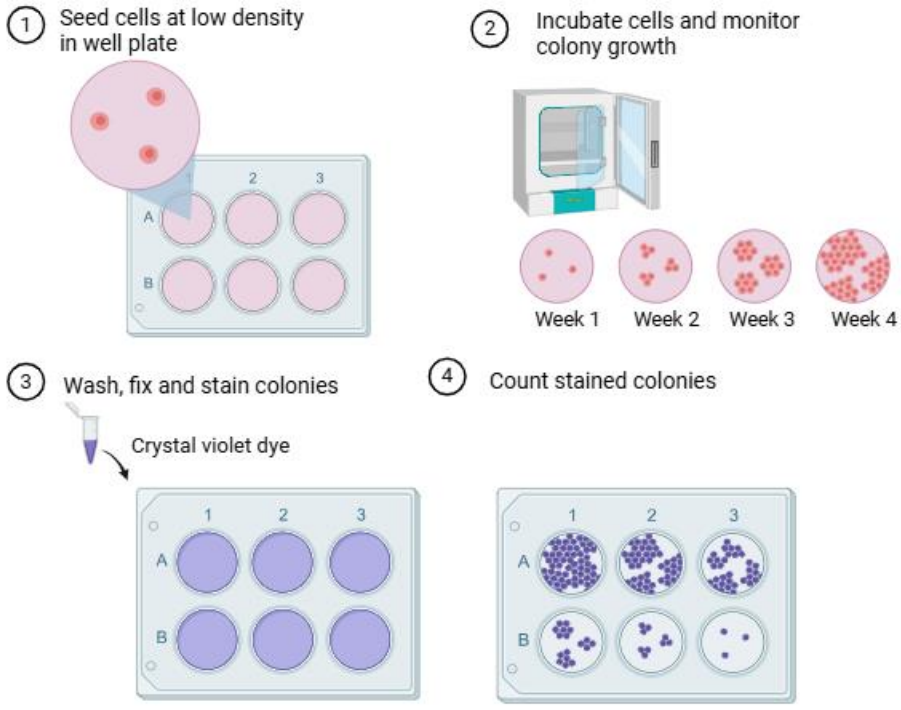


Figure 2.5 Colony formation assay
Schematic representation of the clonogenic assay used to evaluate the proliferation potential and survival ability of cells harbouring PLAG1. (1) Cells previously transfected with PLAG1 or an empty vector (EV) were seeded at a density of 500 cells per well in standard growth medium in a 6-well plate. (2) Cells were incubated for ten days to allow for colony formation. (3) Following incubation, cells underwent two 5-minute washes with PBS, were fixed with 100% Methanol and 100% acetic acid (7:1 ratio) for 5 minutes at room temperature, and stained with a 0.2% Crystal Violet solution for 2 minutes. (4) Post-staining, the plates were rinsed under running tap water for 3 minutes and allowed to air dry. Colonies were quantified using the ImageJ software colony counter. Illustration generated using BioRender.

2.6.5 Apoptosis assay

The Annexin V-FITC Apoptosis Detection Kit (Cat # 4830-01-K; Bio-Techne, UK) was used to identify and measure cellular events linked to programmed cell death (apoptosis) resulting from PLAG1 overexpression. $5x10^5$ HEK293 cells were seeded into 6-well plates and allowed to adhere

overnight at 37°C with 5% CO2. The next day, cells were transfected with PLAG1 and an empty vector and then incubated for 24 hours under the same conditions. Following successful transfection, the apoptosis assay was carried out following the manufacturer's instructions. This experiment involved four control samples, as outlined in Table 2.13. Adherent and floating cells were collected, washed three times with 500 µl cold PBS, and subsequently resuspended in 100 µl of Annexin V Incubation Reagent. The Annexin V Incubation Reagent components and volumes are detailed in Table 2.14. After a 15-minute incubation at room temperature in the dark, 400 µl of 1x Binding Buffer was added to the samples and processed by Cytek® Aurora Flow Cytometer within one hour. The resultant values were then used for statistical analysis using GraphPad Prism software. The experiment was conducted three times in duplicates.

Table 2.13 List of control samples

Control sample	Purpose
Untransfected & unstained	Cell autofluorescence
Transfected & unstained	GFP control
Untransfected & stained with Annexin only	FITC (Annexin) control
Untransfected & stained with Propidium Iodide IP only	PI control

Table 2.14 Annexin V Incubation Reagent Components.

Reaction Component	Volume
10X Binding Buffer	10 µl
Propidium Iodide	10 µl
TACS Annexin V-FITC	1.0 µl
Distilled water	79 µl
Total volume*	100 µl

* *The reagent should be kept on ice in the dark once prepared.*

2.7 Statistical analysis

Data were presented as mean ± standard deviation. An unpaired t-test was used to compare two groups, while a Two-Way Analysis Of Variance (Two-way ANOVA) was applied when comparing

groups, considering two factors. A P-value of less than 0.05 was considered statistically significant. The number of biological repeats is denoted as 'N='. Statistical significance is indicated by asterisks as follows: *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001, and NS for P > 0.05. The statistical analysis was performed using GraphPad Prism 10 software.

2.8 Bioinformatics

2.8.1 Transcriptomic analysis – RNA sequencing

Following successful transfection, 20µl of RNA from three sets of biological repeats was sent to Novogene (Cambridge, UK) for RNA sequencing. The sequencing process and raw data processing were performed at Novogene following standardised protocols to ensure high-quality RNA sequencing data. Initially, quality control of the samples was conducted using two approaches: first, Agarose gel electrophoresis to assess RNA degradation and potential contamination; and second, the Agilent 2100 Bioanalyser to evaluate RNA integrity and quantify the RNA.

For library preparation, mRNA was fragmented by adding a fragmentation buffer. Complementary DNA (cDNA) synthesis was then carried out using an mRNA template and random hexamer primers. The second-strand synthesis was initiated by adding a custom Illumina second-strand synthesis buffer, along with dNTPs, RNase H, and DNA polymerase I. Subsequent processing steps included end repair, A-tailing, and sequencing adapter ligation to prepare the double-stranded cDNA. The final cDNA library was obtained through size selection and PCR enrichment.

Library quality control was performed using three methods: Qubit quantification, Agilent 2100 Bioanalyzer analysis, and qPCR validation. Once the libraries met the required quality standards, they were pooled based on their effective concentration and expected sequencing data output before being loaded onto an Illumina sequencing platform for high-throughput sequencing.

2.8.1.1 Data processing

Row sequencing data in FASTQ format were cleaned by removing low-quality reads, adapter sequences, and poly-N sequences. Quality metrics like Q20, Q30 and GC content were calculated. Clean reads were aligned to a reference genome using Hisat2, a mapping tool that leverages gene model annotations for improved splice junction identification. Read counts were then mapped to genes using FeautreCounts, and FPKM values were calculated to estimate gene expression levels. Next, differentially expressed genes were identified using DESeq2 based on statistical significance and fold change thresholds. P-value were adjusted using the Benjamini Hochberg method to control the false discovery rate (FDR). Gene with an adjusted p-values <= 0.05 were considered differentially expressed.

Expression data was provided in excel sheets, containing read counts, normalised counts FPKM (Fragments Per Kilobase of transcript per Million mapped reads) and statistical analysis results (fold change, p-values and FDR). These data were used for subsequent analysis and visualisation. Volcano plots and heatmaps of the differentially expressed genes were used for data visualisation.

2.8.1.2 Gene enrichment and pathway analysis

To gain insights into the biological significance of our dataset and to identify which biological functions, pathways or processes are overrepresented, functional enrichment analysis was performed. The analysis was carried out using GSEA software v4.3.3 from the Broad Institute (Cambridge, MA, USA) (Subramanian et al., 2005). Normalised expression data was used to perform the analysis. This approach ensures that gene expression levels are comparable across different samples, providing a robust basis for identifying enriched genes.

The analysis compares differentially expressed genes with predefined gene sets. In this study, the molecular data was assessed using the following gene sets:

1. C5: ontology gene sets, which includes gene sets based on the Gene Ontology (GO terms) providing insights into biological processes (BP), cellular components (CC), and molecular functions (MF).

2. C2: curated gene sets, KEGG_LEGACY subset, which focuses on metabolic and signalling pathways essential for understanding cellular processes.
3. C6: oncogenic signature gene sets, which includes genes associated with cancer, helping to identify oncogenic pathways involved in the development and progression of the tumour being studied.

2.8.2 Proteomic analysis – biotin-based proximity labelling

The enriched protein samples prepared according to the procedures outlined in section 2.5 were subjected to Label-Free Quantification (LFQ) Mass Spectrometry analysis at the Biological Mass Spectrometry Facility, University of Sheffield. Four independent biological replicates were analysed.

2.8.2.1 Data processing

MaxQuant quantitative proteomics software was used to analyse the MS data in Sheffield. Label-free abundance values were log2 transformed to bring them closer to a normal distribution. The dataset was filtered such that all proteins contained quantitative information for three replicates of at least one group. The data was then imported into Perseus where missing values were replaced from a normal distribution of the full matrix (width 0.3, down shift 1.8). No normalisation was performed for this data as it is expected that controls and experimental samples should pull down differing total protein abundances, with differing distributions.

2.8.2.2 Statistical analysis

Statistical significance of the quantitative data was determined by a Student's t-test between the conditions of interest. Imputed values were tracked, and qualitative thresholds for significance were applied such that proteins which were identified in all replicates of the condition of interest, and in none of the controls, were considered highly significant. FDR correction was initially applied but found to be too stringent, identifying only nine significant proteins. Given the risk of

excessive false negatives in proteomics, a p-value < 0.05 threshold was used instead to assess statistical significance.

2.8.2.3 Ingenuity Pathway Analysis

Protein expression data with identifiers, expression levels or fold changes, experimental conditions and statistical data were uploaded to Ingenuity Pathway Analysis® (IPA®) (Qiagen Inc.) for downstream pathway analysis. To emphasise the importance of enrichment in the IP condition of interest, and of exclusive protein identification in the IP condition of interest, proteins which were exclusively quantified in the condition of interest were given a p-value of 0, and a LogFC of 10, and all LogFCs below 0 were removed from the dataset, prior to upload to IPA. Core analysis in IPA was performed with the significance cut-off of p-value <0.05 using the IPA Knowledge base reference set.

2.9 Protein-protein interaction analysis using open-source tools (STRING and SIGNOR)

The list of candidate genes obtained from the IPA analysis was searched in the STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) database and SIGNOR (the SIGnaling Network Open Resource) to further investigate and capture any documented interactions between our proteins of interest.

3 Chapter 3 Investigating the downstream effect of PLAG1 overexpression on the gene level: Transcriptomic analysis

3.1 Aims and Objectives

PLAG1, a proto-oncogene, has the potential to trigger the development of pleomorphic adenomas in the salivary glands when expressed ectopically. Being a transcription factor capable of activating gene expression, its overexpression likely disrupts target genes, leading to unregulated cell growth and proliferation. Identifying PLAG1 target genes and exploring their impact on biological processes, functions, and pathways is crucial for comprehending the molecular mechanisms underlying PLAG1-induced tumour development. To achieve this, we examined the alterations in gene expression resulting from PLAG1 overexpression in cell lines. The work outlined in this chapter aimed to:

- Perform gene expression profiling to identify differentially expressed genes (DEGs) following PLAG1 overexpression compared to control cells transfected with an empty vector (EV).
- Conduct gene set enrichment analysis (GSEA) to identify the functional relationships between PLAG1 and the differentially expressed genes DEGs and determine the enriched pathways and biological processes attributed to these genes.
- Validate the most significant differentially expressed genes by qRT-PCR.

3.2 Methods

Following successful transfection and confirmation of PLAG1 protein expression by western blotting, as outlined in Chapter 2, total RNA was extracted, quantified and sent to Novogene (Cambridge, UK) for RNA sequencing and data processing. The list of differentially expressed genes (DEGs) was used for functional enrichment analysis using GSEA software v4.3.3 from the Broad Institute, USA.

3.3 Results

3.3.1 Generation of FLAG Tagged PLAG1 pCDNA 3.1+IREs GFP vector

The FLAG Tagged PLAG1 pCDNA 3.1+IREs GFP vector was generated by inserting a PLAG1 sequence (Biomatik Corporation©, Ontario, Canada) in-frame with the overexpression vector pCDNA 3.1+ IREs GFP (Addgene, USA) (Figure 3.1). An empty pCDNA 3.1+ IREs GFP vector (EV) was used as a negative control.

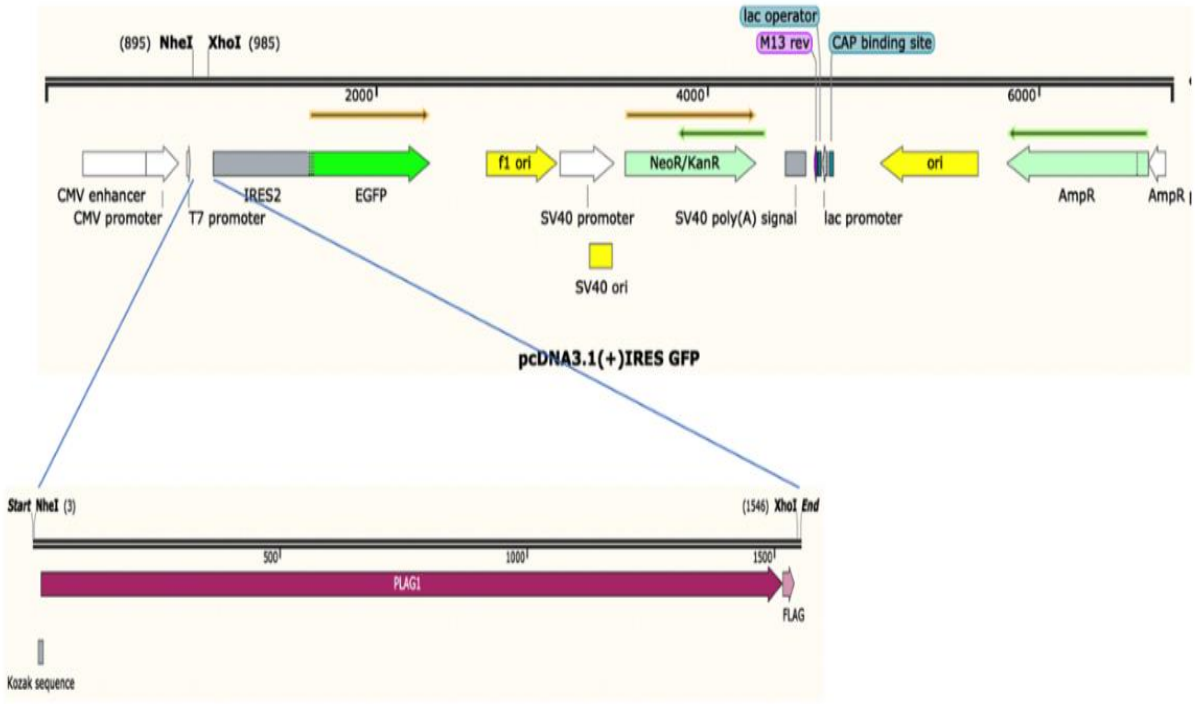


Figure 3.1 Synthesis of PLAG1 pCDNA 3.1+IREs GFP clone.
The FLAG-tagged PLAG1 sequence (represented by the purple arrow) was excised using NheI and XhoI restriction enzymes and subsequently ligated into a predigested pcDNA3.1(+)-IRES-GFP expression vector. The pcDNA3.1(+) vector contains a CMV promoter, which ensures strong transcriptional activation of the inserted PLAG1 gene. The IRES (Internal Ribosome Entry Site) sequence allows for bicistronic expression, enabling the simultaneous translation of EGFP (Enhanced Green Fluorescent Protein), which serves as a reporter. The vector also carries a Neomycin/Kanamycin resistance gene (NeoR/KanR) under the control of an SV40 promoter, allowing for antibiotic selection in mammalian and bacterial cells. Additionally, elements such as the f1 origin of replication (f1 ori) and SV40 polyadenylation signal facilitate replication and transcription termination, respectively. The lac operator and CAP binding site provide regulatory control over bacterial expression. The inserted PLAG1 gene is tagged with a FLAG epitope for easy detection via immunodetection assays. This construct allows for efficient expression and functional analysis of PLAG1 in mammalian systems while enabling GFP-based visualization of transfected cells.

To confirm the correct orientation and integrity of the cloned DNA insert (PLAG1), internal diagnostic digestion was performed. The plasmid was initially cut with XhoI and NheI restriction enzymes to separate the insert from the expression vector. Subsequently the PLAG1 insert was further digested with BamHI, an internal restriction enzyme to cut the insert into 2 smaller fragments. The expected size of the three fragments were 6800 bp representing the pCDNA 3.1+IREs GFP expression vector (XhoI to NheI), 1200 bp (NheI to BamHI) representing the PLAG1 gene fragment and 250 bp (BamHI to BamHI) representing the internally cut PLAG1 gene fragment. The observed band sizes in the gel electrophoresis experiment (Figure 3.2) correspond to these expected values, confirming the successful cloning of The FLAG Tagged PLAG1 pCDNA 3.1+IREs GFP vector.

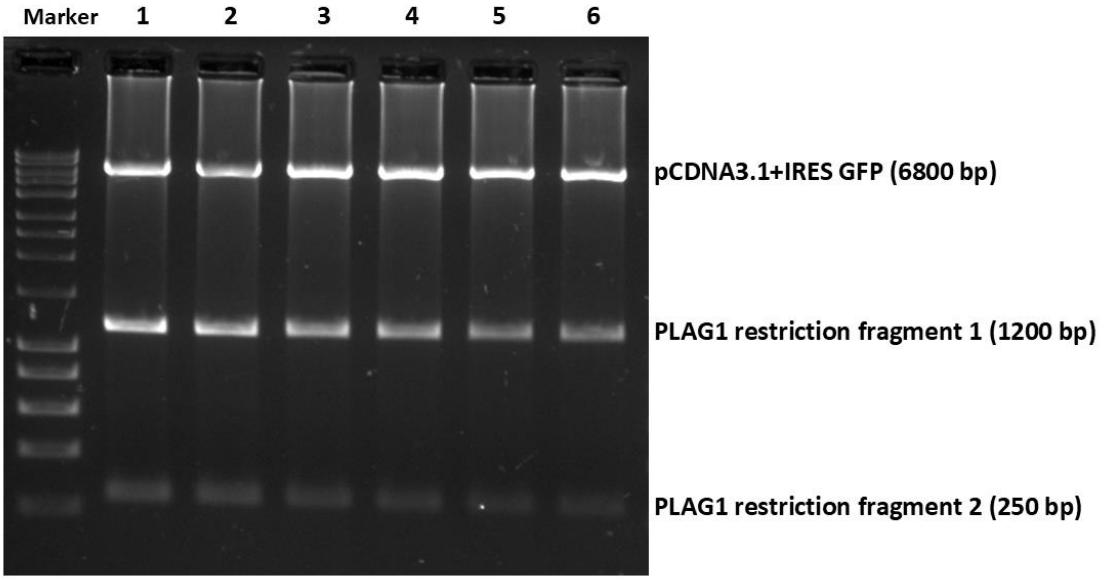


Figure 3.2 Final confirmation of the construct.
The FLAG-tagged PLAG1 pcDNA3.1(+)-IRES-GFP expression vector was digested with NheI and XhoI restriction enzymes, resulting in the separation of the pcDNA3.1(+)-IRES-GFP plasmid backbone (6800 bp, top band) from the PLAG1 insert. To further confirm the presence and integrity of the PLAG1 sequence, a secondary digestion with BamHI was performed, which internally cleaved the PLAG1 insert into two distinct fragments: PLAG1 restriction fragment 1 (1200 bp, middle band) and PLAG1 restriction fragment 2 (250 bp, bottom band). Lane Marker represents a DNA ladder for size reference, while Lanes 1–6 correspond to six independent bacterial colonies screened for successful cloning. The consistent banding pattern across all six samples confirms the successful integration of the PLAG1 insert into the pcDNA3.1(+)-IRES-GFP vector and verifies its correct restriction digestion pattern.

3.3.2 Transfection of Human Primary Salivary Gland Cells

Initial transfection attempts were conducted using primary human (parotid, submandibular and sublingual) salivary gland cells to investigate the role of PLAG1 in PA pathogenesis. However, multiple transfection efforts across different cell batches and donors yielded inconsistent results. Fluorescence microscopy and Western blot analysis failed to detect the expression of the transfected gene or its protein product, indicating unsuccessful transfection (Figure 3.3). Notably, primary parotid gland cells from a specific donor displayed successful transfection (Figure 3.4) and protein expression, but this result was not reproducible across other batches. Further analysis using STR identity profiling revealed contamination with HEK293 cells, questioning the validity of the observed transfection success.

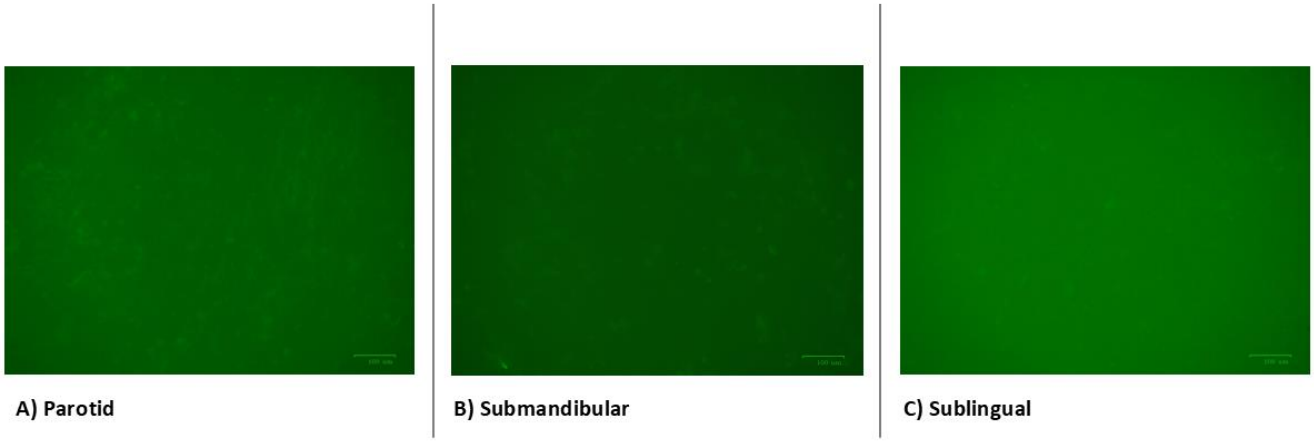


Figure 3.3. Transfection of primary salivary gland cells.
Representative fluorescent microscopy images of transfected primary salivary gland cells: parotid (A), submandibular (B), and sublingual (C). Lack of GFP fluorescence indicates unsuccessful transfection in all three cell types.

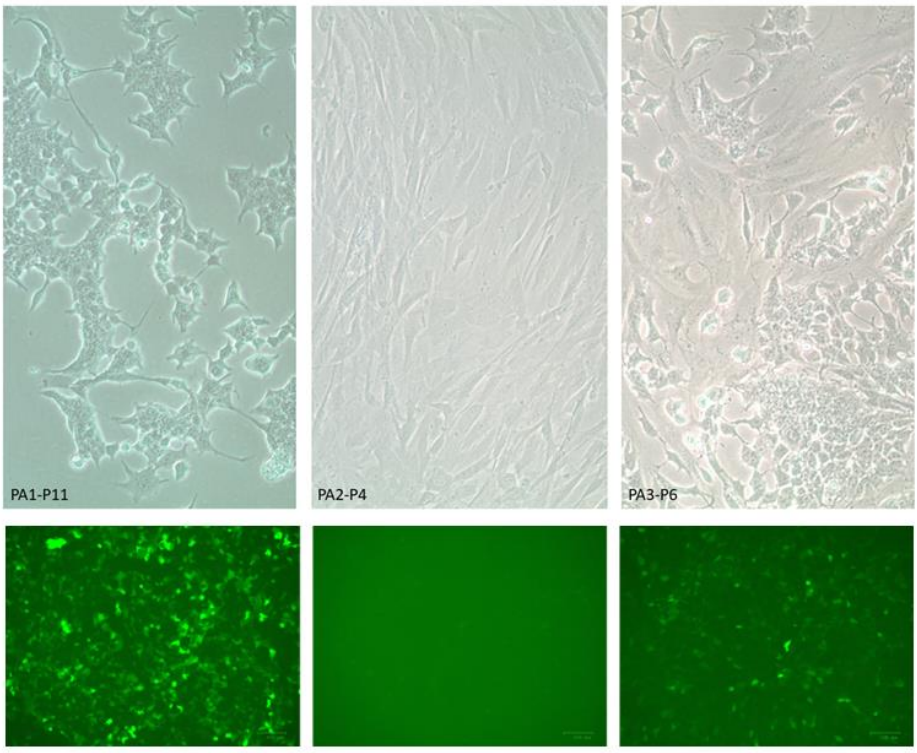


Figure 3.4. Transfection of primary parotid gland cells from different batches.
Representative bright-field microscopy images of three distinct primary parotid cell batches, highlighting phenotypic variability (top). Corresponding GFP fluorescence images demonstrating variable transfection efficiency among the cell populations (bottom).

3.3.3 Transfection of BMI-1 Transduced Parotid and Sublingual Cells

To circumvent the challenges associated with primary salivary gland cells, BMI-1 transduced foetal parotid and sublingual cells, which exhibit extended proliferative capacity while retaining stem cell-like properties, were utilised for transfection. GFP expression confirmed successful transfection (Figure 3.5); however, Western blot analysis repeatedly failed to detect the exogenous protein. Further investigation identified cross-contamination with a GFP-tagged construct during the viral transduction process as the likely source of the observed fluorescence. Given these challenges, HEK293 cells, known for their high transfection efficiency, were subsequently employed for further experimental analyses.

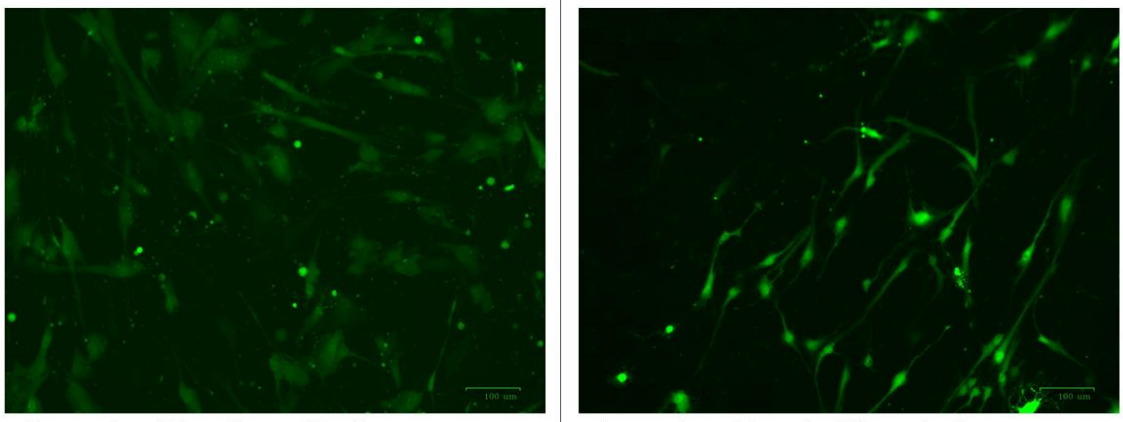
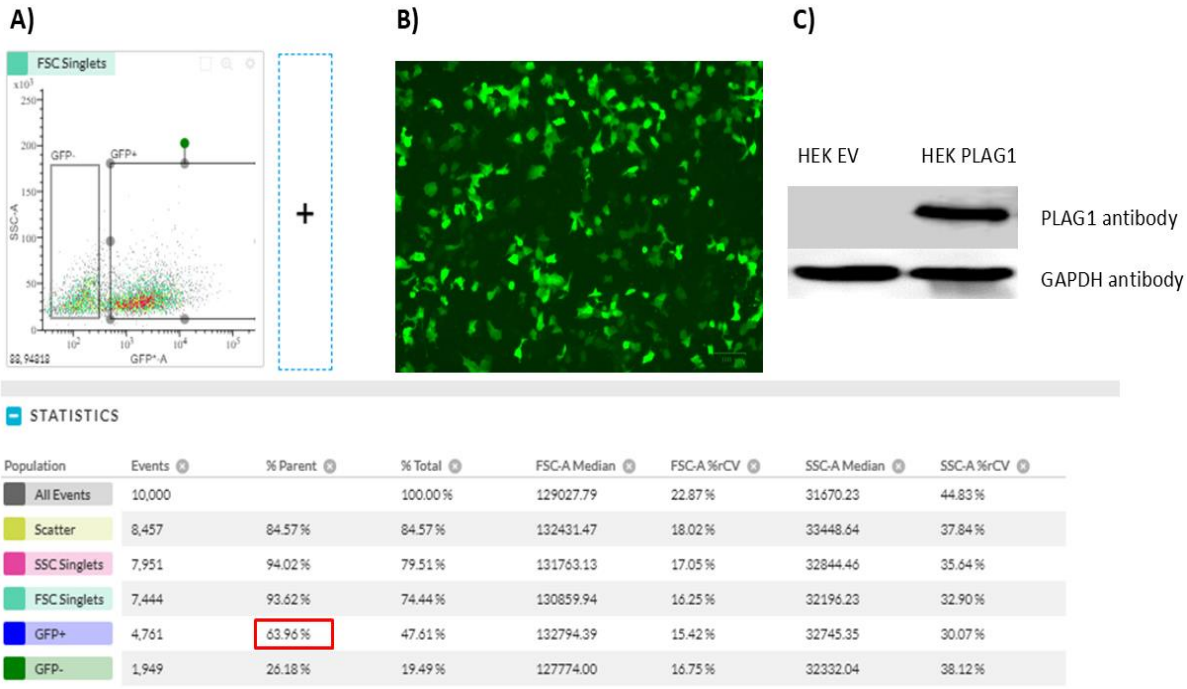


Figure 3.5. Transfection of BMI-transduced parotid and sublingual cells.
Representative fluorescence microscopy images demonstrating successful GFP expression in transduced parotid (A) and sublingual (B) gland cells.

3.3.4 Transfection of HEK293 cells

To investigate the downstream effects of PLAG1 overexpression, RNA extraction was performed on transfected HEK293 cells. These experiments were undertaken on multiple occasions (with transfection monitored by analysis of GFP) and RNA samples from three different experiments were used for our analysis. The transfection efficiency in HEK293 cells was evaluated using flow cytometry, which showed an efficiency rate of 63.96% (Figure 3.6A). GFP expression was further verified through fluorescence microscopy (Figure 3.6B) and Western blot analysis (Figure 3.6C), confirming successful protein expression.



Population	Events	% Parent	% Total	FSC-A Median	FSC-A %rCV	SSC-A Median	SSC-A %rCV
All Events	10,000		100.00 %	129027.79	22.87 %	31670.23	44.83 %
Scatter	8,457	84.57 %	84.57 %	132431.47	18.02 %	33448.64	37.84 %
SSC Singlets	7,951	94.02 %	79.51 %	131763.13	17.05 %	32844.46	35.64 %
FSC Singlets	7,444	93.62 %	74.44 %	130859.94	16.25 %	32196.23	32.90 %
GFP+	4,761	63.96 %	47.61 %	132794.39	15.42 %	32745.35	30.07 %
GFP-	1,949	26.18 %	19.49 %	127774.00	16.75 %	32332.04	38.12 %

Figure 3.6 Transfection of HEK293 cells.
Assessment of transfection efficiency in HEK293 cells using flow cytometry (A), demonstrating a transfection efficiency of 63.96%. Positive GFP expression was confirmed via fluorescence microscopy (B) and Western blot analysis (C), indicating successful protein expression. The data is representative of N = 3 replicate experiments.

3.3.5 RNA quantification

To ensure the quality and quantity of RNA for downstream analysis, RNA was quantified using a Nanodrop 1000 spectrophotometer (Thermo Scientific). The absorbance was measured at 260 nm and 280 nm to determine RNA concentration and purity (Figure 3.7). The RNA concentration was found to be between 900-1212 µg/ml, with an A260/A280 ratio of 2.0, indicating sample purity (details in Table 3.1).

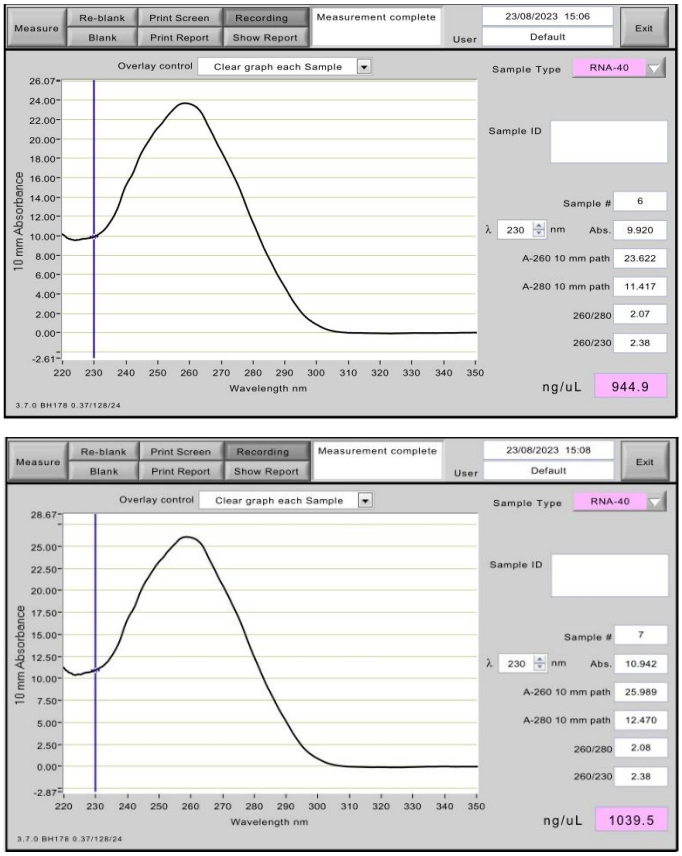


Figure 3.7 NanoDrop measurement of RNA samples.
Quantitative RNA analysis of HEK293 cells transfected with either PLAG1 or empty vector control. RNA concentration (ng/µL) and purity (A260/A280 and A260/A230 ratios) were determined using a NanoDrop spectrophotometer.

Table 3.1 RNA sample quantification by Nanodrop spectrometer

No.	Sample	RNA concentration ng/uL
1	PLAG IRES	900
2	EMPTY IRES	1212
3	PLAG IRES	1123
4	EMPTY IRES	944.9
5	PLAG IRES	990.8
6	EMPTY IRES	1039

3.3.6 Data validation and gene clustering analysis

Principal component analysis (PCA) was performed on the normalised count data to evaluate the degree of similarity between the three biological repeats of the two sets of samples (PLAG1-expressing cells and mock-transfected cells). The ClustVis online tool (Metsalu and Vilo, 2015) was

used for this analysis. The samples were plotted on a 2D plane using their first two principal components (Figure 3.8). This type of plot helps visualise the overall impact of experimental variables and batch effects. The x-axis represents the direction with the most variance, while the y-axis represents the second most. The data exhibited distinct clustering between control and PLAG1-expressing samples, with a slight difference between biological replicates.

The relationship between all samples was further analysed using clustering analysis based on normalised transcript quantification. The heatmap in Figure 3.9 illustrates the distribution and grouping of 1148 genes into two distinct molecular clusters, covering all six samples. Cluster 1 is represented in orange and includes three PLAG1-expressing samples. Cluster 2, shown in pink, consists of the control samples.

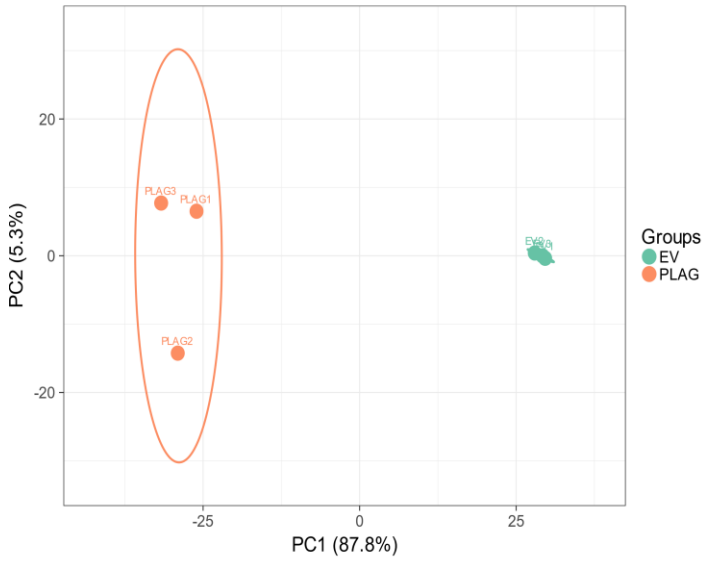


Figure 3.8 PCA plot illustrating sample distribution based on PLAG1 expression.
The PCA plot illustrates the variance in gene expression profiles between PLAG1-overexpressing samples (PLAG1: PLAG1, PLAG2, PLAG3, orange) and empty vector control samples (EV: EV1, EV2, EV3, teal). Principal Component 1 (PC1) accounts for 87.8% of the variance, effectively separating PLAG1-overexpressing samples from control (EV) samples, indicating a strong global transcriptional shift due to PLAG1 overexpression. Principal Component 2 (PC2) explains 5.3% of the variance, capturing within-group variability, particularly among PLAG1-overexpressing replicates. The ellipses represent 95% confidence

intervals, highlighting the distinct clustering of the two experimental groups. The clear separation along PC1 suggests a substantial impact of PLAG1 overexpression on gene expression patterns.

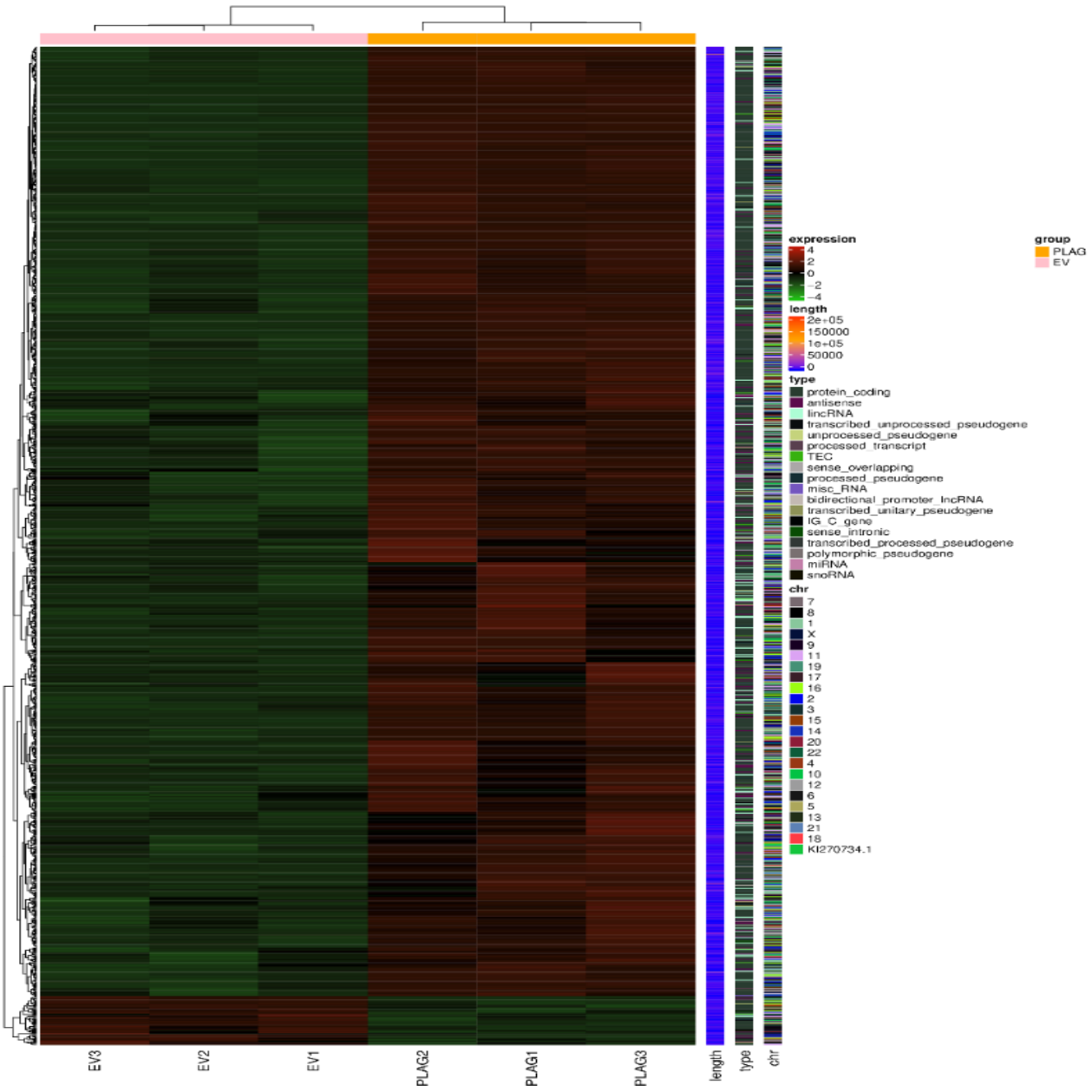


Figure 3.9 Heatmap illustrating unsupervised hierarchical clustering of samples based on gene expression patterns.

Two distinct clusters emerge: Cluster 1, predominantly composed of PLAG1-overexpressing samples, and Cluster 2, comprising control samples. Gene expression levels are colour-coded, with red indicating upregulation and green indicating downregulation, relative to the average expression across all samples.

3.3.7 Differentially expressed genes in response to PLAG1 overexpression

A comparison was conducted between two distinct clusters: samples expressing PLAG1 versus the control samples, as part of the unsupervised analysis. Genes with an FDR < 0.05 and an absolute log2 fold change > 1 were identified as differentially expressed genes for each comparison.

The results showed that 1092 genes were significantly upregulated, while 56 were downregulated. A volcano plot was created to visualise the differentially expressed genes based on their statistical significance (p-value) versus the magnitude of change (fold change) for each gene (see Figure 3.11). Table 3.2 presents the top 20 most significantly upregulated and top 10 downregulated genes. Additionally, to visualise the relationship between the most differentially expressed genes within the two molecular clusters, a heatmap clustering of the top 50 genes was generated, as depicted in Figure 3.12.

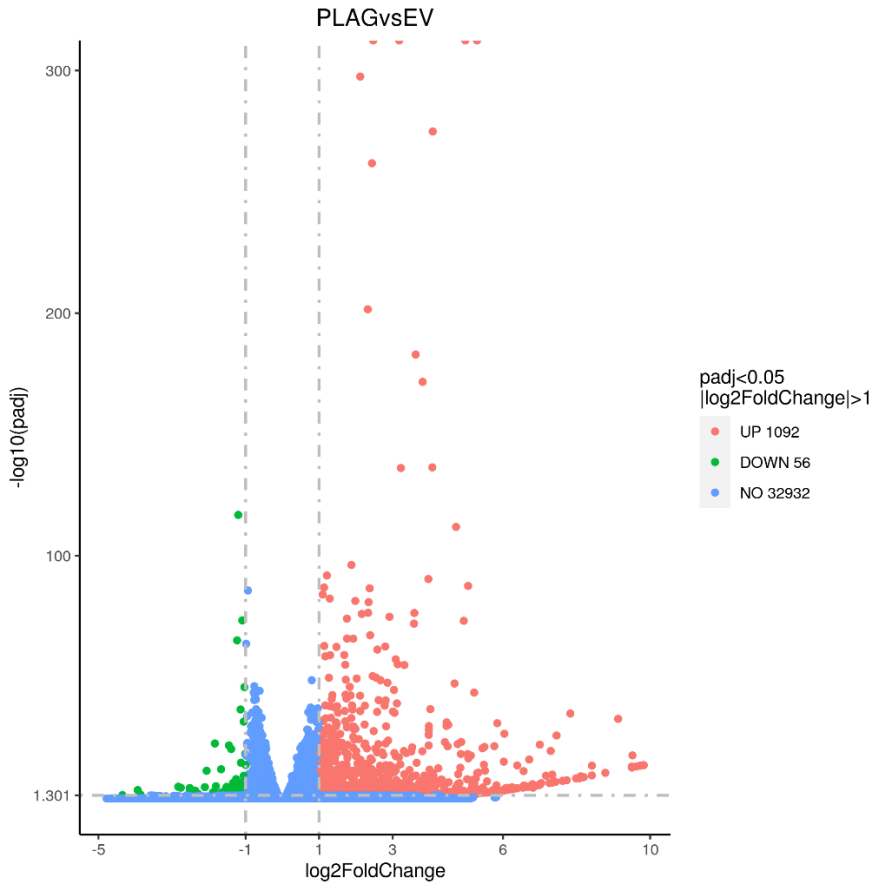


Figure 3.10 Volcano plot displaying fold changes (x-axis) against adjusted p-values (y-axis).

Each data point in the scatter plot represents a gene. Genes with an adjusted p-value of less than 0.05 and a log2 fold change greater than 1 are represented by pink dots, indicating up-regulated genes. Genes with an adjusted p-value of less than 0.05 and a log2 fold change less than -1 indicated by green dots, representing down-regulated genes.

Table 3.2 Most significantly upregulated and downregulated genes.

Gene Name	log2 Fold Change	p-value	FDR	Regulation
PLAG1	5.292783855	0	0	upregulated
ARC	4.975428111	0	0	upregulated
CRABP2	3.176596654	0	0	upregulated
FLNC	2.472017119	0	0	upregulated
PLEC	2.118476879	6.95E-302	3.18E-298	upregulated
SRPK3	4.089506316	3.47E-279	1.32E-275	upregulated
NOTCH1	2.436453725	5.06E-266	1.65E-262	upregulated
AHNAK	2.324579845	8.93E-206	2.55E-202	upregulated
CHD5	3.625099699	4.42E-187	1.12E-183	upregulated
MAFA	3.814784847	9.08E-176	2.07E-172	upregulated
CRLF1	4.078199526	1.67E-140	3.46E-137	upregulated
CDKN1C	3.222811847	3.37E-140	6.42E-137	upregulated
UNC13D	4.723296735	7.27E-116	1.19E-112	upregulated
EPPK1	1.878448427	3.86E-100	5.88E-97	upregulated
MIDN	1.210698396	8.79E-96	1.25E-92	upregulated
PRR35	3.970746305	2.75E-94	3.70E-91	upregulated
FER1L5	5.050608543	2.02E-91	2.57E-88	upregulated
GATA2	1.135297278	8.67E-91	1.04E-87	upregulated
CCDC88B	2.37465585	2.15E-90	2.46E-87	upregulated
FURIN	1.096719238	8.35E-88	8.67E-85	upregulated
EXO1	-1.19909516	7.83E-121	1.38E-117	downregulated
H1F0	-1.087762507	5.40E-77	3.98E-74	downregulated
SNHG12	-1.229978697	1.10E-68	6.80E-66	downregulated
CNOT10	-1.032081899	3.41E-49	1.28E-46	downregulated
C6orf48	-1.136112687	8.46E-40	2.10E-37	downregulated
KIF24	-1.051145732	1.09E-34	2.10E-32	downregulated
NR1D1	-1.83332953	2.04E-25	2.11E-23	downregulated
AC093512.2	-1.451270338	1.41E-24	1.35E-22	downregulated
DHRS2	-1.390816671	4.84E-23	4.11E-21	downregulated
RANBP17	-1.010653072	5.56E-21	4.15E-19	downregulated

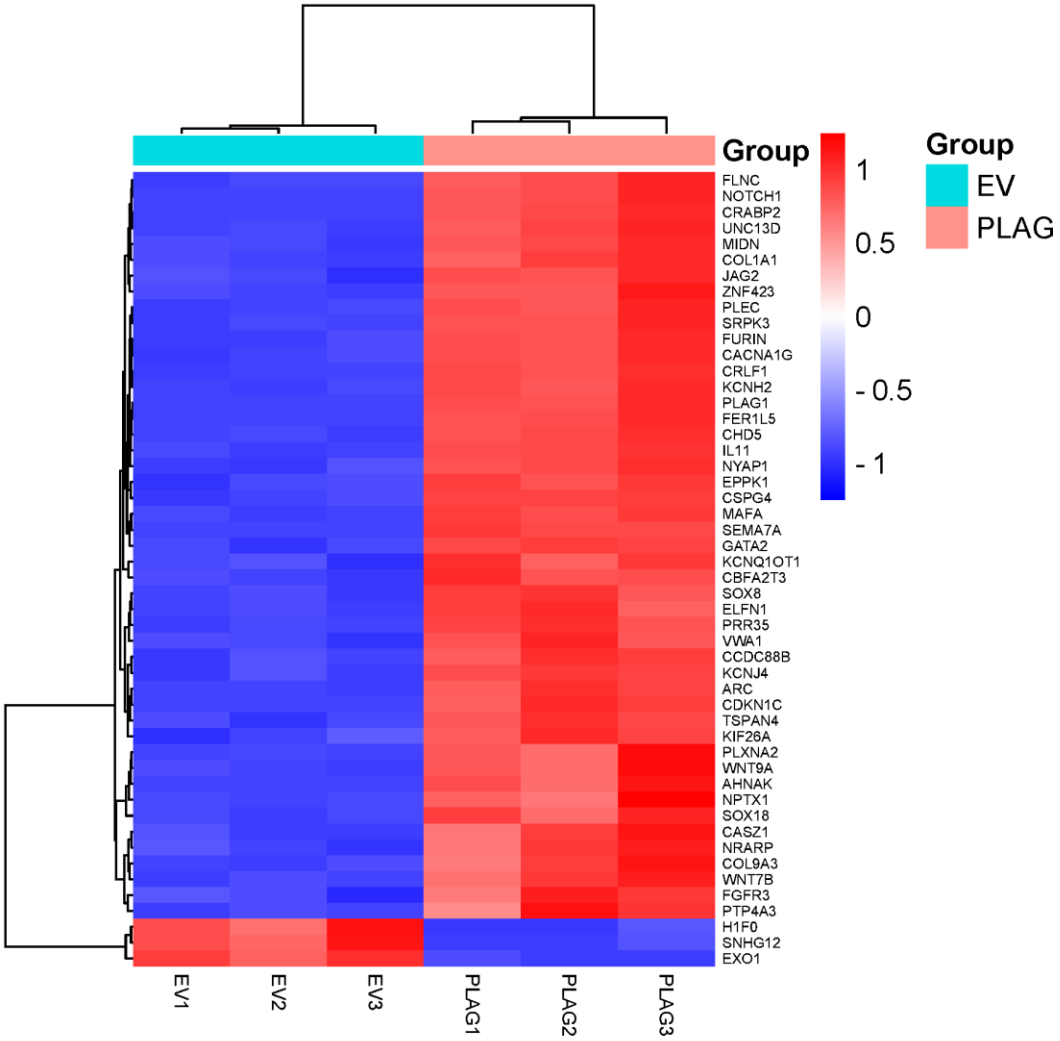


Figure 3.11 Heatmap of the top 50 differentially expressed genes.
Heatmap illustrating the expression patterns of the top 50 differentially expressed genes (DEGs) between PLAG1-overexpressing and control samples. Red and blue colour gradients represent up- and downregulation, respectively.

3.3.8 Functional enrichment analysis

Functional enrichment analysis was performed to identify key biological processes and pathways among the differentially expressed genes, simplifying data interpretation, generating hypotheses and highlighting the potential therapeutic targets. By categorising genes into functional groups, it reveals critical pathways involved in altered functions such as cell proliferation, survival and differentiation. It also facilitates comparative analysis across studies.

GSEA software was used to conduct gene set enrichment analysis by comparing our DEGs to predefined gene sets. These gene sets represent a collection of genes that share common biological functions, locations or regulation. The analysis identifies whether the genes in these gene sets are overrepresented in our dataset. In our study, GSEA was performed using C2 Curated gene sets, C5 Ontology gene sets, and C6 Oncogenic signature gene sets.

3.3.8.1 C5 subcollection GO: Gene Ontology

The C5 subcollection includes gene sets from Gene Ontology (GO) annotations: Biological Processes (BP), Molecular Function (MF) and cellular components (CC). This help identify significant biological roles, activities and locations of genes providing insights into the biological relevance of gene expression data.

The 1148 differential expressed genes were used for comprehensive enrichment analyses based on the GO terms. Our findings revealed significant enrichment of terms related to the extracellular matrix (ECM) across all three components. This encompassed external encapsulating structure, collagen-containing extracellular matrix, and extracellular matrix binding. Additionally, there was an enrichment of terms related to growth factor binding and activity in GO molecular function terms. This enrichment suggests that the ECM and growth factor related genes play a significant role in the biology of pleomorphic adenoma and may indicate the PA development and progression are heavily influenced by the interaction of pathways related to these genes.

Table 3.3 presents various Gene Ontology (GO) terms that are significantly enriched in our dataset. Each term is associated with a specific biological process, cellular component, or molecular function, along with the size of the gene set, p-value, and false discovery rate (FDR). These terms provide valuable insights into the biological significance of the genes involved,

particularly within the context of oncogenesis. Figure 3.12 represents the enrichment plots of key GO terms.

Table 3.3 GO enrichment in phenotype PLAG1.
Most significantly enriched terms (CC, BP and MF) with a P-value less than 0.05 and FDR less than 0.05.

Name	Size	p-value	FDR	Description (Kanehisa et al., 2023)
GOCC_COLLAGEN_TRIMER	74	0.000	0.001	Refers to a trimeric form of collagen, a structural protein in the extracellular matrix. Collagen alterations can influence tumour progression and metastasis by modifying the tumour microenvironment.
GOCC_EXTERNAL_ENCAPSULATING_STRUCTURE	482	0.000	0.000	Encompasses structures like the extracellular matrix that encapsulate cells. These structures are crucial in maintaining tissue integrity and can be disrupted in cancer, aiding in tumour invasion and metastasis.
GOCC_COLLAGEN_CONTAINING_EXTRACELLULAR_MATRIX	371	0.000	0.001	Indicates the presence of collagen within the extracellular matrix. Changes in this matrix can affect cell behaviour and contribute to cancer progression.
GOBP_EXTERNAL_ENCAPSULATING_STRUCTURE_ORGANIZATION	294	0.000	0.010	Involves the organisation of structures like the extracellular matrix. Proper organisation is essential for tissue homeostasis, and its disruption can facilitate cancer cell invasion.
GOBP_CELLULAR_RESPONSE_TO_RETINOIC_ACID	64	0.000	0.007	Refers to cellular responses to retinoic acid, a derivative of vitamin A. Retinoic acid can influence cell differentiation and proliferation, and its dysregulation is implicated in cancer.
GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT	144	0.000	0.000	Refers to proteins that are part of the extracellular matrix. These proteins provide structural support and can influence cell signalling pathways involved in cancer.
GOMF_EXTRACELLULAR_MATRIX_STRUCTURAL_CONSTITUENT_CONFERRING_ TENSILE_STRENGTH	43	0.000	0.000	Specific to extracellular matrix proteins that provide tensile strength. These proteins are crucial for tissue integrity and their alteration can promote cancer cell invasion.

GOMF_EXTRACELLULAR_MATRIX_BINDING	49	0.000	0.003	Involves proteins that bind to the extracellular matrix. These interactions are important for cell adhesion and migration, processes that are often dysregulated in cancer.
GOMF_GROWTH_FACTOR_BINDING	126	0.000	0.003	Refers to proteins that bind growth factors. Growth factors are key regulators of cell proliferation and survival, and their dysregulation is a hallmark of cancer.
GOMF_GROWTH_FACTOR_ACTIVITY	136	0.000	0.003	Involves proteins with growth factor activity. These proteins can stimulate cell growth and division, and their overexpression or mutation can lead to uncontrolled cell proliferation in cancer.
GOMF_NOTCH_BINDING	24	0.001	0.004	Refers to proteins that bind to Notch receptors. Notch signalling is involved in cell differentiation and proliferation, and its dysregulation is linked to various cancers.

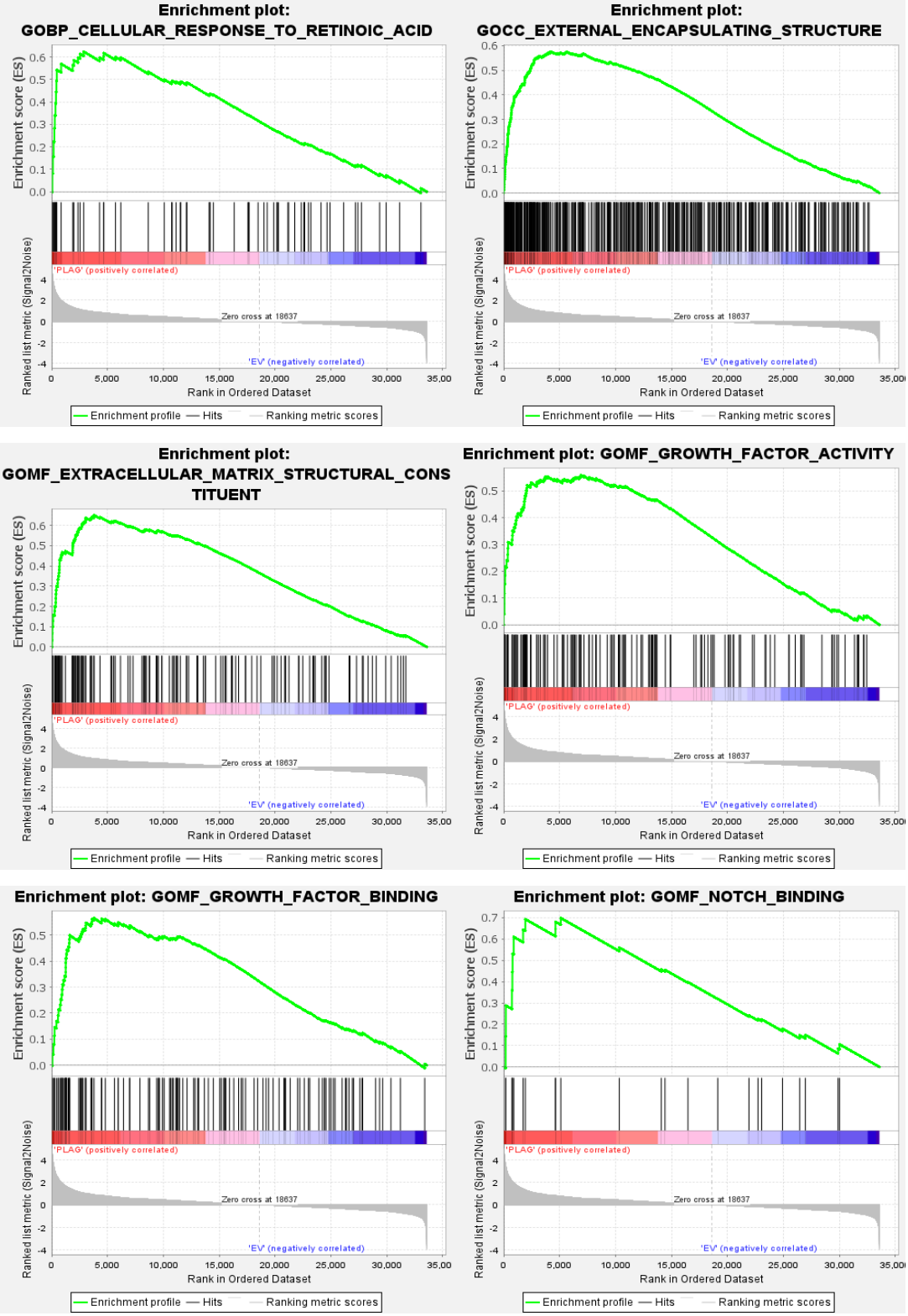


Figure 3.12 Representative GO gene set enrichment plots.

The enrichment plot displays the Enrichment Score (ES) as the peak of the green curve, which indicates the degree of over- or under-expression of the gene set. The middle section shows the ranked list metric, with red and blue bars indicating the positions of the genes in the gene set. The green line represents the running enrichment score, and the vertical lines at the bottom indicate the locations of the genes in the ranked list.

3.3.8.2 C2 subcollection CP: Canonical pathways (KEGG Legacy Sets)

To analyse pathways and identify significantly enriched pathways in the dataset, GSEA was conducted using the C2 KEGG Legacy Sets. This method utilises curated gene sets, particularly canonical pathways, from the KEGG pathway database (Kyoto Encyclopaedia of Genes and Genomes), offering insights into the most relevant biological processes and pathways related to the data. Our data analysis revealed that 117 out of the 172 gene sets in the C2 Kegg Legacy Sets database showed enrichment in the PLAG1 phenotype. Thirty-five gene sets were significantly enriched at a false discovery rate (FDR) of less than 25% and thirty-three at a p-value of less than 5%.

Table 3.4 summarises the analysis findings and provides a brief explanation of the importance of the enriched pathways. The observed significant enrichment pertained to the activation of the Basal Cell Carcinoma and Hedgehog signalling pathways (Figure 3.13). These pathways demonstrate interconnections with the WNT signalling pathway, which plays a crucial role in PA development. This is supported by the common expression of different members of the WNT gene family, such as WNT7B, WNT9A, WNT6, and WNT11, by these interacting pathways.

Table 3.4 Enriched pathways in PLAG1 samples using KEGG Legacy Gene Sets.

Name	Description (Kaneshia et al.,2023)
KEGG_BASAL_CELL_CARCINOMA	This pathway, consisting of 54 genes, shows significant enrichment with a p-value and FDR of 0. It is crucial in the development of basal cell carcinoma, a common type of skin cancer.
KEGG_MELANOGENESIS	Comprising 95 genes, this pathway is significantly enriched with both p-value and FDR at 0. Melanogenesis is involved in the production of melanin, and its dysregulation can contribute to melanoma.
KEGG_HEDGEHOG_SIGNALING_PATHWAY	This pathway, with 53 genes, shows significant enrichment with a p-value and FDR of 0. The

	Hedgehog signalling pathway is essential for cell differentiation and proliferation, and its aberrant activation is linked to various cancers.
KEGG_CELL_ADHESION_MOLECULES_CAMS	Encompassing 116 genes, this pathway is significantly enriched with a p-value of 0 and an FDR of 0.002. Cell adhesion molecules are critical for cell-cell and cell-matrix interactions, and their dysregulation can lead to tumour metastasis.
KEGG_TASTE_TRANSDUCTION	This pathway, consisting of 38 genes, shows significant enrichment with a p-value of 0 and an FDR of 0.009. While primarily involved in taste perception, alterations in this pathway can have implications in cancer.
KEGG_NEUROACTIVE_LIGAND_RECEPTOR_INTERACTION	With 213 genes, this pathway is significantly enriched with a p-value of 0 and an FDR of 0.008. This pathway involves neurotransmitter receptors and their ligands, which can influence cancer progression through neuroendocrine signalling.
KEGG_HEMATOPOIETIC_CELL_LINEAGE	This pathway, comprising 72 genes, shows significant enrichment with a p-value of 0 and an FDR of 0.011. It is crucial for the development of blood cells, and its dysregulation can lead to hematologic malignancies.
KEGG_MATURITY_ONSET_DIABETES_OF_THE_YOUNG	This pathway encompasses 24 genes and is significantly enriched, with a p-value of 0.001 and an FDR of 0.017. While primarily related to diabetes, alterations in this pathway can also impact cancer metabolism.
KEGG_CALCIUM_SIGNALING_PATHWAY	This pathway, with 162 genes, shows significant enrichment with a p-value of 0 and an FDR of 0.026. Calcium signalling is vital for various cellular processes, and its dysregulation can contribute to cancer.
KEGG_TYPE_II_DIABETES_MELLITUS	Comprising 45 genes, this pathway is significantly enriched with a p-value of 0.009 and an FDR of 0.032. Type II diabetes is linked to cancer through metabolic and inflammatory pathways.
KEGG_NOTCH_SIGNALING_PATHWAY	This pathway, consisting of 47 genes, shows significant enrichment with a p-value of 0.002 and an FDR of 0.038. The Notch signalling pathway is involved in cell differentiation and proliferation, and its dysregulation is associated with various cancers.

KEGG_COMPLEMENT_AND_COAGULATION_CASCADES	This pathway encompasses 60 genes and is significantly enriched, with a p-value of 0.004 and an FDR of 0.037. These cascades play a role in immune response and inflammation, which can influence cancer progression.
KEGG_AXON_GUIDANCE	This pathway, with 126 genes, shows significant enrichment with a p-value of 0 and an FDR of 0.038. Axon guidance is crucial for neural development, and its dysregulation can contribute to cancer metastasis.
KEGG_DILATED_CARDIOMYOPATHY	This pathway comprises 82 genes and is significantly enriched, with a p-value of 0 and an FDR of 0.042. While primarily related to heart disease, alterations in this pathway can also impact cancer.
KEGG_ECM_RECEPTOR_INTERACTION	This pathway, consisting of 79 genes, shows significant enrichment with a p-value of 0 and an FDR of 0.041. ECM (extracellular matrix) receptor interactions are critical for cell adhesion and migration, and their dysregulation can lead to tumour invasion and metastasis.

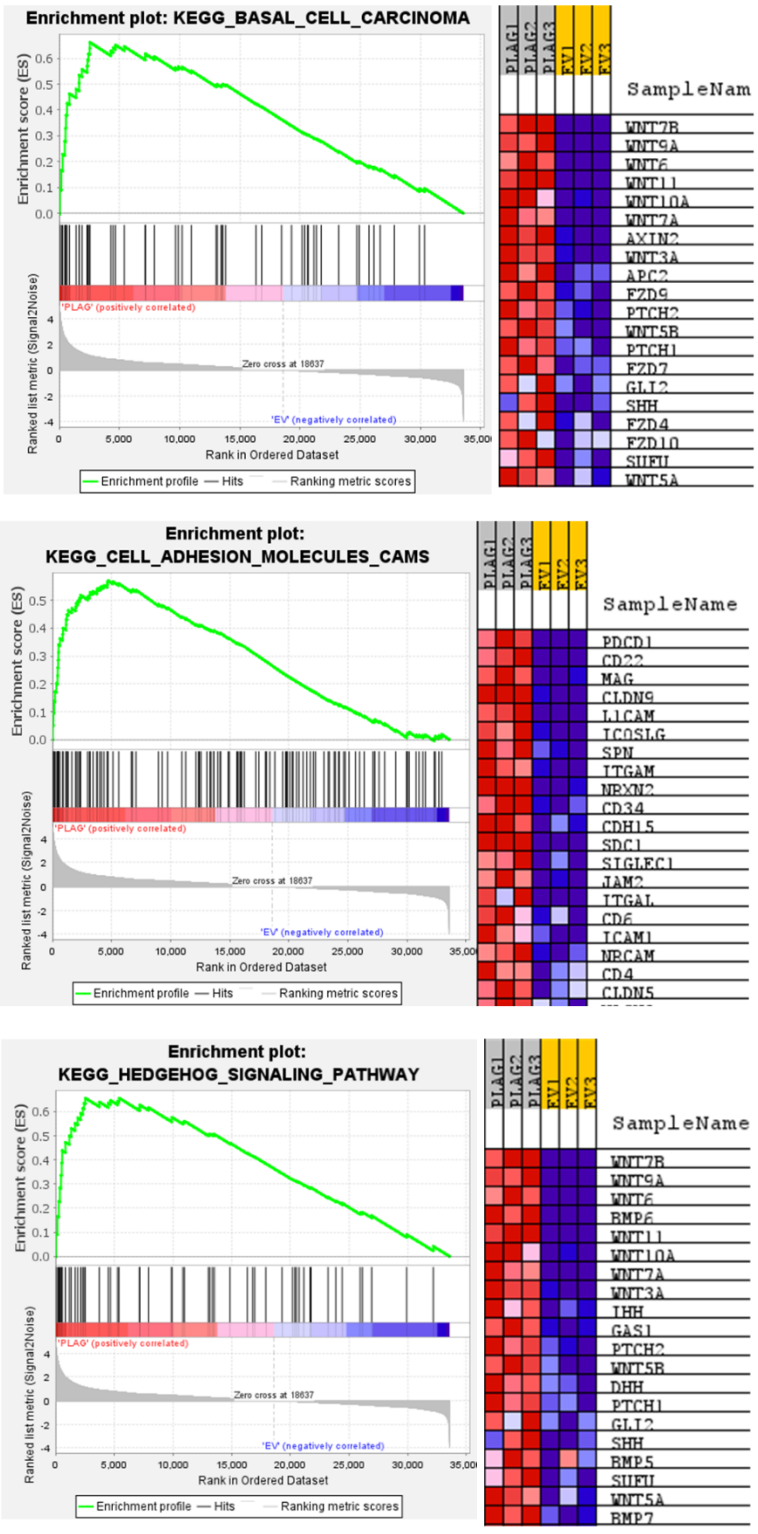


Figure 3.13 Enrichment plots and heatmaps of key pathways.

The figures present the enrichment plots of the most significantly enriched pathways, including the Basal Cell Carcinoma Pathway, Hedgehog Signalling Pathway, and Cell Adhesion Molecules. Each plot is accompanied by a heatmap representing the top 20 genes involved in these pathways.

3.3.8.3 C6 Collection: Oncogenic Signature Gene Sets

The "oncogenic signature gene set" represents pathways frequently dysregulated in cancer. In this category, our findings revealed that 165 out of 187 gene sets exhibited upregulation in the PLAG phenotype. Among these, 112 gene sets demonstrated significance at FDR < 25%, indicating that gene sets are genuinely involved in the development of pleomorphic adenoma and play a critical role in the tumour phenotype, with a relatively low probability of false positives. Additionally, 62 gene sets displayed significant enrichment at a nominal p-value < 1% and 82 gene sets were significantly enriched at a nominal p-value < 5%.

Table 3.5 provides an overview of the top 15 gene sets enriched in the PLAG1-expressing samples, highlighting their significance in tumour development. Key pathways identified in this analysis include those associated with stemness, DNA damage repair, and oncogenic signalling. The ESC_V6.5_UP_EARLY.V1_DN gene set is downregulated in PLAG1-expressing cells, indicating a suppression of genes associated with pluripotency and stem-like characteristics. This suggests a shift towards differentiation, aligning with the benign and controlled growth pattern of PA, in contrast to malignant tumours that maintain a high degree of stemness for invasive potential. LEF1_UP.V1_UP and WNT_UP.V1_UP gene sets were significantly enriched, indicating activation of Wnt/β-catenin signalling, a pathway critical for cell proliferation, differentiation, and tumour maintenance. While this activation may support the structured architecture of PA, excessive Wnt signalling has been associated with enhanced tumour plasticity and a potential for recurrence. Additionally, the CTIP_DN.V1_UP gene set, linked to DNA damage repair suppression, was upregulated, raising concerns about genomic stability. While reduced DNA repair activity may limit the adaptability of the tumour, prolonged genomic instability could predispose PA cells to malignant transformation.

KRAS signalling remained subdued, as indicated by the enrichment of the KRAS.600_UP.V1_DN gene set. This suggests that KRAS-driven proliferative signals are not dominant in PLAG1-

expressing tumours, aligning with their slow-growing and well-differentiated nature. Enrichment plots derived from this analysis are shown in Figure 3.14.

Table 3.5 Enriched sets in PLAG1 cluster using Oncogenic Signature Gene Set.

Name	Size	p-value	FDR	Description (Liberzon et al.,2011)
ESC_V6.5_UP_EARLY.V1_DN	148	0	0	This gene set is associated with early embryonic stem cell differentiation. Dysregulation in these genes can lead to uncontrolled cell growth and differentiation, contributing to tumorigenesis.
LEF1_UP.V1_UP	187	0	0	LEF1 is a transcription factor involved in the Wnt signalling pathway, which is crucial for cell proliferation and differentiation. Overexpression of LEF1 can lead to increased cell proliferation and cancer development.
CTIP_DN.V1_UP	108	0	0	CTIP is involved in DNA damage response and repair. Downregulation of CTIP can impair DNA repair mechanisms, leading to genomic instability and cancer.
KRAS.600_UP.V1_DN	223	0	0	KRAS is a key oncogene in the RAS/MAPK pathway, which regulates cell growth and survival. Mutations in KRAS are common in various cancers, leading to uncontrolled cell proliferation.
KRAS.KIDNEY_UP.V1_DN	107	0	0	This gene set includes genes upregulated in kidney cells with KRAS mutations. KRAS mutations in kidney cells can lead to renal cell carcinoma.
KRAS.LUNG_UP.V1_DN	112	0	0	This set includes genes upregulated in lung cells with KRAS mutations. KRAS mutations are common in non-small cell lung cancer.
KRAS.PROSTATE_UP.V1_DN	111	0	0	This gene set includes genes upregulated in prostate cells with KRAS mutations. KRAS mutations can contribute to prostate cancer development.
KRAS.LUNG.BREAST_UP.V1_DN	115	0	0.001	This set includes genes upregulated in both lung and breast cells with KRAS mutations. KRAS mutations can drive tumorigenesis in both lung and breast cancers.
BMI1_DN.V1_UP	135	0.001	0.001	BMI1 is part of the Polycomb group of proteins involved in maintaining stem cell properties. Downregulation of BMI1 can lead to loss of cell cycle control and cancer.

KRAS.300_UP.V1_DN	113	0	0.001	This gene set includes genes upregulated in cells with a specific KRAS mutation. KRAS mutations are significant drivers of cancer.
ATM_DN.V1_DN	118	0	0.001	ATM is involved in the DNA damage response. Downregulation of ATM can lead to impaired DNA repair and increased cancer risk.
KRAS.50_UP.V1_DN	37	0	0.001	his set includes genes upregulated in cells with a specific KRAS mutation. KRAS mutations are critical in cancer development.
JAK2_DN.V1_UP	143	0	0.001	JAK2 is involved in the JAK/STAT signalling pathway, which regulates cell growth and immune function. Dysregulation of JAK2 can lead to hematologic malignancies.
MEL18_DN.V1_UP	133	0	0.001	MEL18 is a Polycomb group protein involved in gene silencing. Downregulation of MEL18 can lead to loss of cell cycle control and cancer.
WNT_UP.V1_UP	158	0	0.001	The Wnt signalling pathway is crucial for cell proliferation and differentiation. Overactivation of Wnt signalling is common in many cancers.
RELA_DN.V1_UP	133	0	0.001	RELA is a subunit of NF-κB, a transcription factor involved in immune response and cell survival. Dysregulation of NF-κB signalling can contribute to cancer.
CAHOY_ASTROGLIAL	87	0	0.001	This gene set includes genes specific to astroglial cells. Dysregulation in these genes can contribute to gliomas and other brain cancers.
IL15_UP.V1_UP	169	0	0.001	IL-15 is a cytokine involved in immune response. Overexpression of IL-15 can lead to chronic inflammation and cancer.
PGF_UP.V1_DN	168	0	0.001	PGF (Placental Growth Factor) is involved in angiogenesis. Overexpression of PGF can promote tumour growth by enhancing blood supply.
TGFB_UP.V1_UP	175	0	0.001	TGFB is involved in cell growth and differentiation. Dysregulation of TGFB signalling can lead to cancer progression and metastasis.

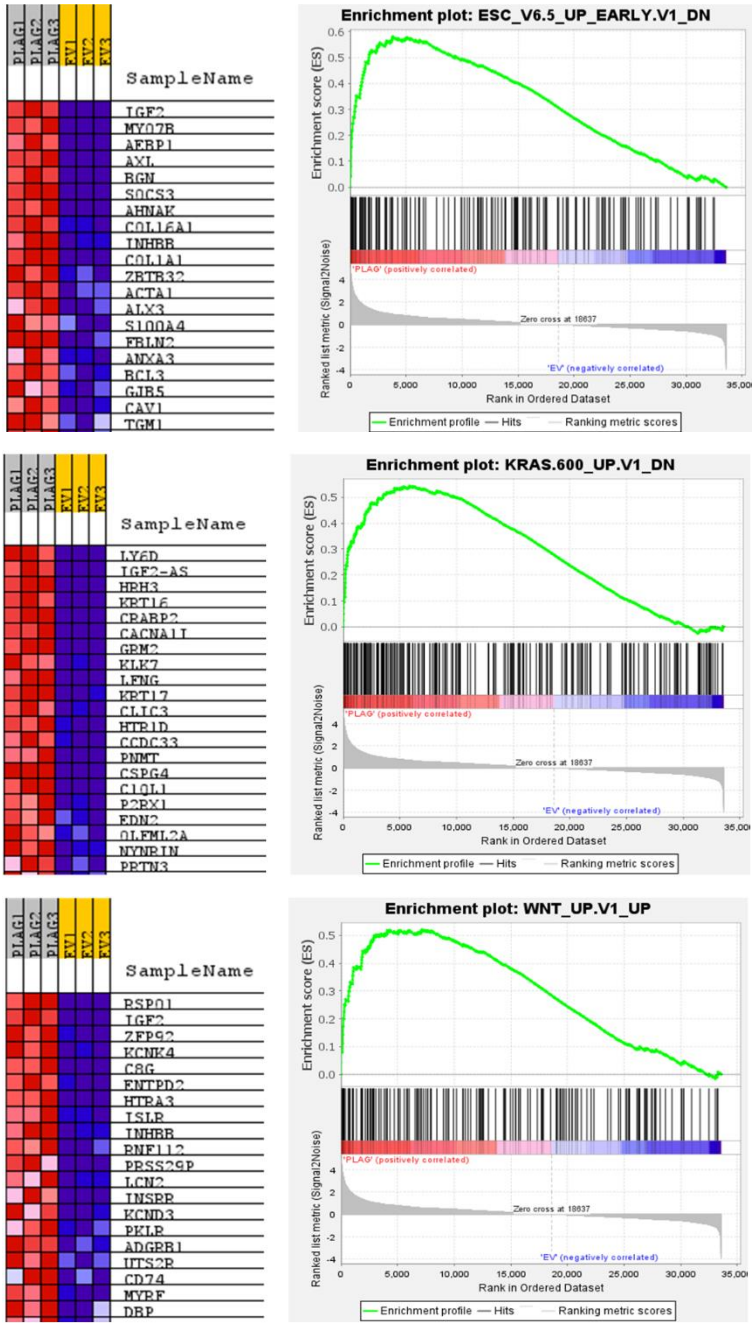


Figure 3.14 Enriched sets in PLAG1 cluster using oncogenic signature gene set.
Representative enrichment plots of key enriched gene sets with heatmaps of the most significantly enriched genes.

3.3.9 Validation of upregulated genes by qPCR

To validate the most significantly upregulated genes in response to PLAG1 overexpression, qPCR was performed using TaqMan probes (ThermoFisher Scientific). The analysis targeted the three

top statistically upregulated genes—Filamin C (FLNC), Cellular Retinoic Acid-Binding Protein 2 (CRABP2), and Activity-Regulated Cytoskeleton-Associated Protein (ARC)—along with Pleomorphic Adenoma Gene 1 (PLAG1), selected based on their statistical significance in differential expression analysis. cDNA samples derived from the three biological replicates used for RNA sequencing were subjected to qPCR analysis. Results demonstrated a significant upregulation of all four target genes in PLAG1-expressing samples compared to the control group, consistent with the RNA sequencing data (Figure 3.15).

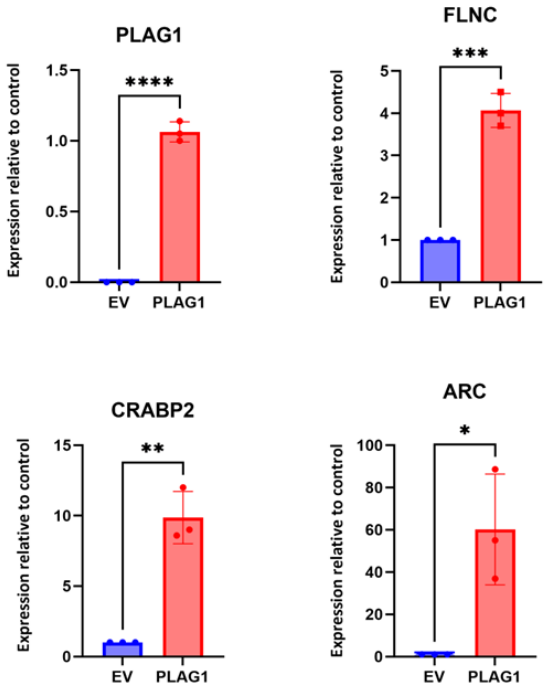


Figure 3.15 Validation of transcriptomics data by qPCR.
Relative gene expression levels of PLAG1, FLNC, CRABP2, and ARC in PLAG1-overexpressing cells compared to control cells. Gene expression was quantified by qPCR and normalised to a housekeeping gene. Data represent mean ± SEM (n = 3 biological replicates). Statistical significance was determined by Student's t-test. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

3.4 Discussion

The overexpression of PLAG1, a transcription factor, has been identified as a key driver for developing pleomorphic adenomas in salivary glands. This abnormal expression disrupts target genes, necessitating the identification of these genes and the investigation of the biological processes affected by PLAG1 and its downstream targets. This is a crucial step in our ongoing efforts to understand the tumorigenic effects induced by PLAG1.

This study aimed to explore the role of the PLAG1 gene in PA by modelling its tumorigenic process in primary human salivary gland cells. Various attempts to transfect these cells failed, as confirmed by fluorescence microscopy and Western blot analysis. Transfection of primary salivary gland cells from multiple donors produced inconsistent results, and contamination with HEK293 cells cast doubt on previous findings. To address these issues, BMI-1 transduced cells, with extended proliferative capacity, were used but the exogenous protein could not consistently be detected. Consequently, the decision was made to employ HEK293 cells, renowned for their robust transfection efficiency, as a model system for subsequent analyses.

RNA sequencing was employed to conduct a comprehensive gene expression analysis following the induction of PLAG1 overexpression in HEK293 cells, thereby identifying the altered functions contributing to the tumorigenesis of PA.

Total RNA from three biological repeats was isolated and sequenced. Following the identification of differentially expressed genes, dataset validation was confirmed by conducting principal component analysis PCA and gene clustering analysis. Our data showed two distinct clusters representing PLAG1-treated cells and control cells. A comparison between the two conditions uncovered 1148 differentially expressed genes (DEGs). We showed that 1092 genes were significantly upregulated, while 56 genes were downregulated. The list of significantly upregulated genes includes Filamin C (FLNC), cellular Retinoic Acid Binding Protein 2 (CRABP2), and Activity-Regulated Cytoskeleton-Associated Protein (ARC). According to Gene Ontology (GO) Annotations, these genes play vital roles in various critical biological processes and functions.

The FLNC gene is a key component in cell junction organisation. It assists in cross-linking actin filaments, which is crucial for cell movement and stability. Changes in FLNC expression can impact

cell motility, a crucial factor in cancer metastasis. Overexpression of FLNC has been linked to enhanced invasiveness and poor patient prognosis in certain cancers, including glioblastoma multiforme (GBM; (Kamil et al., 2019).

Changes in FLNC expression can impact cancer cells' ability to degrade and move through the ECM, which is essential for cancer cell invasion and metastasis. Additionally, FLNC is involved in various signalling pathways, including the MAPK signalling pathway, which is known to impact cell proliferation, differentiation, and survival. Dysregulation of these pathways may contribute to cancer progression (Yang et al., 2017).

Gene Ontology analysis revealed that CRABP2 is involved in the transport of retinoic acid (RA) from the cytosol to retinoic acid receptors in the nucleus. Retinoic acid is a vital signalling molecule that plays a crucial role in the process of cell differentiation during embryonic development and tissue maintenance (Yurdakok-Dikmen et al., 2017). It exerts its effects by binding to the retinoic acid receptor (RAR) and subsequently inducing the transcription of target genes. This, in turn, leads to the differentiation of responsive cells, thereby promoting tissue-specific functions.

In the context of cancer, RA has been observed to induce cell differentiation, while research in the field of head and neck squamous cell carcinomas has demonstrated that retinoic acid has the ability to promote differentiation and prevent growth through the regulation of genes related to cell cycle control and apoptosis (Osanai et al., 2023). Interestingly, RA can exhibit dual effects; while it promotes cell differentiation in certain contexts, it has also been found to enhance the stemness and progenitor cell growth in other cancer cell types. In a study conducted by Mishra et al. in 2018, the role of retinoic acid in the proliferation of neural stem and progenitor cells (NSPCs) in the adult hippocampus was investigated. The study demonstrated that RA signalling plays a crucial role in maintaining NSPC populations and promoting the generation of new neurons. It was shown that inhibiting RA synthesis or signalling significantly reduces NSPC proliferation by affecting cell-cycle kinetics and regulators. Furthermore, RA contributes to apoptosis in specific cell types and is also known to induce cell-cycle arrest, thereby preventing uncontrolled cell proliferation (Dobrotkova et al., 2017). In the case of breast cancer, RA has been

shown to inhibit breast cancer cell proliferation by blocking cell cycle progression, specifically in the G1 phase (Jiménez-Lara et al., 2010).

ARC was significantly upregulated in this study and although ARC is primarily involved in synaptic function and memory consolidation within neurons, some studies have looked into its role in solid cancers particularly renal cell carcinoma (RCC) (Toth et al., 2017). ARC plays a crucial role in determining cell fate and whilst initially, it was thought to primarily inhibit apoptosis, later studies demonstrated that it also has a regulatory role in other types of cell death such as autophagy-related pathways where cells degrade and recycle their own components (Zhang et al., 2021).

As mentioned above, ARC expression suppresses tumour cell apoptosis in renal cell carcinoma. ARC is highly overexpressed in RCCs and contributes to resistance to therapy. It achieves this by cooperating with anti-apoptotic Bcl-2 family members to exert its strong anti-apoptotic effects (Toth et al., 2017).

In brief, FLNC might help stabilise tumour cells by maintaining their structure and cellular architecture. This could prevent them from becoming invasive or spreading to other parts of the body. CRABP2 is involved in retinoic acid signalling, promoting cell differentiation and regulating growth, which are characteristics of non-cancerous tumours. On the other hand, ARC allows tumour cells to escape programmed cell death and continue to grow by inhibiting apoptosis. The seemingly conflicting actions of PLAG1 downstream targets in tumour formation may explain why PLAG1 specifically leads to the growth of non-cancerous tumours, such as pleomorphic adenomas of the salivary glands. Therefore, it could be suggested that the balance between PLAG1 targets, which both promote and hinder tumour formation, leans towards modestly stimulating cell growth, as shown in pleomorphic adenomas.

To corroborate the RNA sequencing findings, qPCR was performed on the same panel of genes identified as being significantly upregulated in response to PLAG1 overexpression. Consistent with the transcriptomic data, a significant increase in expression levels of PLAG1, FLNC and CRABP2 was observed in PLAG1-overexpressing cells compared to controls. While ARC also demonstrated a significant increase in expression ($*p < 0.05$), the magnitude of upregulation was less pronounced compared to the other genes. This discrepancy might be attributed to a larger degree

of variation between biological replicates for this specific gene, potentially influencing the statistical power of the analysis. Nonetheless, the consistent upregulation of all three genes in response to PLAG1 overexpression strongly supports a regulatory role for PLAG1 in their expression.

The Gene Ontology (GO) of the DEGs showed significant enrichment of the gene sets related to the ECM. These include the external encapsulating structure, collagen-containing extracellular matrix, and extracellular matrix binding. The extracellular matrix (ECM) is a highly intricate network composed of a diverse range of macromolecules that envelop the cells within the body (Romer et al., 2021). Of these macromolecules, collagens stand out as the primary constituent of the ECM, accounting for approximately 30% of the total protein mass (Frantz et al., 2010). Studies have suggested that tumour-associated alterations to the ECM have the potential to influence all hallmarks of cancer as the remodelling of the ECM leads to a multitude of biophysical and biochemical changes that impact cell signalling, ECM stiffness, cell migration, and the progression of tumours (Egeblad et al., 2010; Pickup et al., 2014).

PA are characterised by a diverse combination of epithelial and myoepithelial elements encased within a fibrous capsule. Histologically, these tumours manifest as sheets of cells, papillary projections, cords, and individual cells, often concurrently present within the same tumour. The transition from PA to Ca ex PA is marked by crucial ECM alterations (Antony et al., 2012). Notably, Ca ex PA demonstrates heightened stiffness compared to PA. In vitro experiments using organ-derived ECM hydrogels with varying degrees of stiffness have revealed that increased stiffness can enhance proliferative and invasive characteristics of tumour cells (Harmsen et al., 2024). Consequently, In the context of Ca ex PA, ECM stiffness may significantly influence cancer cell behaviour and potentially contribute to the progression of Ca ex PA.

The gene sets GOMF_GROWTH_FACTOR_BINDING, and GOMF_GROWTH_FACTOR_ACTIVITY displayed significant enrichment, attributed to the upregulation of various growth factors and their corresponding binding proteins. Notably, enriched genes in these categories include IGF2, transforming growth factor beta (TGFB), platelet-derived growth factor beta (PDGFB), bone morphogenic protein 6 (BMP6), and their pertinent binding proteins.

The significant role of IGF2 as a potent stimulator of cell proliferation during both embryonic development and tumorigenesis has been extensively documented in the literature (DeChiara et al., 1990; Baker et al., 1993; Toretsky and Helman, 1996; Burns and Hassan, 2001). There is strong evidence to suggest that the oncogenic potential of PLAG1 may, at least in part, be attributed to its ability to induce IGF2 expression. PLAG1 binds to the IGF2 promoter, enhancing its transcription and resulting in elevated levels of IGF2 protein, which in turn activates downstream signalling pathways such as the AKT and MAPK pathways which promote cell cycle progression and inhibit apoptosis, thereby contributing to tumour growth. The combination of increased cell proliferation and decreased apoptosis creates an environment favourable for tumour formation (Voz et al., 2000).

TGFB is a versatile cytokine that controls cell growth, differentiation, and immune responses. In the context of cancer, TGFB plays a dual role. During the early stages, it functions as a tumour suppressor by inhibiting cell proliferation, however, in the later stages, it can facilitate tumour progression, invasion, and metastasis by triggering epithelial-mesenchymal transition (EMT) and immune evasion (Massagué, 2008).

In cancer, the activation of PDGFB signalling can promote tumour growth and metastasis by stimulating angiogenesis and recruiting stromal cells that support the tumour microenvironment (Andrae., et al 2008). BMP6, a member of the TGF-beta superfamily, plays a role in bone and cartilage development, as well as cell differentiation. Its impact on cancer varies depending on the context. In some cases, BMP6 can hinder tumour growth by prompting differentiation and apoptosis. Conversely, in other scenarios, it may actually facilitate tumour progression (Ye et al., 2009).

The expression of these growth factors alone may not be sufficient to definitively determine the tumour's nature and behaviour. Nevertheless, their enrichment, along with the activation of the ECM pathways, suggests a potentially more aggressive tumour phenotype.

Gene set enrichment analysis of KEGGS pathways uncovered functional enrichment of several canonical pathways, including BASAL_CELL_CARCINOMA and HEDGEHOG_SIGNALING_PATHWAY.

The pathogenesis of basal cell carcinoma is closely associated with the constitutive activation of the sonic hedgehog (Hh) signalling pathway. This pathway is recognised as a pivotal regulator during embryogenesis, governing fundamental processes such as cellular proliferation, differentiation, and tissue patterning (di Magliano et al., 2003; Marini et al., 2011). In the adult organism, it plays a critical role in maintaining stem cells, tissue repair, and regeneration (Petrova and Joyner, 2014). Furthermore, its involvement in various human cancers is well-documented, where it facilitates the expansion and proliferation of tumour cells. Notably, effector molecules implicated in tumour cell proliferation, including BMP and cyclin, which are enriched in our dataset, have been identified as target genes of the Hh pathway or downstream effectors (Paiva et al., 2010). Sonic hedgehog (SHH) serves as the ligand protein in the Hh signalling pathway and binds to the Patched (PTCH) receptor upon secretion, relieving its inhibitory effect on the Smoothened (SMO) transmembrane protein. Consequently, SMO activation triggers the upregulation of the glioma-associated oncogene (GLI) family of transcription factors, which orchestrate processes including proliferation, differentiation, and interactions with the ECM (Caro and Low, 2010). Additionally, crosstalk has been reported between the Hh signalling pathway and other signalling cascades, such as Notch, Wnt, and Ras pathways (Paiva et al., 2010; Shi et al., 2011; Min et al., 2011), all of which were noted to be enriched in this present analysis.

In the present study, our findings indicate an upregulation of all previously mentioned Hh signalling pathway members (SHH, PTCH, SMO and GLI) and downstream targets (BMP and cyclin). Notably, a study conducted by Vidal et al. (2016) exhibited significant expression of Hh signalling pathway components in PA, ACC, and MEC, suggesting the involvement of the Hh signalling pathway in the tumorigenesis of these neoplasms. These observations were consistent with those reported by Xu et al., (2012) who also documented heightened expression of Hh signalling components in benign thyroid tumours. Furthermore, analogous studies have reported elevated expression of Hh pathway components in other benign tumours including gastric adenomas (Lee et al., 2007), intestinal adenomas (Oniscu et al., 2004), and ameloblastomas (Kanda et al., 2013). These collective findings propose that the Hh pathway may be active in the early stages of tumour development, potentially contributing to the pathogenesis of benign neoplasms within the salivary gland, including PA.

Another enrichment was observed in the CELL_ADHESION_MOLECULES_CAMS pathway. Cell adhesion molecules (CAMs) are proteins found on the cell surface that participate in cell-cell and cell-ECM interactions. They are vital for maintaining tissue structure, regulating cell movement, and facilitating signalling. In healthy tissues, the expression and activity of CAMs are tightly regulated. However, in cancer, changes in the expression and function of CAMs can lead to the development of tumours, invasion, and metastasis (Kyoto Encyclopaedia of Genes and Genomes, 2024).

Activation of the CAM pathway indicates a potential involvement in tumour development. They can enhance tumour growth by enabling cell-cell interactions and offering anchorage for tumour cells (Harjunpää et al., 2019).

CAMs play a role in mediating the interactions between tumour cells and the surrounding stroma, thereby influencing tumour angiogenesis, immune response, and drug resistance (Harjunpää et a., 2019). Additionally, CAMs enable tumour cells to adhere to and migrate through the ECM and blood vessels, playing a crucial role in the metastatic spread of cancer cells to distant organs (Huang et al., 2021).

While there may be limited studies and conflicting information on CAMs in PA (Andreadis et al., 2020), it is reasonable to speculate that CAMs could play a role in maintaining the tumour's structure and potentially influencing its growth and recurrence, given the tumour's complex architecture involving epithelial and myoepithelial components. Therefore, it is important to conduct a more in-depth analysis of specific CAMs enriched in our transcriptomic data to understand their exact role in the context of PA. This could serve as a guide for future studies in this area.

In the analysis of Oncogenic Signature Gene Sets, it was observed that PLAG1-induced DEGs exhibited enrichment in the ESC_V6.5_UP_EARLY.V1_DN gene set. This gene set includes genes involved in the early stages of differentiation of embryoid bodies that are derived from V6.5 embryonic stem cells. These genes are normally highly expressed in undifferentiated embryonic stem cells (ESCs) but become downregulated as the cells start to differentiate. This

downregulation is vital for the shift from a pluripotent state to a more specialised cell type (Sene et al., 2007; Tamayo, 2024)

In this study, the ESC_V6.5_UP_EARLY.V1_DN gene set is downregulated, indicating suppression of genes associated with stemness and pluripotency in PLAG1-expressing cells. In the context of PA, the downregulation of this gene set suggests that PA cells have reduced stem-like features compared to ESCs, aligning with the tumour's benign nature. Unlike malignant tumours that rely on stemness for invasion and metastasis, PA exhibit a more controlled growth pattern. The decreased expression of pluripotency genes implies a shift towards differentiation, which may contribute to the tumour's characteristic architectural complexity and the presence of both epithelial and myoepithelial components. This differentiation is further associated with a more indolent growth pattern, distinguishing PA from aggressive malignancies. This finding supports the tumour's classification as a localised, non-invasive entity with minimal metastatic potential. Furthermore, tumours with a pronounced stem-like profile are often linked to malignancy, recurrence, or resistance to therapy (El-Tanani et al., 2025). The observed downregulation in this study indicates a lower risk of PA transitioning into Ca Ex PA, reducing the likelihood of malignant transformation.

Another significant activation was observed in the LEF1_UP.V1_UP gene set. This gene set consists of genes that are upregulated in response to increased Lymphoid Enhancer-Binding Factor 1 (LEF1) activity, which is a transcription factor involved in the Wnt/β-catenin signalling pathway, that regulates cell proliferation, differentiation, and stem cell maintenance (Medici et al., 2019). The enrichment of this gene set in the context of pleomorphic adenoma suggests heightened LEF1 activity and activation of its downstream pathways.

LEF1 is essential for embryogenesis, stem cell maintenance, and tissue regeneration and in cancers, LEF1 overactivation enhances proliferation, survival, and stem-like properties, potentially contributing to oncogenesis (Clevers and Nusse, 2012). In PA, its increased activity may influence tumour architecture and growth patterns (Swid et al., 2023).

Elevated LEF1 activity in PA has both beneficial and detrimental implications. On the positive side, LEF1 signalling helps maintain the epithelial and myoepithelial components of the tumour,

preserving its structural integrity and benign nature. Additionally, while LEF1 promotes proliferation, in this context, it may support slow, controlled tumour growth without invasion. However, increased LEF1 activity could also pose risks in that it may enhance stem-like properties and plasticity, which could lead to tumour recurrence after incomplete resection. Furthermore, excessive Wnt/β-catenin signalling through LEF1 could promote EMT, a hallmark of malignancy, increasing the potential for progression to Ca Ex PA (Matsumiya-Matsumoto et al., 2022).

In a benign, localised PA, enrichment of LEF1_UP.V1_UP likely reflects a structured and organised growth pattern. However, if signs of recurrence or atypical cellular behaviour are observed, heightened LEF1 activity could indicate early molecular changes associated with increased proliferation or stem-like features. While LEF1 is not inherently indicative of malignancy, its dysregulation warrants careful monitoring, particularly in recurrent or atypical cases (de Lima-Souza et al., 2025).

The CTIP_DN.V1_UP gene set comprises genes upregulated when CTIP (also known as RBBP8) activity is downregulated or suppressed. CTIP is crucial for DNA damage repair through homologous recombination, cell cycle regulation, and tumour suppression. It plays a key role in maintaining genomic stability by facilitating double-strand break repair.

The upregulation of this gene set may disrupt DNA repair mechanisms and alter cell cycle regulation, leading to uncontrolled cell proliferation, which is a critical process in tumorigenesis. Additionally, the activation of this gene set can inhibit apoptosis, making cells more resistant to cell death and supporting their growth (Furuta et al., 2006).

Reduced CTIP activity in PA has both beneficial and detrimental implications. A potential positive outcome is that limited DNA repair capacity may restrict the tumour's ability to adapt to genomic stress, preventing aggressive growth and malignant transformation. Additionally, CTIP suppression may contribute to tumour stability by reducing genomic instability and supporting the slow-growing nature of PA. However, reduced CTIP function could also result in genomic instability, increasing the risk of chromosomal rearrangements or mutations. Although PAs are generally benign, long-term genomic instability may predispose them to malignant transformation into Ca Ex PA. Furthermore, CTIP plays a role in regulating cell cycle checkpoints,

and its suppression may alter the balance between proliferation and differentiation, potentially allowing tumour cells to persist in a dysregulated state (Xu, et al., 2020).

Clinically, the enrichment of CTIP_DN.V1_UP aligns with the benign nature of PA, as reduced CTIP activity may limit the tumour's ability to develop aggressive features. However, long-term suppression of DNA repair mechanisms could increase the potential for malignant progression, particularly if additional genetic or environmental factors contribute to instability. If the tumour remains benign, the enrichment of this gene set likely reflects its slow-growing and non-invasive nature. However, if signs of progression emerge, enrichment of CTIP_DN.V1_UP may indicate early molecular changes that warrant closer monitoring.

The conditional overexpression of PLAG1 also results in the enrichment of multiple gene sets responsive to KRAS, as indicated in Table 3.5. The KRAS.600_UP.V1_DN gene set consists of genes that are typically downregulated when the KRAS oncogene is activated (MSigDB, 2024). Its enrichment in this study suggests that these genes, which are normally suppressed during KRAS activation, remain active, indicating a lower level of KRAS signalling. This provides insight into the role of KRAS in the biology of PA, suggesting that its signalling is not fully engaged in this tumour type.

KRAS is a key regulator in the RAS/MAPK signalling pathway, driving cell proliferation, survival, and differentiation. In many cancers, KRAS mutations or overactivation lead to uncontrolled growth and tumour progression (Pylayeva-Gupta et al., 2011; Kinsey et al., 2019; Wang et al., 2021). The presence of genes from the KRAS.600_UP.V1_DN set in PA implies that KRAS signalling remains subdued, preventing the unchecked proliferation often associated with malignancy. This aligns with the slow-growing and non-invasive nature of PAs, where low KRAS activity may help maintain the characteristic epithelial and myoepithelial architecture. The absence of strong KRAS signalling may also contribute to the tumour's benign nature by reducing the likelihood of malignant transformation, a process frequently linked to KRAS hyperactivation in other cancers.

While reduced KRAS signalling supports a stable and differentiated tumour phenotype, it could also have implications for tumour persistence. A limited proliferative drive might contribute to slow but continuous tumour growth, leading to recurrence if not completely resected.

Additionally, the suppression of KRAS may lead to the activation of alternative pathways, such as PI3K/AKT or Wnt/β-catenin, which could support tumour maintenance through different mechanisms. If the tumour were to acquire further genetic alterations, these compensatory pathways might influence its progression (Wang et al., 2013).

From a clinical perspective, the enrichment of KRAS.600_UP.V1_DN reinforces the idea that PAs are typically benign and stable, with a low risk of aggressive behaviour. The absence of strong KRAS activation suggests that the tumour remains differentiated and less prone to malignant transformation. However, in cases where recurrence or progression is observed, it may indicate that alternative signalling pathways have compensated for the lack of KRAS activity, warranting further investigation into their role in disease progression.

The enrichment of the WNT_UP.V1_UP gene set in this study suggests the activation of WNT signalling, a critical pathway involved in regulating processes such as cell differentiation, proliferation, migration, and stem cell maintenance (Clevers & Nusse, 2012; Nusse & Clevers, 2017). The WNT_UP.V1_UP gene set refers to genes that are upregulated when WNT signalling, particularly the canonical pathway, is activated. This activation typically promotes cell proliferation and can influence differentiation and stemness by regulating the expression of target genes associated with these processes (Zhan et al., 2017).

One of the positive outcomes of this activation is the promotion of differentiation in tumour cells. In some contexts, WNT activation can help maintain the differentiated epithelial and myoepithelial components that are characteristic of PAs and contribute to their benign nature (Kleinsmith & Martin, 2019). In PA, moderate activation of WNT signalling may support controlled tumour growth and maintenance, ensuring that the tumour grows in a regulated manner and does not progress to malignancy. This activation helps maintain the structural integrity of the tumour by balancing cell proliferation and differentiation.

However, there are potential negative consequences associated with the overactivation of WNT signalling. Overactivation of the canonical WNT pathway in particular is associated with uncontrolled cell proliferation, which could contribute to tumorigenesis (Polakis, 2012). While PAs are generally benign, excessive WNT signalling may lead to increased growth and, in rare cases,

facilitate malignant transformation, such as Ca Ex PA. Dysregulated WNT signalling could also promote the survival of undifferentiated or cancer stem-like cells, which might increase the risk of tumour recurrence (Reya & Clevers, 2005). Additionally, WNT signalling is linked with the maintenance of stem cell-like properties in cancer cells. Overactivation of this pathway could enhance the stemness of PA cells, making them more resistant to therapies and potentially increasing the likelihood of recurrence or malignancy (Zhong and Virshup, 2020). The non-canonical WNT signalling pathway, which regulates cell migration and invasion, could also contribute to a more aggressive tumour phenotype if aberrantly activated (Katoh & Katoh, 2007). While PAs are typically non-invasive, excessive WNT signalling might induce invasive characteristics if the tumour undergoes transformation or if signalling is misregulated (Song et al., 2024).

The BMI1_DN.V1_UP gene set comprises genes upregulated when BMI1 activity is diminished, indicating pathways related to differentiation, cell cycle arrest, or senescence are active. The enrichment of the BMI1_DN.V1_UP gene set in this study suggests a downregulation of BMI1, a component of the Polycomb Repressive Complex 1 (PRC1) known for its role in stem cell maintenance, proliferation, and senescence regulation. BMI1 is crucial for cell cycle progression and protecting cells from senescence. In various cancers, BMI1 overexpression has been linked to tumorigenesis (Park et al., 2004).

In the context of PA, the suppression of BMI1 activity leads to the upregulation of genes associated with differentiation, senescence, and cell cycle arrest. This downregulation enhances the differentiation of tumour cells, supporting the characteristic epithelial and myoepithelial components of PAs, thereby contributing to their benign nature and relatively slow growth. Additionally, reduced BMI1 activity can induce cell cycle arrest and senescence, preventing uncontrolled cell proliferation and limiting the tumour's growth potential (Liu et al., 2009). This aligns with the benign behaviour of PAs, as decreased proliferation helps maintain the tumour's localised and non-aggressive characteristics. Furthermore, the loss of BMI1 activity diminishes stem cell-like properties, making PAs less likely to exhibit aggressive behaviours associated with cancer stem cells (Park et al., 2004).

While senescence generally halts cancer cell proliferation, however, premature senescence could impede tissue repair and contribute to tumour stasis or fibrosis, potentially affecting long-term tissue integrity (Molofsky et al., 2003). Furthermore, excessive differentiation might disrupt a tumour's ability to maintain its characteristic cellular architecture, possibly limiting growth. Nonetheless, such outcomes are unlikely in PAs as these tumours are typically benign and stable (Guo et al., 2006).

The PGF_UP.V1_DN gene set represents genes that are typically upregulated in response to PGF activation but are found to be downregulated in PLAG1 expressing cells. PGF is a member of the VEGF family, known for its role in angiogenesis, tumour progression, and immune modulation (Ceci et al., 2020). The downregulation of this gene set suggests that the biological processes normally induced by PGF signalling are suppressed in our dataset.

A key consequence of PGF_UP.V1_DN downregulation is the potential reduction in angiogenesis, the process by which new blood vessels form. Since PGF plays a crucial role in promoting vascularisation, its suppression could impair a tumour's ability to establish an effective blood supply, potentially limiting growth and metastasis. This could have significant implications for tumour progression as reduced angiogenesis may result in a less aggressive tumour phenotype.

In the context of PA, such downregulation could lead to reduced angiogenesis within the tumour microenvironment, potentially limiting tumour growth and recurrence. This is particularly relevant given that increased expression of growth factors such as PDGF-A, PDGF-B, and FGF-2 has been associated with recurrent PAs (Soares et al., 2012).

Moreover, the suppression of PGF-responsive genes might influence a tumour's potential for malignant transformation. Enhanced angiogenesis has been observed in Ca Ex PA, especially in cases progressing to salivary duct carcinoma. Therefore, downregulation of the PGF pathway could theoretically reduce the risk of such malignant transformations by limiting the vascular support necessary for tumour progression (Suzuki et al., 2021).

Beyond its role in angiogenesis, PGF also contributes to tumour cell survival and inflammatory signalling. The downregulation of PGF-responsive genes could indicate a weakened tumour-promoting microenvironment, which may influence the recruitment of immune cells and the

overall inflammatory response. This shift could either hinder tumour growth by reducing supportive signals or, conversely, make the tumour more resistant to immune surveillance mechanisms, depending on the specific context of the tumour type and surrounding tissue (Adini et al., 2002; Kim et al., 2012; Albonici et al., 2019).

The enrichment of the TGFB_UP.V1_UP gene set indicates active TGF-β signalling, which plays a multifaceted role in tumour biology. As mentioned earlier, TGF-β is recognised for its dual functions: acting as a tumour suppressor during the initial stages of tumorigenesis by inhibiting cell proliferation and inducing apoptosis, and promoting tumour progression in later stages through mechanisms such as EMT and immune modulation (Massagué, 2008). The specific impact of TGF-β is highly dependent on the tumour microenvironment and cellular context. In the case of PA, elevated TGF-β signalling presents both advantageous and potentially adverse implications.

On the beneficial side, TGF-β contributes to the regulation of epithelial and myoepithelial cell differentiation, which is essential for the characteristic mixed cellular composition observed in PAs. This regulatory function aids in maintaining tumour stability and prevents uncontrolled cellular proliferation, thereby preserving the benign nature of the tumour. Moreover, TGF-β plays a role in ECM remodelling and the maintenance of stromal integrity, reinforcing the tumour's structural organisation and reducing the likelihood of invasive behaviour. By suppressing excessive cell growth, TGF-β further supports the well-differentiated and localised phenotype typical of Pas (Hao et al., 2029).

However, excessive activation of TGF-β signalling in PAs could potentially result in the induction of EMT, a process where epithelial cells acquire mesenchymal characteristics, leading to increased motility and invasiveness (Peinado et al., 2003). While PAs are benign, the initiation of EMT could elevate the risk of recurrence or, in rare instances, malignant transformation into Ca Ex PA. Additionally, TGF-β is known to suppress immune responses by inhibiting the activity of cytotoxic T cells and natural killer (NK) cells, which may enable tumour cells to evade immune surveillance, thereby increasing the potential for persistence and regrowth. Furthermore, TGF-β's involvement in promoting angiogenesis and fibroblast activation, while necessary for tumour maintenance,

could, if unregulated, create conditions favourable for tumour progression (Rak et al., 2000; Miyazono et al., 2012).

The transcriptomic analysis has unveiled a complex interplay of molecular pathways triggered by PLAG1 overexpression in the context of PA. This overexpression leads to alterations in ECM composition, growth factor signalling, and oncogenic pathways, ultimately influencing tumour characteristics. The remodelling of the ECM and the resultant increase in its stiffness may contribute to the progression towards Ca ex PA.

Upregulation of growth factors such as IGF2, TGFB, PDGFB, and BMP6 suggests enhanced proliferation and potential tumour maintenance. While IGF2 promotes growth TGFB's dual role in suppression and EMT-driven progression highlight a risk for malignant transformation. Wnt/β-catenin and Hedgehog pathway activation further suggest developmental signalling involvement in PA growth.

Enriched CAMs may support tumour structure but also influence recurrence. Limited KRAS activation and BMI1 suppression indicate a benign, slow-growing nature, while reduced angiogenesis (PGF suppression) may constrain tumour expansion. However, suppressed DNA repair (CTIP) and excessive TGFB activity raise concerns for genomic instability, recurrence, and malignancy.

Overall, these findings support a largely stable PA phenotype but highlight pathways—ECM remodelling, growth factor signalling, and Wnt/LEF1 activation—that warrant further investigation for recurrence and malignant potential. Clinically, these insights may aid in prognosis, risk assessment, and therapeutic targeting.

Voz et al. (2004) conducted a comprehensive study to identify target genes regulated by the PLAG1. Since PLAG1 is a transcription factor known to activate gene expression by binding to specific DNA motifs, its deregulated expression is thought to contribute to tumorigenesis through the aberrant activation of downstream genes.

To investigate the molecular mechanisms underlying PLAG1-induced tumorigenesis, the study utilised oligonucleotide microarray analysis to assess gene expression changes in two key models:

(1) human fetal kidney 293 cell lines engineered for inducible PLAG1 expression and (2) PA compared to normal salivary gland tissue. This approach enabled the identification of 47 genes upregulated and 12 genes downregulated upon PLAG1 induction.

Among the most significantly upregulated genes were growth factors and cytokines, including insulin-like growth factor II (IGF2) and cytokine-like factor 1 (CLF1), both of which are implicated in cell proliferation and survival. In silico analysis revealed that many upregulated genes contained multiple PLAG1-binding motifs in their promoters, suggesting direct transcriptional activation by PLAG1. Furthermore, several transcription factors, ECM-related proteins, and oncogenes were also induced, reinforcing the role of PLAG1 in altering cell signalling and tumour microenvironment interactions.

The study also compared gene expression profiles between PA and normal salivary gland tissue, identifying 12 consistently upregulated genes in both the in vitro and in vivo settings. This finding provided strong evidence that these genes are key mediators of PLAG1-driven tumorigenesis in PA. The deregulation of IGF2 was particularly notable, as its overexpression is known to activate the AKT and MAPK pathways, thereby promoting cell cycle progression and inhibiting apoptosis.

Overall, Voz et al. (2004) provided critical insights into the oncogenic role of PLAG1, highlighting its impact on growth factor signalling, transcriptional regulation, and extracellular matrix remodelling. These findings laid the groundwork for subsequent studies exploring the mechanistic basis of PA development and potential therapeutic targets for PLAG1-driven tumours.

This findings of our study align with the findings of Voz et al. (2004), further confirming the role of PLAG1 in PA tumorigenesis. Both studies demonstrate that PLAG1 overexpression leads to the upregulation of key growth factors and transcription factors, including IGF2 and CRABP2, which play crucial roles in promoting cell proliferation and survival. Additionally, both studies highlight the involvement of the PI3K-AKT and Wnt signalling pathways in PA, reinforcing the idea that PLAG1 contributes to tumorigenesis by regulating genes involved in cell growth and survival. Furthermore, extracellular matrix remodelling was found to be a significant factor in both studies, indicating that PLAG1 influences not only intracellular signalling but also the tumour

microenvironment. These shared findings strengthen the understanding of PLAG1-driven oncogenesis in PA and suggest potential therapeutic targets for future research.

Song et al. (2011) conducted a study to identify differentially expressed genes in PA compared to normal salivary gland tissue. Using cDNA microarrays, they identified a total of 447 DEGs including, Gli2 and CTNNB1, which were significantly upregulated and were identified as being at the core of the DEG network. Gli2 gene is a part of the Hedgehog signalling pathway which plays a crucial role in cell growth and differentiation. CTNNB1 gene encodes β-catenin, a protein involved in the Wnt signalling pathway which is important for cell proliferation and differentiation.

A more recent study by Chen et al. (2023) focused on the role of ECM remodelling in the progression of PA to Ca ex PA. RNA sequencing and bioinformatic analysis of patient derived organoids revealed significant changes in the ECM- associated genes. Among the identified key genes were COLA 1A1, MMP2 and FN1 which were also significantly expressed in our PLAG1 samples. The COLA 1A1 gene encodes type I collagen, a major component of the ECM and upregulation of this gene suggests increased ECM production and remodelling, potentially contributing to tumour stiffness and progression. MPP2 encodes matrix metalloproteinases-2, an enzyme involved in the breakdown of ECM components and so altered expression of MMP2 could lead to ECM remodelling and tumour invasion. The FN1 gene encodes fibronectin, a glycoprotein that helps in cell adhesion and migration and changes in FN1 expression have been associated with alterations in the ECM and tumour metastasis.

The consistency of our findings with previous studies underscores the robustness of our data and the critical role of these pathways in PA. It is important to note, however, that this analysis presents a preliminary overview, and further in-depth analysis is necessary to fully elucidate the underlying mechanisms.

4 Chapter 4 Identifying Protein Interactions following *de novo* expression of PLAG1

4.1 Aims and Objectives

Having observed a clear role for PLAG1 in mediating the crosstalk between the different pathways and biological processes involved in pleomorphic adenoma, as supported by our previous findings (Chapter 3) and those of other studies (Zhao et al., 2006), it was fundamental to identify which proteins were directly associated with PLAG1 and understand the downstream effect of these proteins in the context of this benign tumour. A biotin-based proximity labelling technique was used to explore the network of interactions between PLAG1 and any proximal molecules. Such molecules will likely be implicated in the biological function of PLAG1 and may be implicated in PA.

To our knowledge, this is the first study to investigate the PLAG1 interactions network using this approach. The work outlined in this chapter aimed to:

- Identify PLAG1 interacting proteins using PLAG1 overexpression and biotinylation of PLAG1 associated proteins.
- Identify the enriched canonical pathways in the dataset, assess their activation state, and determine the most common genes shared by these pathways.
- Identify the upstream regulators expressed in the dataset, predict their effect (activation/inhibition) on the downstream target genes, and build networks describing
- potential molecular interactions of these proteins.
- Identify key biological processes and functions influenced by differentially expressed genes following the biotinylation of PLAG1 proximal proteins and build networks to explain the impact of upstream molecules on downstream biology.

4.2 Principle

TurboID (Figure 4.1) works by labelling nearby proteins in a cell using biotin, a natural coenzyme that strongly binds to glycoproteins like avidin or streptavidin. This allows the tagged proteins to be easily purified and identified using avidin-coated techniques. TurboID is an engineered BirA biotin-protein ligase. It efficiently labels neighbouring proteins in just 10 minutes.

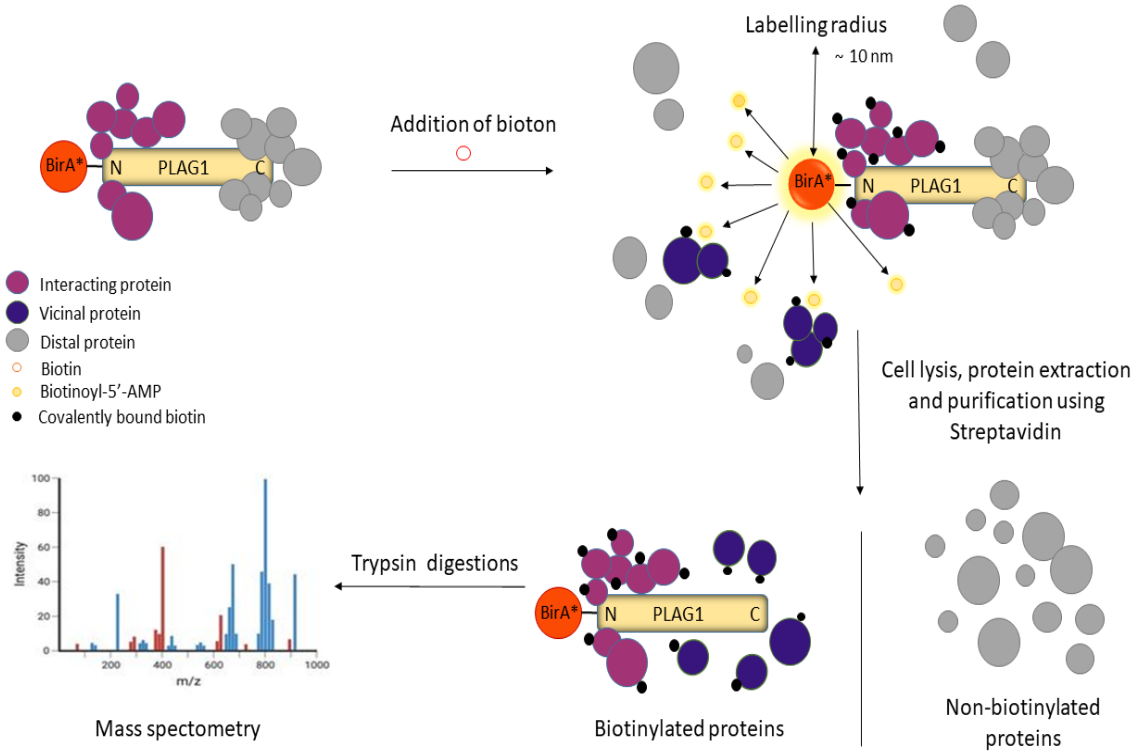


Figure 4.1 Schematic Representation of the TurboID Experimental Workflow.
In the TurboID experiment, the protein of interest, PLAG1, is genetically fused with a mutant E. coli biotin ligase (BirA*) construct and subsequently expressed in the target cells, specifically HEK293 cells. Upon expression, PLAG1 forms protein complexes with its endogenous interactors. The introduction of biotin then leads to the biotinylation of the solvent-accessible lysine side chains of the protein complex partners within the defined labelling radius. Notably, protein interactors situated beyond the labelling radius, despite their direct interaction with PLAG1, remain unaltered. Moreover, non-interacting proteins within the 10 Å radius are prone to false positive labelling. Following cell lysis, the biotinylated proteins are pulled down using streptavidin-coated beads through affinity purification. Subsequently, these beads undergo trypsin treatment, yielding tryptic peptides, which are then subjected to mass spectrometry analysis for comprehensive identification and quantification.

4.3 Methods

The relevant materials and methods used in this study are detailed in Chapter 2. Human embryonic kidney (HEK293) 293 cells were transiently transfected with two distinct constructs: TurboID empty vector (EV) as a negative control and FlagTurboID PLAG1 fusion construct. Following transfection, HEK293 cells expressing PLAG1 were visualised by fluorescent microscopy, treated with 0.5mM Biotin for 20 minutes and western blot and immunofluorescence staining used to confirm protein biotinylation, PLAG1 expression, and localisation. Streptavidin beads were used to pull the biotinylated proteins from the cell lysate, and purified protein lysates were subjected to mass spectrometry analysis. The list of the differentially expressed proteins (DEPs) was used to perform functional analysis using Ingenuity Pathway Analysis (IPA) software to identify relationships, functions, and pathways relevant to the DEPs.

4.4 Results

4.4.1 Generation of FLagTurboID PLAG1 construct

To facilitate proximity labelling and interaction studies, a fusion construct was designed by inserting the PLAG1 sequence in-frame with the TurboID biotin ligase in the FLAG-TurboID plasmid. This ensures that the expressed FLAG-TurboID-PLAG1 fusion protein retains both the transcriptional activity of PLAG1 and the biotinylation capability of TurboID, allowing for efficient identification of interacting proteins. The FLAG-TurboID-PLAG1 construct was commercially synthesised and purchased from Biomatik Corporation (Ontario, Canada). As a negative control, an empty FLAG-TurboID vector (EV) was used, which lacks the PLAG1 insert but retains all other vector elements.

To confirm the correct insertion and orientation of the PLAG1 sequence, the construct was subjected to restriction enzyme digestion analysis using NheI (GCTAGC) and XhoI (CTCGAG). These enzymes specifically cleave the vector at predefined sites, allowing for verification of successful cloning and sequence integrity. Proper digestion patterns were analyzed by agarose gel electrophoresis, ensuring the correct integration of the PLAG1 insert into the TurboID expression system (Figure 4.2).

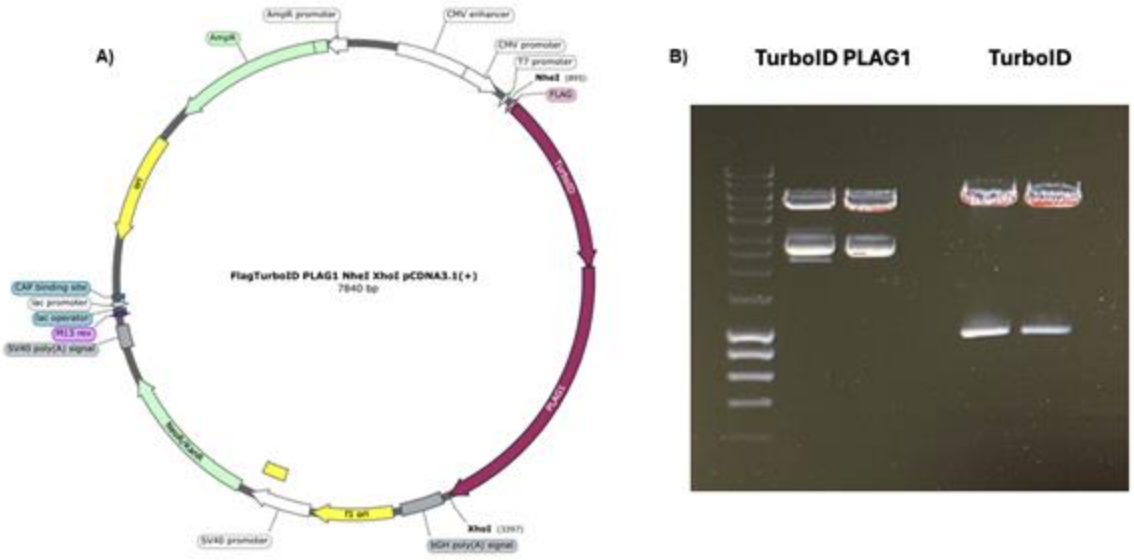


Figure 4.2 Construct map and restriction digestion.

A. The map of the FlagTurboID PLAG1 construct generated by Biomatik. B. Duplicate plasmid preps of TurboID PLAG1 and TurboID empty vector were digested with NheI and XhoI to release the synthesised fragments.

4.4.2 Characterisation of TurboID-PLAG1 construct

4.4.2.1 Fluorescence microscopy imaging

FlagTurboID PLAG1 was transiently transfected into human embryonic kidney cells (HEK293 cells 293) using jetOPTIMUS DNA transfection reagent (Polyplus, France). Twenty-four hours post-transfection, construct expression and transfection efficiency were assessed using an inverted fluorescent microscope (ZOE Fluorescent Cell Imager, Bio-Rad, UK) (Figure 4.3). Quantifying PLAG1 expression helps determine whether sufficient levels of the fusion protein are present for downstream biotinylation and interaction studies. If expression is too low, subsequent analysis may not yield meaningful results.

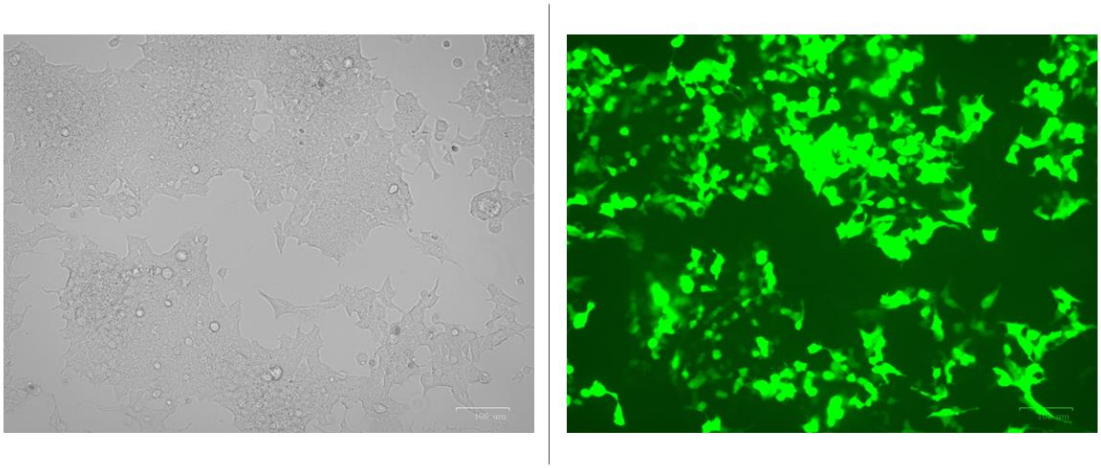


Figure 4.3 Fluorescence microscopy images of transiently transfected cells HEK293
A representative image of transiently transfected cells HEK293 with FlagTurboID PLAG1. The brightfield image (left) and the eGFP expression (right) captured 24 hours post-transfection demonstrate notably high transfection efficiency rates.

4.4.2.2 Plasmid localisation and detection of PLAG1 protein expression

After successful transfection, PLAG1 protein expression and construct localisation were evaluated using anti-PLAG1 antibodies for Western blotting and immunofluorescence staining. Western blot analysis (Figure 4.4) verified the successful expression of the PLAG1 fusion protein at the expected

molecular weight (~56 KDa), ensuring that the construct was properly translated and functional. Immunofluorescence staining (Figure 4.5) confirmed the correct subcellular localisation of PLAG1, predominantly in the nucleus, consistent with its known biological role. These validation steps are critical quality control measures before initiating biotin-based proximity labelling, as they ensure that the observed protein interactions reflect physiologically relevant processes. Establishing expression, stability, and appropriate localisation of PLAG1 is essential for the reliability of downstream analyses.

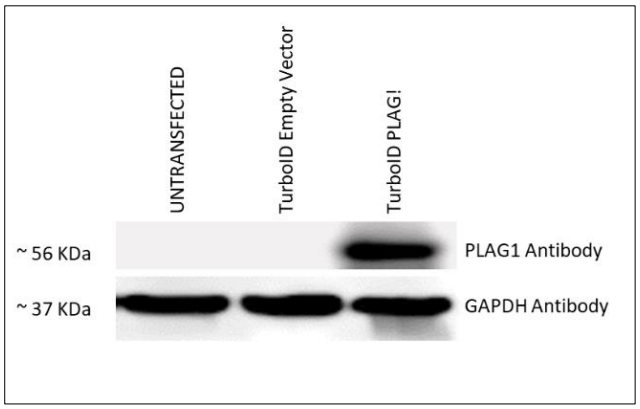


Figure 4.4 PLAG1 protein expression by western blot.
The presence of PLAG1 protein was confirmed in samples transfected with TurboPLAG1, while untransfected and EV samples exhibited no production of PLAG1.

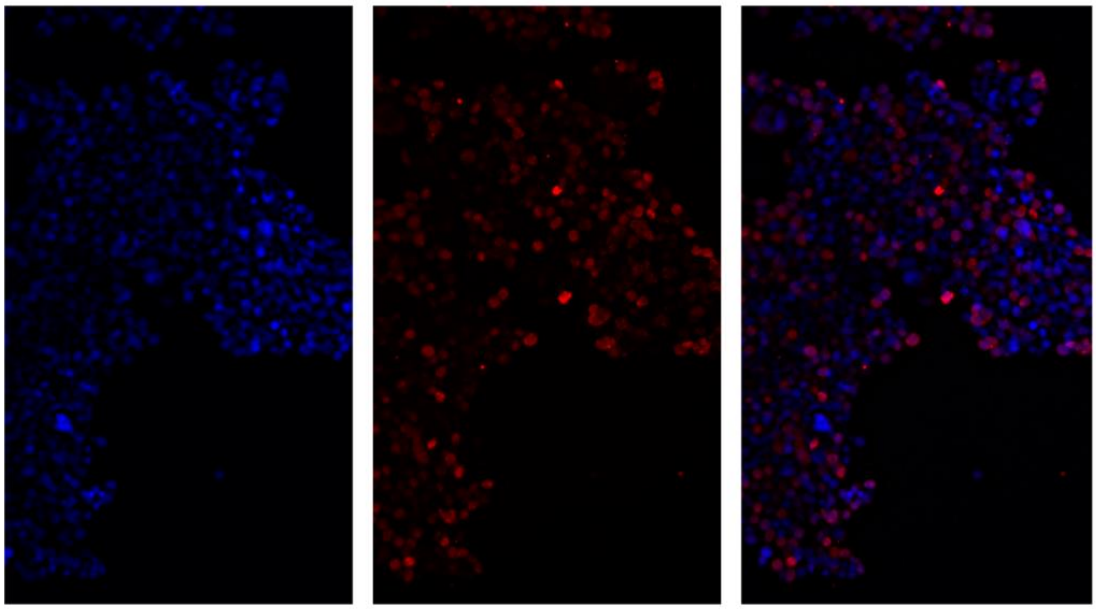


Figure 4.5 Localisation of FLagTurboID PLAG1 construct by IF.
DAPI staining (blue, left panel) marks the nuclei, while immunofluorescence staining of FlagTurboID PLAG1 (red, middle panel) visualises its cellular localisation. The merged image (right panel) demonstrates strong colocalisation of PLAG1 with DAPI, indicating that the FlagTurboID PLAG1 construct is predominantly localised in the nucleus. Minimal red fluorescence is observed outside the nuclei, suggesting limited or absent cytoplasmic expression. Images were captured at 20× magnification.

4.4.3 Streptavidin purification of biotinylated proteins

After verifying the appropriate localisation of the FLagTurboID PLAG1, the activity of the construct was evaluated by western blot analysis of biotinylated proteins from the whole-cell lysate using Streptactin (Precision Protein StrepTactin-HRP Conjugate, 1:5000, #1610381 -Bio Rad, UK) (Figure 4.6) and by immunofluorescence staining of biotinylated proteins using Streptavidin conjugated with Alexa Fluor (Streptavidin, Alexa Fluor™ 647 conjugate, 1:100, Catalogue number: S21374) (Figure 4.7). Western blotting using StrepTactin detection allows for the verification of biotinylated protein presence, specificity, and enrichment efficiency, distinguishing between target proteins and potential background noise or non-specific interactions. By comparing protein profiles before and after enrichment, it is possible to assess the efficiency of streptavidin-based purification and validate the selective isolation of biotinylated proteins.

Similarly, immunofluorescence staining using StrepTactin antibodies serves as an additional confirmation step, enabling the visualisation of subcellular localisation of biotinylated proteins. This approach ensures that the biotinylation process has occurred in a physiologically relevant manner and that the enriched proteins retain their expected localisation. Together, these validation techniques confirm the success of the biotinylation and enrichment process, ensuring the reliability of subsequent proteomic and functional analyses.

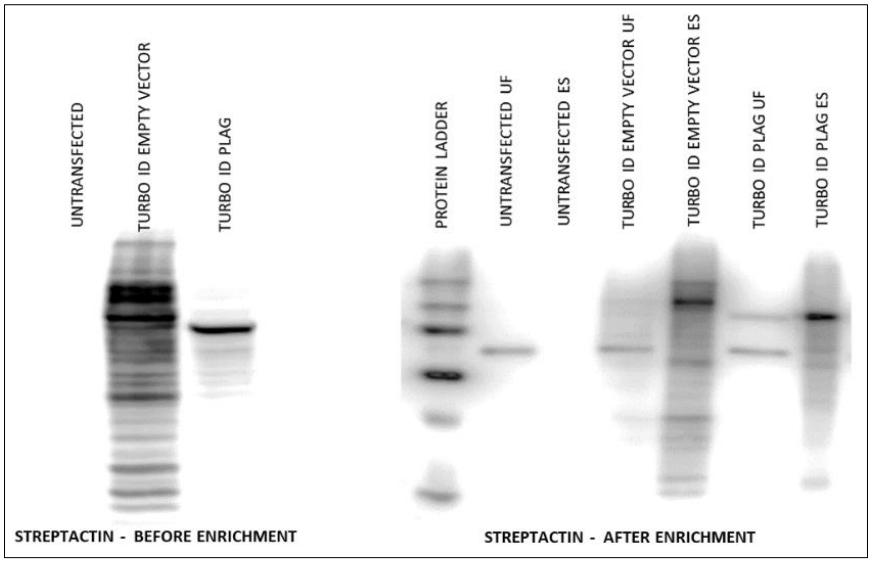


Figure 4.6 Evaluation of protein biotinylation before and after enrichment (purification) using Streptavidin antibody

The left panel (before enrichment) shows total protein lysates from untransfected cells, TurboID empty vector (EV), and TurboID PLAG1-expressing cells. Minimal signal is observed in untransfected cells, while multiple bands in TurboID EV indicate background biotinylation or non-specific binding. TurboID PLAG1 displays a distinct banding pattern, suggesting successful biotinylation of target proteins.

The right panel (after enrichment) presents protein fractions following streptavidin-based purification. The unbound fraction (UF) contains proteins that did not bind to the streptavidin beads, while the enriched sample (ES) represents successfully isolated biotinylated proteins. TurboID PLAG1 ES exhibits defined bands, confirming successful enrichment of biotinylated proteins, whereas TurboID EV ES retains some bands, likely due to non-specific binding. The untransfected ES sample shows no detectable bands, confirming enrichment specificity.

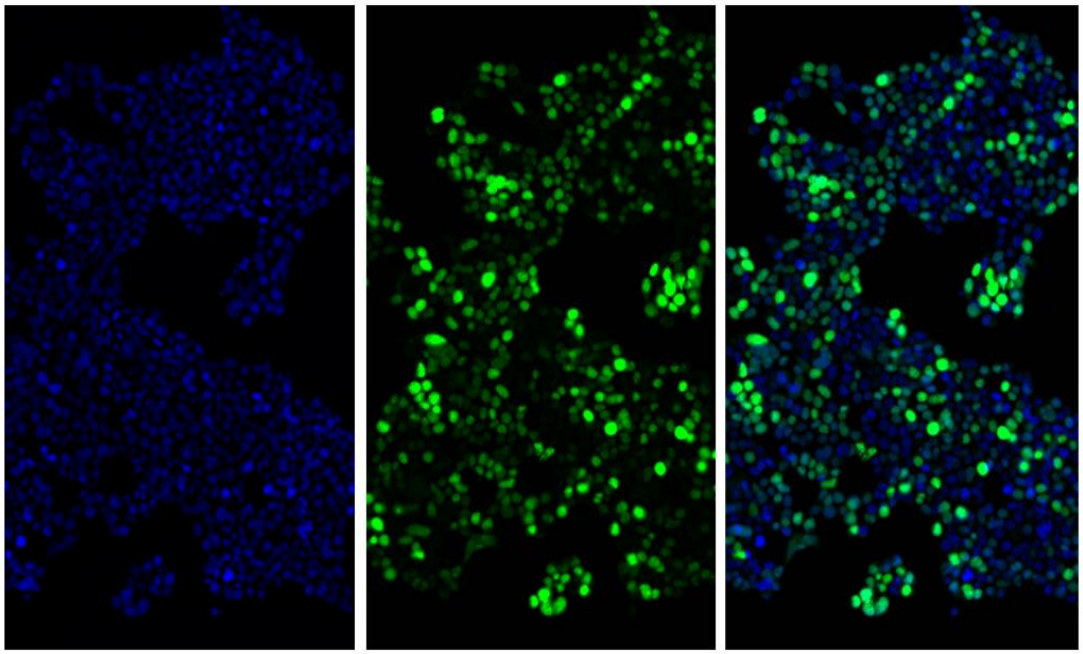


Figure 4.7 Immunofluorescence staining of biotinylated proteins using StrepTactin antibody.
The left panel (blue channel - DAPI staining) shows nuclear staining with DAPI, serving as a reference for cell nuclei. The middle panel (green channel - StrepTactin staining) visualises biotinylated proteins detected using a StrepTactin antibody conjugated to a fluorescent dye. The right panel (merged image - DAPI + StrepTactin) overlays the DAPI-stained nuclei with the StrepTactin signal, illustrating the subcellular localisation of biotinylated proteins. In the merged panel, most of the green signal overlaps with the blue DAPI-stained nuclei, suggesting that the biotinylated proteins are primarily localised in the nucleus. There is minimal green fluorescence outside the nuclei, indicating that cytoplasmic expression is limited or absent. This confirms successful biotin labelling and retention of physiological localisation.

4.4.4 Identification of proteins associated with PLAG1

4.4.4.1 Unsupervised clustering methods

The MS data generated from the TurboID experiments was firstly subjected to principal component analysis, a starting point for most testing methods, to evaluate the similarity between the four biological replicates and provide insight into sample grouping based on the distance matrix. Secondly, the relationship between all samples was further studied by clustering analysis (heatmap) based on quantities of the DEPs. The unsupervised cluster analysis of the two sample sets delineated the presence of two distinct molecular clusters, representing the control (Empty TurboID Vector EV) and PLAG1 expressing cells (TurboID PLAG1), with the exception of sample 4, which exhibited a slight deviation from the PLAG1 expressing group (Figure 4.8). Despite meticulous maintenance of controlled conditions, the observed variance may stem from technical

variability introduced during sample handling and processing. Discrepancies in transfection efficiency, protein labelling, or purification across samples might account for experimental differences. However, our data demonstrates consistent PLAG1 expression values (LQF intensity) across the four biological replicates, indicative of uniform transfection efficiency.

Additionally, a P-value histogram (Figure 4.9) was used to inspect data distribution and assess the validity of the results.

The initial comparison between the two distinct clusters (TurboID PLAG and EV expressing samples), provided with the MS data revealed nine differentially expressed proteins with an FDR < 0.05 and a fold change > 1 as shown in the volcano plot in Figure 4.10 (A). However, following discussion with a bioinformatician we recognised that this criterion was too stringent and might lead to false negative results. Therefore, the analysis was adjusted to use a significance cut-off of p-value < 0.05. The top 50 significantly upregulated proteins are illustrated in the heatmap shown in Figure 4.10 (B).

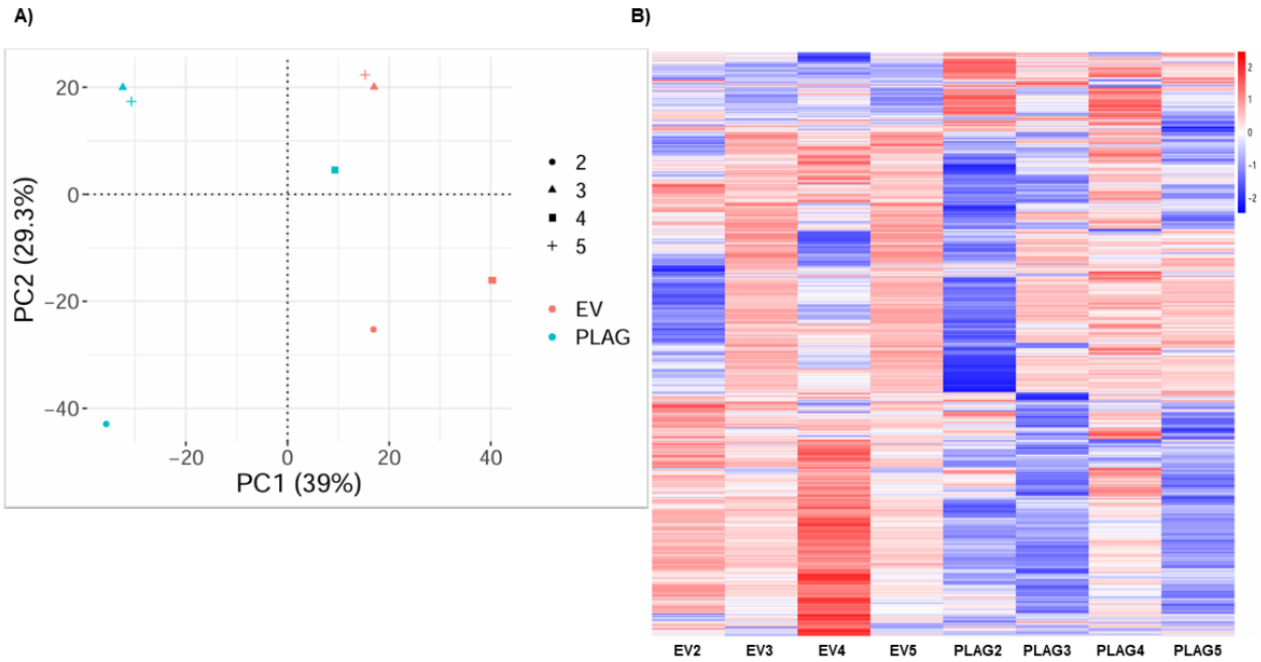


Figure 4.8 Clustering analysis
In Panel A, the principal component analysis (PCA) demonstrates the division of all samples into two distinct clusters, denoted by the red and blue shapes. Significantly, the samples expressing PLAG1

predominantly align to the left of the Y axis, with the exception of sample 4 situated on the right. Conversely, the EV samples exhibit clustering to the right of the Y axis. In Panel B, a heatmap derived from unsupervised cluster analysis depicts the categorisation of eight samples based on all identified proteins. The colour red denotes enriched proteins, while blue indicates proteins with reduced enrichment.

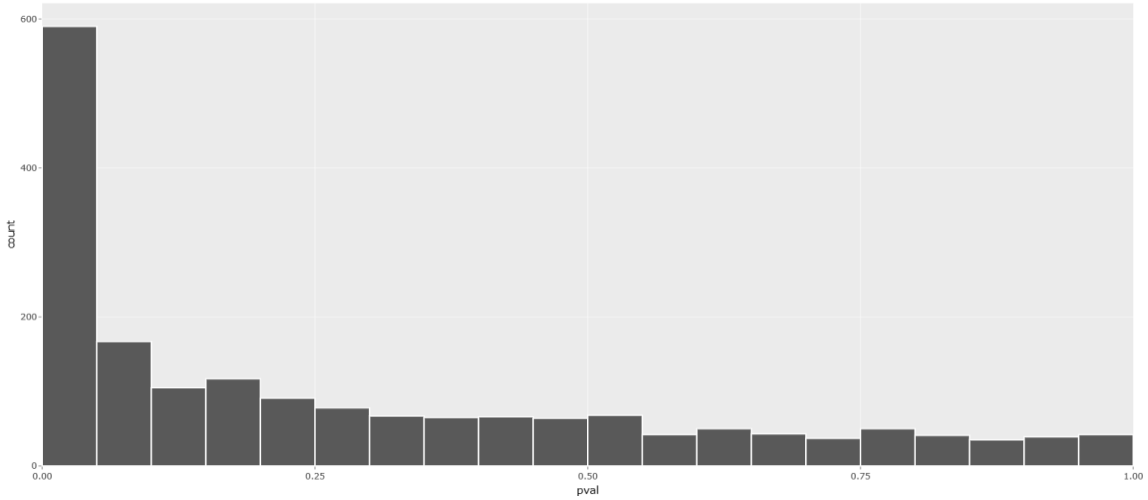


Figure 4.9 P-value histogram plot.

A representative histogram showing an enrichment of small P-values to the left. The x-axis (p-val) represents p-values ranging from 0.00 to 1.00, binned into intervals (0.00, 0.25, 0.50, 0.75, 1.00). The y-axis represents the number of genes within each p-value bin, ranging from 0 to 600. The histogram shows a strong enrichment of genes with very low p-values, as indicated by the tallest bar near 0.00, representing approximately 600 genes. This suggests that a significant number of genes are enriched following PLAG1 expression. The number of genes progressively declines as p-values increase, indicating that fewer genes show non-significant enrichment. This distribution suggests that PLAG1 expression drives significant gene enrichment, with many genes exhibiting strong statistical significance.

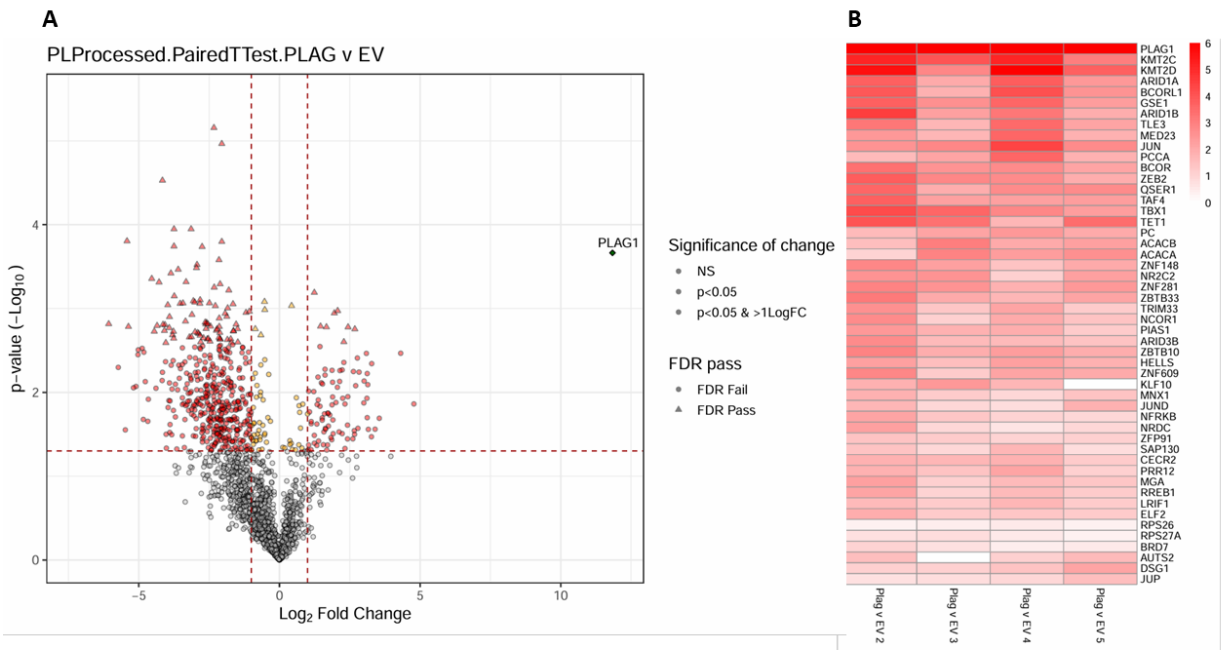


Figure 4.10 Differentially Expressed Proteins (DEPs) from the TurboID experiments
Panel (A) represents a volcano plot**.** The Empty Vector condition on the left side non-specifically labelled numerous proteins with lower abundances. In contrast, the PLAG1 protein interactors on the right side presented fewer proteins but with higher abundances. A cutoff of the p-value < 0.05, along with a log2 fold change of 1, was applied to determine significantly regulated proteins in each pairwise comparison. Only nine proteins were FRD significant (Δ FDR Pass) in the PLAG1 overexpression arm. Panel (B) represents a heatmap of the top 50 most abundant proteins.

4.4.5 Ingenuity pathway analysis (IPA)

Only proteins demonstrating upregulation with real values and meeting the specified cut-off criteria (p-value < 0.05) were considered for the IPA analysis. In order for a protein to be included in the analysis, it needed to be expressed in a minimum of three PLAG1-expressing samples as compared to no expression in any EV samples: 3 v 0, 4 v 1 or 4 v 0. The latter condition represents the ideal inclusion criteria, indicating that the protein is expressed in all PLAG1-expressing samples while showing no expression in any of the control samples.

A detailed overview of the proteins included in the IPA analysis, referred to as analysis-ready molecules, along with their expression data and activation predictions, is provided in Appendix Table 8.1.

4.4.5.1 Canonical pathways

The canonical pathway analysis of differentially expressed proteins using ingenuity pathway analysis (IPA) unveiled the upregulation of several key pathways, such as the activation of anterior HOX genes in hindbrain during early embryogenesis and transcriptional regulation by RUNX1 (Table 4.1 and Figure 4.11). Conversely, our findings indicated a significant inhibition of the PPARα/RXRα Activation and Sumoylation Pathways, as evidenced by the strong negative z-score, suggesting that the proteins enriched in the dataset exert inhibitory effects on these pathways. The z-score, a calculated statistical measure employed by IPA, is utilised to forecast the activation state of any molecule, biological function, or pathway. It assesses the direction of change (upregulation or downregulation) of known targets of a regulator in the dataset against what is anticipated from existing literature. Z-scores greater than 2 or smaller than -2 are commonly deemed significant.

Table 4.1 Top 5 Canonical Pathways

Name	p-value	Overlap
Activation of anterior HOX genes in hindbrain during early embryogenesis	2.87E-12	12.3 % 10/81
Transcriptional regulation by RUNX1	8.26E-11	7.2 % 11/152
Transcriptional regulation of white adipocyte differentiation	3.89E-09	9.5 % 8/84
Formation of WDR5-containing histone-modifying complexes	3.11E-08	14.3 % 6/42
Glucocorticoid Receptor Signalling	4.68E-08	2.6 % 15/582

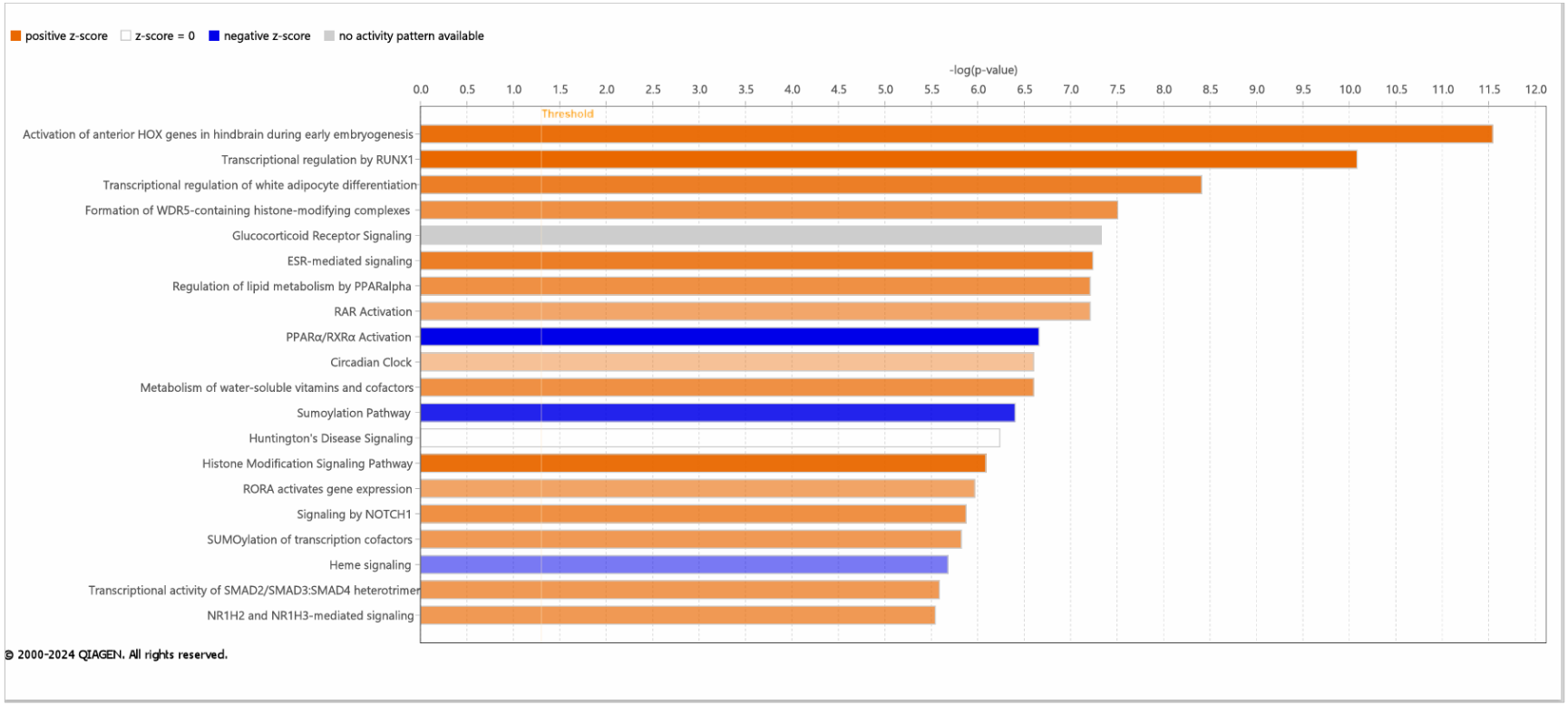


Figure 4.11 Top 20 canonical pathways significantly enriched in PLAG1 expressing cells.
The findings of the canonical pathway analysis are depicted in a horizontal bar chart showcasing the top 20 most significant Canonical Pathways. Pathway names are indicated on the y-axis, while the x-axis represents the negative logarithm of the p-value. The significance of each pathway is reflected in the height of its respective bar, with greater heights signalling increased importance. The chart is arranged so the most significant pathways are at the top. As evidenced in the figure, the Activation of Anterior HOX genes and Transcriptional Regulation by RUNX1 emerge as the most notably enriched pathways across the dataset, displaying positive activity patterns. Conversely, PPARα/RXRα Activation and Sumoylation Pathways exhibited inhibition, evidenced by negative z-scores of -1.89 and -1.63, respectively. Notably, while the Glucocorticoid Receptor Signalling is among the significantly enriched pathways, the software did not provide an activity prediction. It is essential to note that an unfavourable z-score does not diminish a pathway's significance.

1. **Upregulation of anterior HOX genes in hindbrain during early embryogenesis pathway**

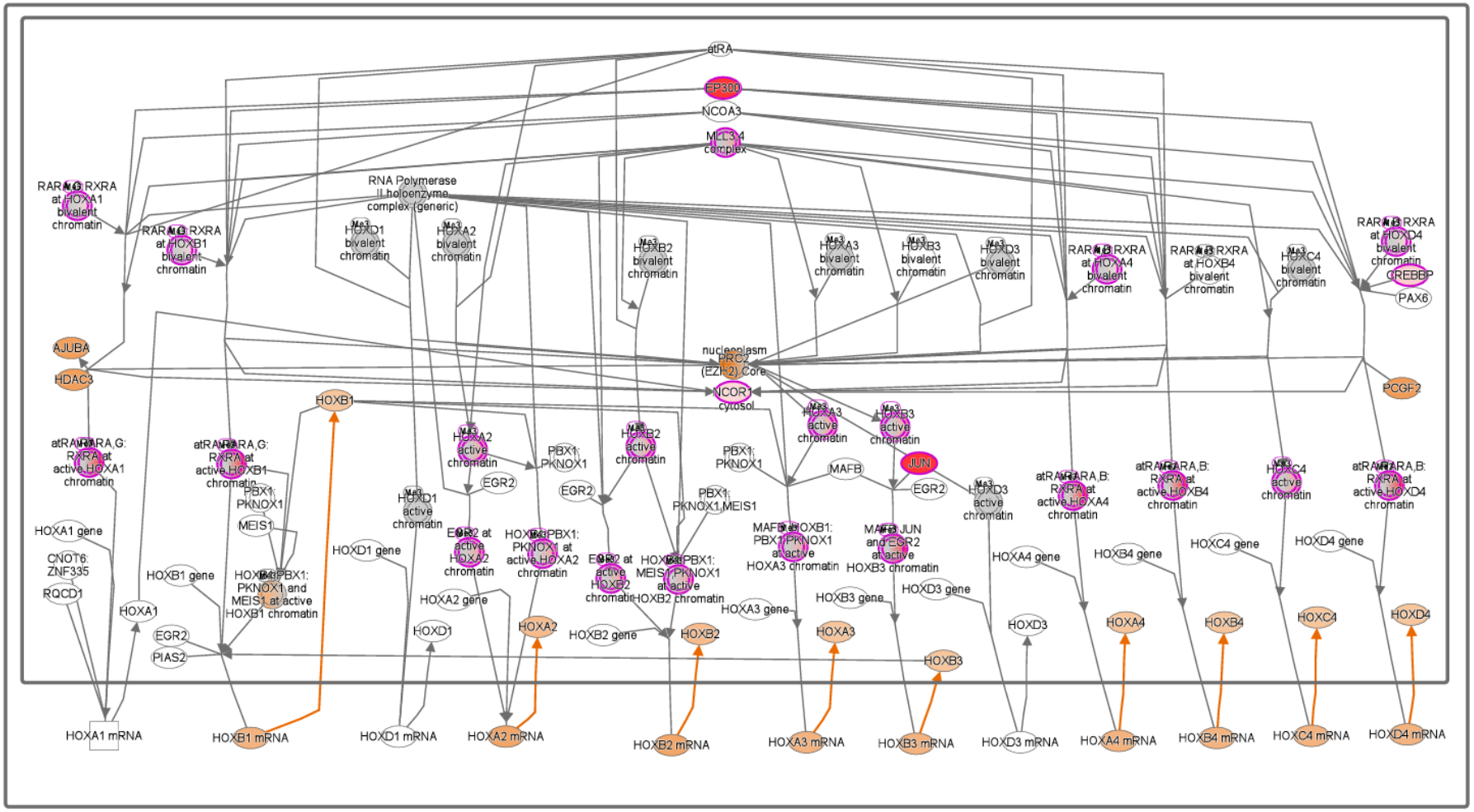


Figure 4.12 Activation of anterior HOX genes in hindbrain during early embryogenesis.

The schematic illustration delineates molecular interactions within the pathway, portraying genes, molecules, or complexes through nodes. Connections between the nodes are indicative of established relationships sourced from the Ingenuity Knowledge Base. Nodes outlined in purple denote molecules identified as enriched in our dataset, with the intensity of the coloured fill reflecting the degree of enrichment. Additionally, blue and orange-coloured molecules and lines denote predicted activation states produced by the Molecular Activity Predictor function in Ingenuity Pathway Analysis (IPA). Blue indicates a predicted inhibition, whereas orange signifies a predicted activation. Yellow lines signify discrepancies between our findings and the state of the downstream molecule. Notably, the geometric shapes of the molecules within the pathway correspond to their functional classifications: Nested Circle/Square = Group/Complex, Horizontal ellipse = Transcriptional Regulator, Vertical Ellipse = Transmembrane Receptor, Vertical Rhombus = Enzyme, Square = Cytokine/Growth Factor, Triangle = Kinase, Circle = Other. Furthermore, the delineation of molecules through edges serves to differentiate direct (solid lines) and indirect (dashed lines) relationships.

HOX genes, acting as master regulators, control cellular processes such as proliferation, adhesion, migration, apoptosis, and tumour progression. Their expression is tightly regulated by transcription factors, including EP300, CREBBP, NCOR1, and JUN, as well as epigenetic modifiers like HDAC3 and KDM6A, which influence chromatin states. In their bivalent chromatin state, HOX genes remain poised for either activation or repression, depending on the regulatory environment. Once activated, these genes undergo transcription, producing HOX mRNA, which drives downstream protein synthesis and cellular effects.

A detailed list of proteins associated with the activation of anterior HOX genes, including their expression levels and predicted activation status, is provided in Appendix Table 8.2.

2. Upregulation of Transcriptional regulation by RUNX1 pathway

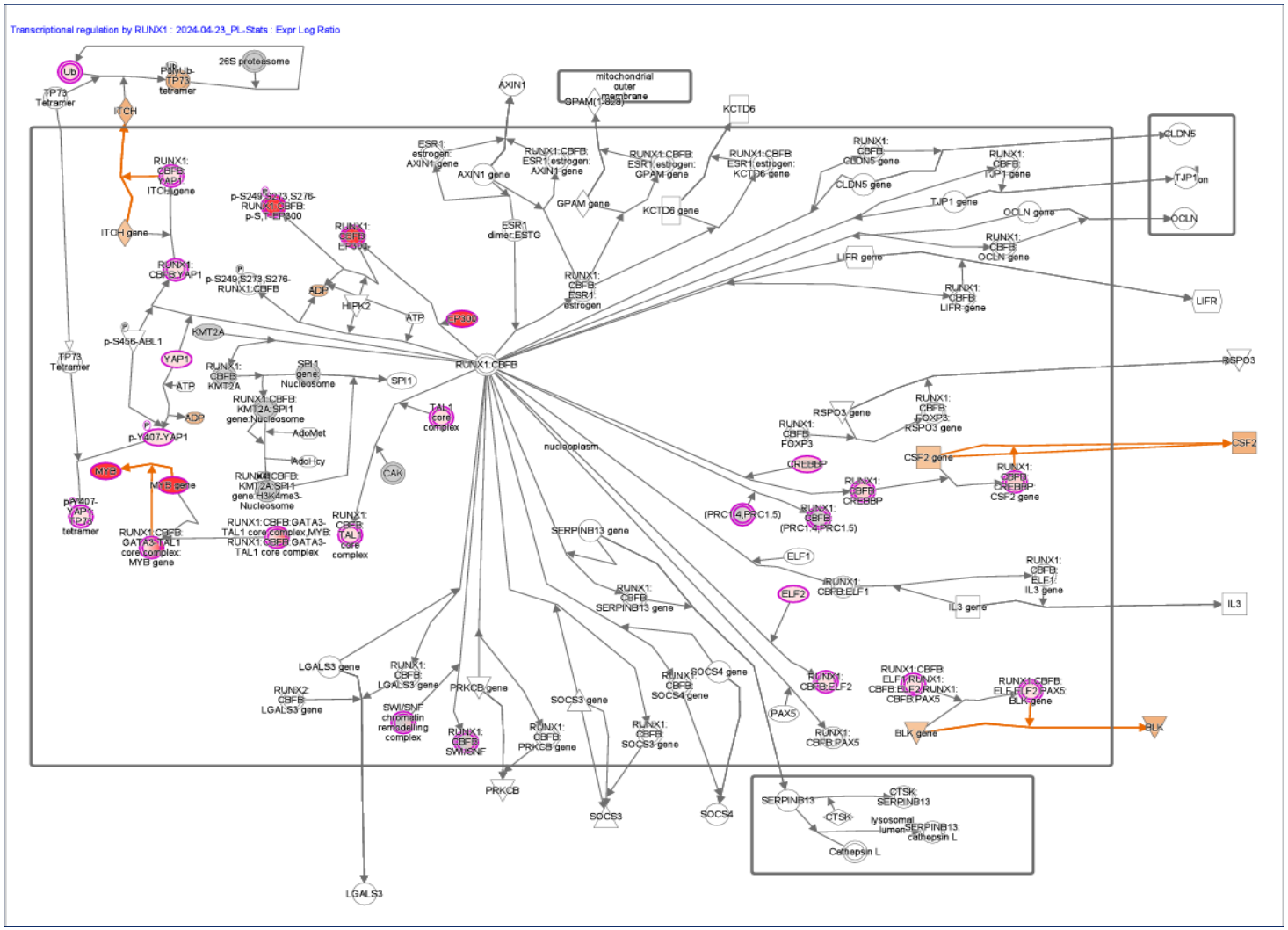


Figure 4.13 Transcriptional regulation by RUNX1.

The diagram highlights RUNX1-mediated transcriptional regulation and its interactions with key factors influencing tumour progression, particularly in pleomorphic adenoma (PA). Genes and proteins directly regulated by RUNX1, such as MYB, EP300, ELF1, and CBFB, are represented by pink ovals, playing roles in chromatin modification and oncogenic signalling. Orange arrows indicate activated oncogenic pathways, including CSF2 and ITGA, which contribute to immune cell infiltration and tumour microenvironment remodelling, while purple diamonds denote essential transcription factors and chromatin regulators like EP300, which modulate chromatin structure and gene expression. Solid black arrows represent direct transcriptional activation or repression by RUNX1, whereas dashed arrows illustrate indirect regulatory interactions, including the potential RUNX1-MYB collaboration, which is essential in hematopoiesis and tumour progression. RUNX1 can function as either a tumour suppressor or oncogene, depending on cellular context and co-factor interactions. Additionally, the involvement of RUNX1-IT1 (a long non-coding RNA) in carcinoma ex pleomorphic adenoma suggests its regulatory influence on miR-195 and CyclinD1, which contribute to tumour growth and metastasis.

Appendix Table 8.3 lists proteins involved in RUNX1-mediated transcriptional regulation, along with their expression changes and predicted activation status**.**

3. Inhibition of 'PPARα/RXRα Activation pathway'

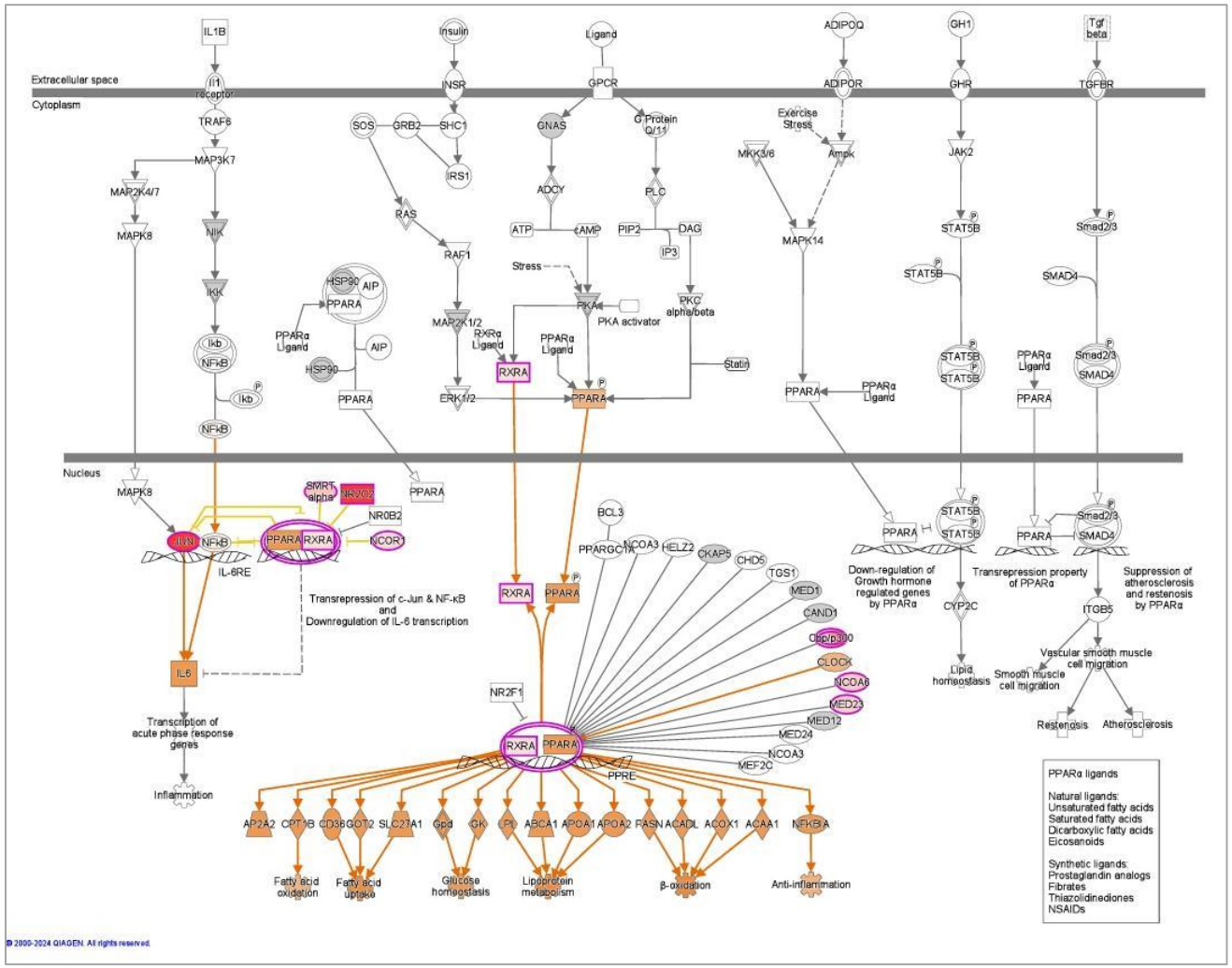


Figure 4.14 PPARα/RXRα pathway map.

In the diagram, pink ovals represent nuclear receptors and transcription factors, such as PPARα (NR1C1) and RXRα, which normally form a heterodimer to regulate target genes. Red rectangles indicate molecules whose altered expression, such as SIRT1 and NR0B2, leads to pathway inhibition, while orange rectangles mark transcriptional targets involved in anti-inflammatory and metabolic processes, including IL6, β-oxidation genes, and glucose homeostasis genes, which are downregulated due to pathway silencing. Grey circles denote upstream regulators, such as cytokines and growth factors, that influence PPARα/RXRα activation. The pathway's normal function is illustrated by orange arrows, which show how PPARα/RXRα promotes lipid metabolism, anti-inflammatory responses, and tumour suppression, whereas yellow arrows depict cross-talk with NF-κB, where PPARα typically inhibits pro-inflammatory cytokines, a regulation that is lost upon inhibition. The consequences of pathway suppression are shown with dashed black lines, representing dysregulated processes such as enhanced angiogenesis, vasculogenesis, endothelial cell proliferation, and reduced apoptosis, all of which contribute to tumour progression. The inhibition of PPARα disrupts its anti-tumorigenic effects, preventing cell-cycle arrest (p18, p21, p27), apoptosis induction, and Bcl-2 suppression, thereby facilitating cancer growth. Additionally, angiogenesis and tumour microenvironment remodelling are enhanced due to the loss of CYP2C suppression, promoting vascularization and tumour expansion. The pathway's inhibition also increases pro-inflammatory signalling by removing PPARα's negative regulation of NF-κB, leading to higher IL6 expression and extracellular matrix protein accumulation, which contribute to fibrosis and chronic inflammation. Lastly, the loss of PPARα's metabolic regulatory function results in reduced fatty acid oxidation and impaired glucose homeostasis, creating a metabolic landscape that may favour tumour survival and proliferation.

Appendix Table 8.4 presents the proteins related to the PPARα/RXRα signalling pathway, highlighting their expression changes and the predicted down-regulation of this pathway.

4. Inhibition of Sumoylation Pathway

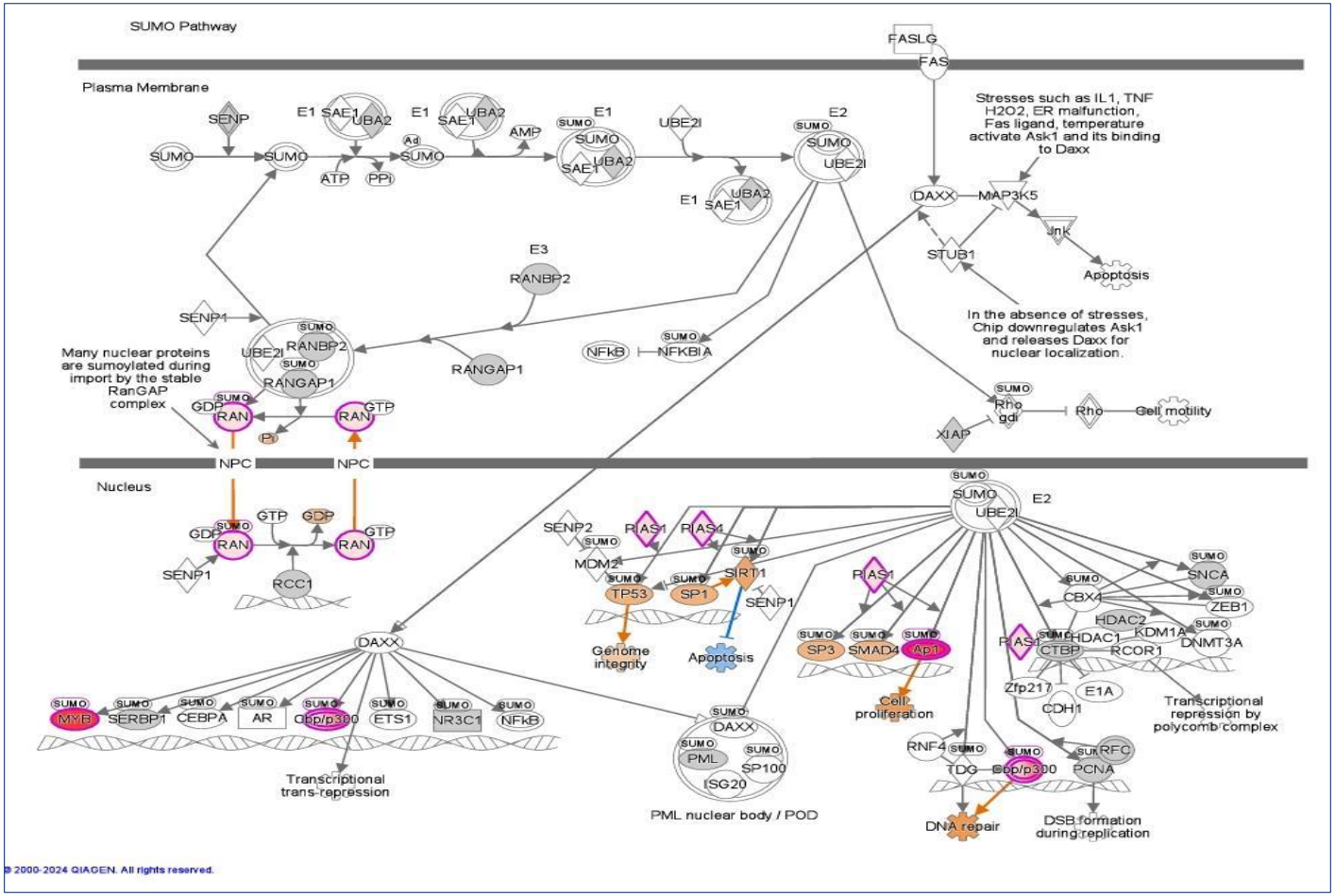


Figure 4.15 Sumoylation Pathway map.

Pink ovals represent SUMO-modified proteins, which normally undergo sumoylation to regulate transcription, DNA repair, apoptosis, and cell proliferation. Orange shapes highlight key downstream cellular functions affected by SUMO modifications, such as genome integrity, transcriptional repression, and stress response. The purple diamonds indicate transcription factors and regulatory proteins that interact with the sumoylation machinery, including SP3, SMAD4, and RAS, which are involved in cell proliferation and differentiation. The black dashed lines illustrate SUMO-dependent transcriptional repression, a process that controls gene expression by modifying transcription factors. The solid black arrows depict normal sumoylation events, where SUMO proteins (SUMO1, SUMO2/3) are conjugated to substrates via the E1 (SAE1/SAE2), E2 (UBE2I), and E3 ligases (such as PIAS and RANBP2).

In the context of PLAG1 overexpression in pleomorphic adenoma, the inhibition of the sumoylation pathway (depicted with grey crossed-out SUMO labels) removes a key regulatory checkpoint that typically suppresses PLAG1's transcriptional activity. This disruption leads to increased activation of PLAG1 target genes, which enhances cell proliferation, inhibits apoptosis, and contributes to tumorigenesis. The loss of SUMO-mediated transcriptional repression also results in the dysregulation of oncogenic pathways, leading to genomic instability and abnormal cell cycle progression. Additionally, key tumour-suppressive processes, such as DNA repair and NF-κB signalling regulation, are impaired, further promoting tumour growth. The enhanced transcriptional activity of PLAG1 due to reduced sumoylation allows it to activate mitogenic and pro-survival genes, driving uncontrolled cellular proliferation and differentiation.

A comprehensive summary of proteins associated with the SUMOylation pathway, including expression changes indicating pathway inhibition, is provided in Appendix Table 8.5.

4.4.5.2 Overlapping canonical pathways

To visualise the shared biology in pathways through the common proteins participating in these pathways, overlapping canonical pathway analysis was conducted. This involved identifying clusters of pathways that share common proteins, predicting potential crosstalk between those pathways, and uncovering hidden connections within the dataset. Figure 4.16 visually outlines the interplay among the top 10 significant canonical pathways, while Table 4.2 summarises the most frequently shared proteins within these pathways.

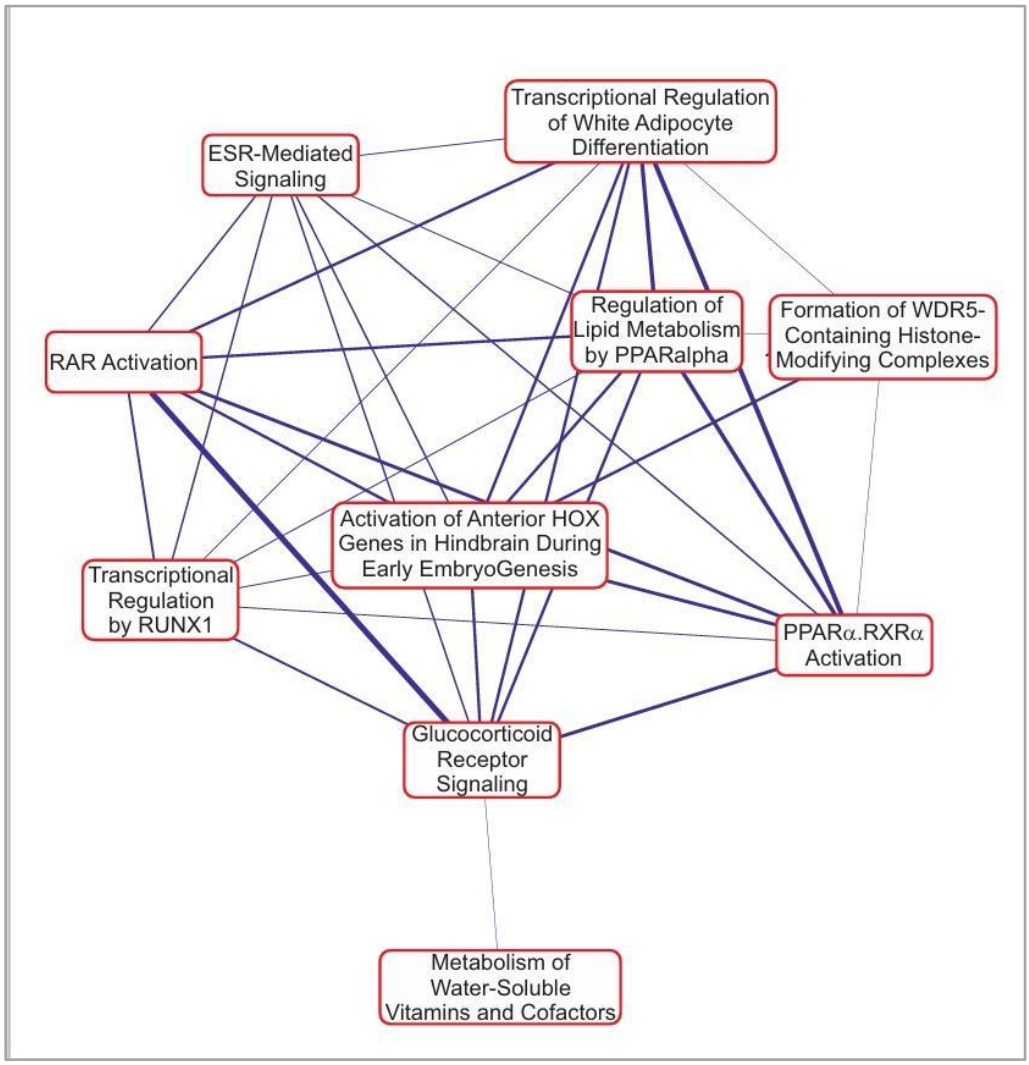


Figure 4.16 Overlapping canonical pathways.

The network of overlapping canonical pathways displays each pathway as a single node, with a line connecting any two pathways when they share at least one data set molecule. The thickness of the lines

is directly proportional to the number of proteins common to both pathways. The map highlights multiple overlapping pathways, with the Activation of Anterior HOX Genes in Hindbrain During Early Embryogenesis emerging as the most interconnected regulatory hub. This pathway interacts extensively with transcriptional regulators (RUNX1, RAR activation, ESR-mediated signalling), metabolic networks (PPARα/RXRα activation, lipid metabolism), and epigenetic modulators (WDR5-containing histone-modifying complexes), underscoring its role in development and oncogenesis. The dysregulation of HOX genes—potentially driven by PLAG1 overexpression—may contribute to pleomorphic adenoma by altering transcriptional programs, sustaining oncogenic transcription, and modifying chromatin accessibility. Additionally, its strong connectivity with glucocorticoid receptor signalling suggests an influence on tumour microenvironment responses. By integrating differentiation, proliferation, apoptosis, and metabolic reprogramming, HOX gene activation appears central to the molecular mechanisms underlying pleomorphic adenoma progression, reinforcing its significance as a primary oncogenic driver within this pathway network.

Table 4.2 Most common proteins associated with the top 10 significant canonical pathways.

Symbol	Entrez Gene Name	Expr Log Ratio	Expr p-value	Expr FDR (q-value)
CREBBP	CREB binding protein	2.416	0.0359	0.129
EP300	E1A binding protein p300	10	0	0
NCOR1	nuclear receptor corepressor 1	1.866	0.0141	0.0818
NCOR2	nuclear receptor corepressor 2	3.121	0.02	0.0972
RXRA	retinoid X receptor alpha	1.64	0.0275	0.112

4.4.5.3 Upstream regulators

The upstream regulator analysis was performed to identify potential upstream regulators that may have contributed to the observed changes in differential protein expression. The Ingenuity Pathway Analysis (IPA) can detect various molecules, such as transcription factors, miRNA, chemicals, drugs, or compounds, that have an impact on the dataset identifiers. These molecules may not be overexpressed in the dataset (no Exp Log ratio shown), however, the activation of such upstream regulators could be due to post-transcription/translation modification without inducing their expression, leading to changes in the expression of the target molecules in the dataset. Furthermore, IPA can predict the activation or inhibition state of these regulators based on their z-score, with a Z-score greater than 2 or smaller than -2 often considered significant**.** Table 4.3 lists the most significant overlapped upstream regulators ranked by their p-value. Figure 4.17 illustrates the most significantly activated and inhibited upstream regulators.

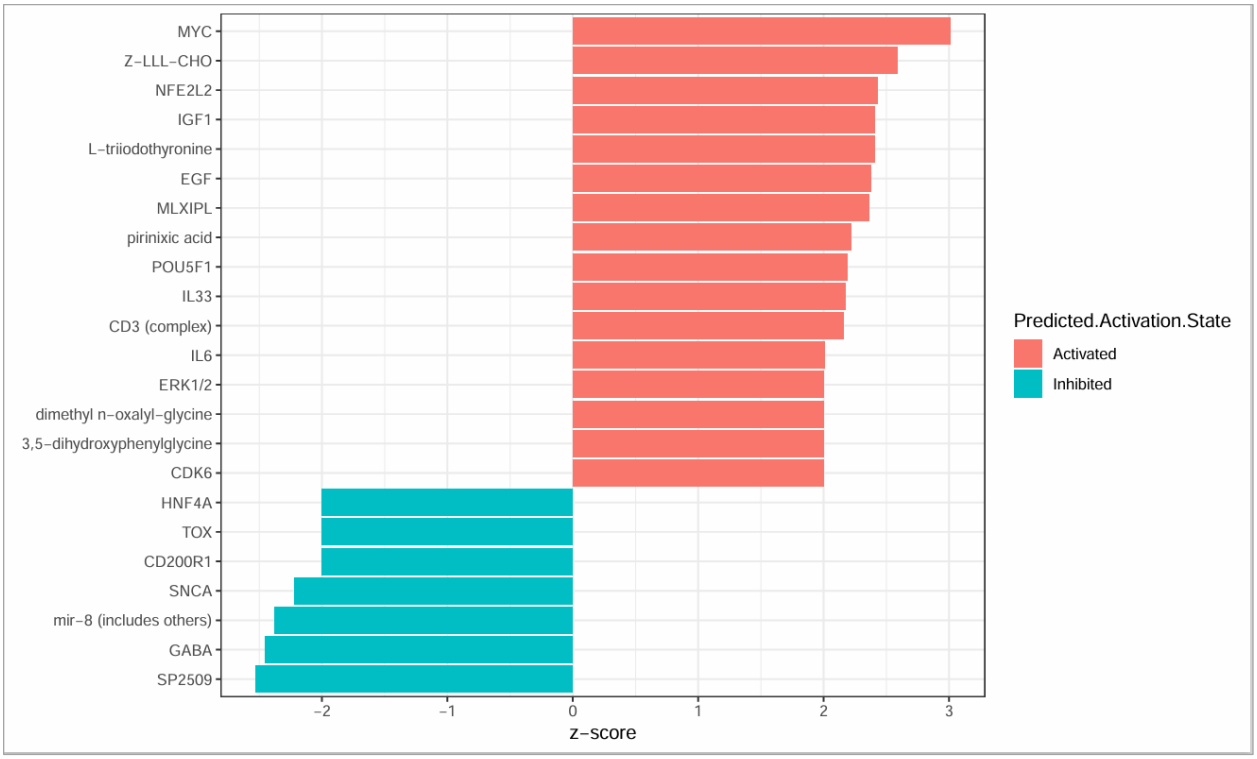


Figure 4.17 Most significantly activated and inhibited upstream regulators.
The bar chart highlights the most significantly activated and inhibited upstream regulators based on their predicted activation state. Activated regulators, shown in red with positive z-scores, include MYC, NFE2L2, IGF1, EGF, and IL6, among others, indicating their potential role in driving the observed biological response. In contrast, inhibited regulators, shown in blue with negative z-scores, include CDK6, HNF4A, TOX, SNCA, and GABA, suggesting their suppression in this context. The x-axis represents the z-score, reflecting the strength of activation or inhibition, while the y-axis lists the specific regulators. This visualization provides insights into the key molecular drivers influencing the system.

Biological upstream regulators (expressed in the dataset with Expr Log Ratio)

As mentioned earlier, IPA can identify any molecule that has a downstream effect on the differentially expressed proteins, irrespective of their expression status. However, this study aimed to primarily define the potential biological upstream regulators expressed in the dataset following the biotinylation of PLAG1 proximal interactors. Table 4.3 lists the most significantly enriched upstream regulators in the dataset based on the expression log ratio and p-value of overlap between the DEPs in the dataset and the data from the Ingenuity Knowledge Base.

Candidate upstream regulators

To further investigate and validate the downstream effect of these upstream regulators, a shorter list of candidate proteins was established based on the following selection criteria:

1. Level of expression (Exp Log Ratio)

   Regulators that induce significant changes in protein expression are likely to exert a substantial impact on biological functions. Additionally, biological validation, such as immunohistochemistry staining, can only be performed if the regulators are actively expressed
2. Statistical significance (p-value)
3. Biological context

   Proteins that are shared among the most significant pathways or connecting multiple downstream genes indicate their importance. (Table 4.3 in bold).

Table 4.3 Biological upstream regulators.
List of upstream regulators expressed in the dataset, along with their Expression Log Ratio and p-value of overlap. Candidate genes for further investigation are highlighted in **bold.**

Upstream Regulator	Molecule Type	Expr Log Ratio	Predicted Activation State	p-value of overlap
MYB	transcription regulator	10		1.34E-03
JUN	transcription regulator	10		6.72E-03
EP300	transcription regulator	10		1.51E-02
TBX1	transcription regulator	10		4.47E-02
NCOA6	transcription regulator	2.667		4.87E-02
NCOA1	transcription regulator	2.45		2.64E-02
CREBBP	transcription regulator	2.416		2.93E-02
ACACA	Enzyme	2.253		3.67E-02
ACACB	Enzyme	2.231		1.97E-03

1. MYB (Myeloblastosis Proto-Oncogene)

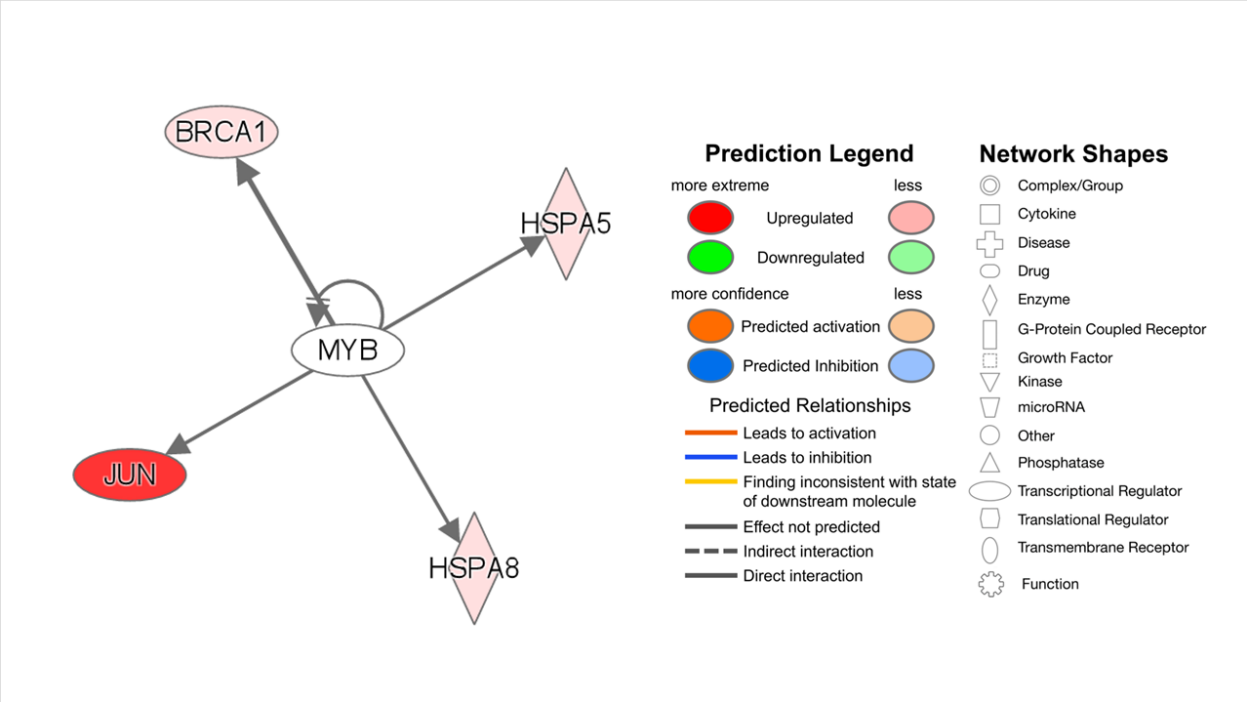


Figure 4.18 MYB network of interactions.
Visualisation of MYB as an upstream regulator with downstream targets in the dataset. Node shapes represent functional classes, and arrows denote the relationship between the nodes. In this specific dataset, MYB is shown to regulate key molecules, including BRCA1, JUN, HSPA5, and HSPA8, which are involved in crucial biological processes such as transcriptional regulation, stress response, and oncogenesis. This network provides insights into the functional impact of MYB and its potential role in cellular pathways, including cancer progression.

Table 4.4 MYB upstream regulator and its downstream targets

Genes in dataset	Prediction (based on measurement direction)	Expr Ratio	Log	Findings
JUN	Affected	10		Regulates (1)
MYB	Activated	10		Upregulates (27)
HSPA5	Affected	0.721		Regulates (1)
BRCA1	Activated	0.686		Upregulates (1)
HSPA8	Affected	0.572		Regulates (1)

2. JUN (Jun Proto-Oncogene, AP-1 Transcription Factor Subunit)

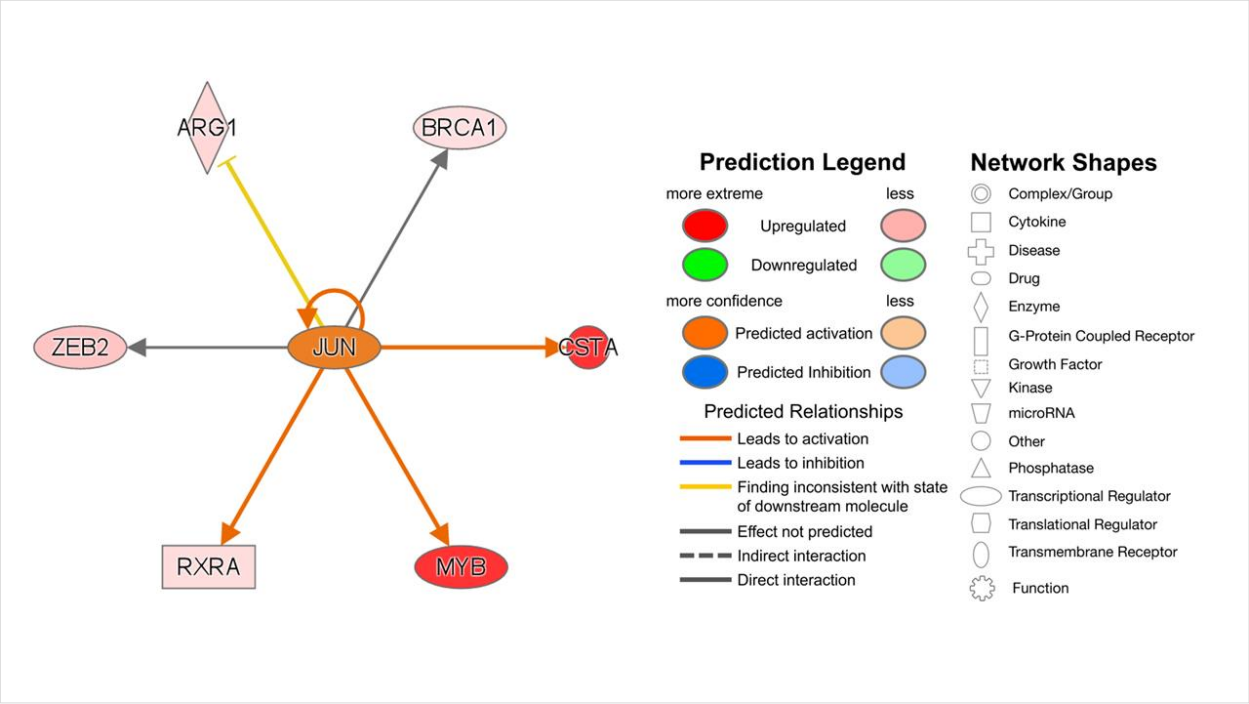


Figure 4.19 JUN network of interactions.
In this dataset, JUN appears to be an active regulatory hub, influencing multiple downstream targets such as CSTA, MYB, RXRA, ZEB2, BRCA1, and ARG1. Notably, CSTA and MYB are upregulated, while JUN itself is predicted to activate several targets, suggesting a pro-survival and transcriptionally active role. This network highlights JUN's involvement in transcriptional regulation, cellular stress response, and potential oncogenic pathways, providing insight into its functional impact within the context of PA.

Table 4.5 JUN upstream regulator and its downstream targets.

Genes in dataset	Prediction (based on measurement direction)	Expr Log Ratio	Findings
JUN	Activated	10	Upregulates (84)
CSTA	Activated	10	Upregulates (2)
MYB	Activated	10	Upregulates (1)
ZEB2	Affected	2.829	Regulates (1)
ARG1	Inhibited	1.902	Downregulates (1)
RXRA	Activated	1.64	Upregulates (2)
BRCA1	Affected	0.686	Regulates (1)

3. EP300 (E1A-Binding Protein p300)

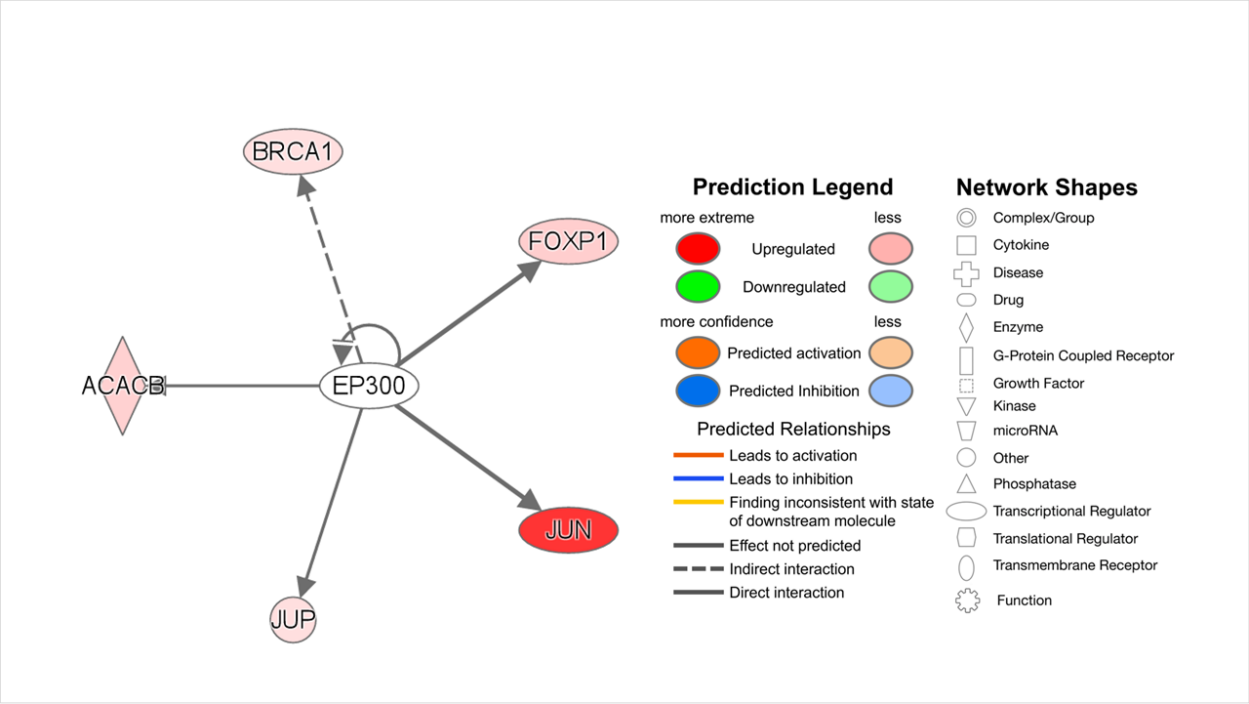


Figure 4.20 EP300 network of interactions.
This figure illustrates EP300 as a key regulatory molecule, interacting with multiple downstream targets. EP300 is connected to BRCA1, FOXP1, JUN, JUP, and ACACB, all of which are upregulated to varying degrees. JUN exhibits the most extreme upregulation, highlighting its potential role in transcriptional activation, cell proliferation, and stress responses. EP300 is a well-known transcriptional coactivator, and its involvement in these interactions suggests a regulatory function in gene expression, chromatin remodelling, and possibly oncogenic signalling pathways. These findings emphasise EP300's role in transcriptional regulation, cellular differentiation, and tumorigenesis, offering insights into its functional significance within this dataset.

Table 4.6 EP300 upstream regulator and its downstream targets.

Genes in dataset	Prediction (based on measurement direction)	Expr Log Ratio	Findings
JUN	Activated	10	Upregulates (6)
EP300	Activated	10	Upregulates (42)
FOXP1	Activated	2.33	Upregulates (1)
ACACB	Affected	2.231	Regulates (1)
JUP	Affected	1.038	Regulates (1)
BRCA1	Affected	0.686	Regulates (1)

4. CREBBP (cAMP Response Element-Binding Protein (CREB) Binding Protein)

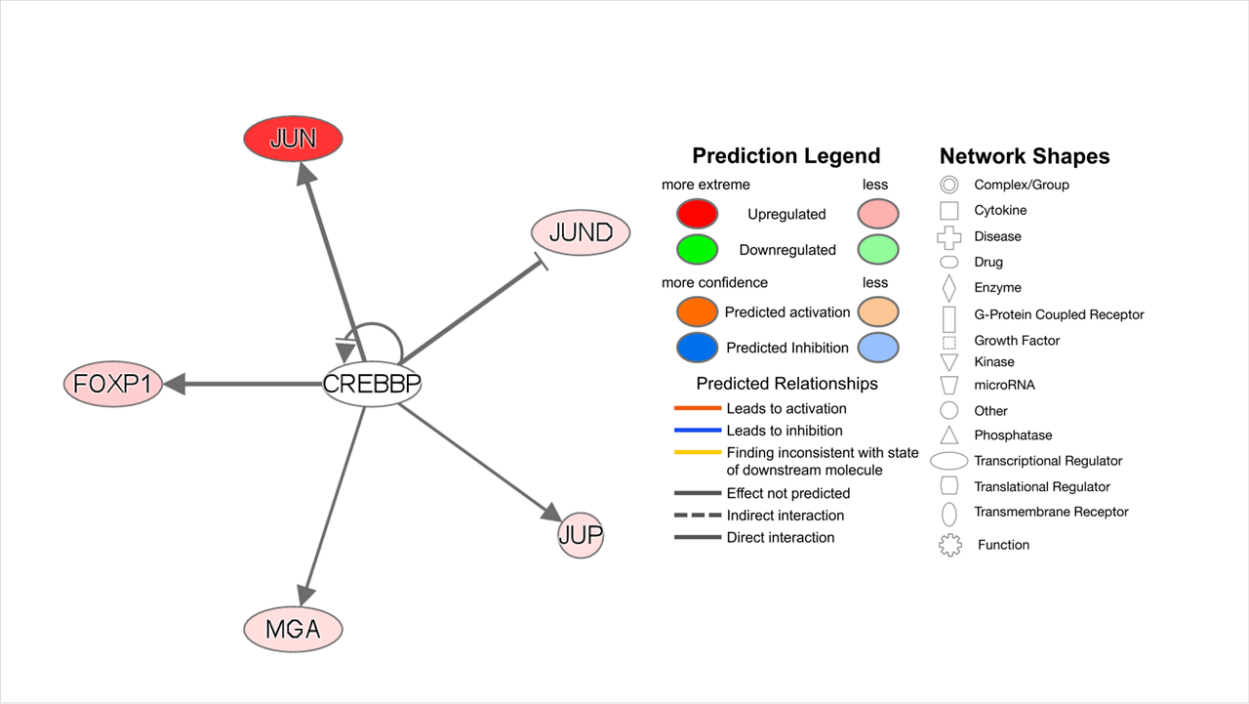


Figure 4.21 CREBBP network of interactions.
This network diagram illustrates the molecular interactions of CREBBP (CREB-binding protein) and its associated transcriptional regulators, highlighting its role as a key coactivator in gene regulation, chromatin remodelling, and transcriptional activation. CREBBP directly interacts with several transcriptional regulators, including JUN, which is strongly upregulated, as well as JUND, FOXP1, MGA, and JUP, which exhibit moderate upregulation. Additionally, CREBBP shows a self-regulatory interaction (loop), emphasizing its central role in transcriptional regulation. These interactions suggest that CREBBP may contribute to signalling pathways involving JUN family transcription factors, potentially influencing key regulatory networks.

Table 4.7 CREBBP upstream regulator.

Genes in dataset	Prediction (based on measurement direction)	Expr Log Ratio	Findings
JUN	Activated	10	Upregulates (2)
FOXP1	Activated	2.33	Upregulates (1)
MGA	Affected	1.617	Regulates (1)
JUND	Inhibited	1.345	Downregulates (1)
JUP	Affected	1.038	Regulates (1)

4.4.5.4 Mechanistic network

To explore potential connections or interactions between multiple upstream regulators, we conducted a Mechanistic Network Analysis. This analysis involved a statistical enrichment method aimed at uncovering significant relationships and overlaps between upstream regulators and the targets present within the experimental dataset.

The IPA has identified MYB and JUN as influential upstream regulators within the sample group that displays PLAG1 expression. In response, a detailed mechanistic network analysis has been constructed to provide a thorough understanding of the interconnected upstream regulators driving altered gene expression through downstream effector molecules (Figures 4.22 and 4.23).

MYB Mechanistic Network

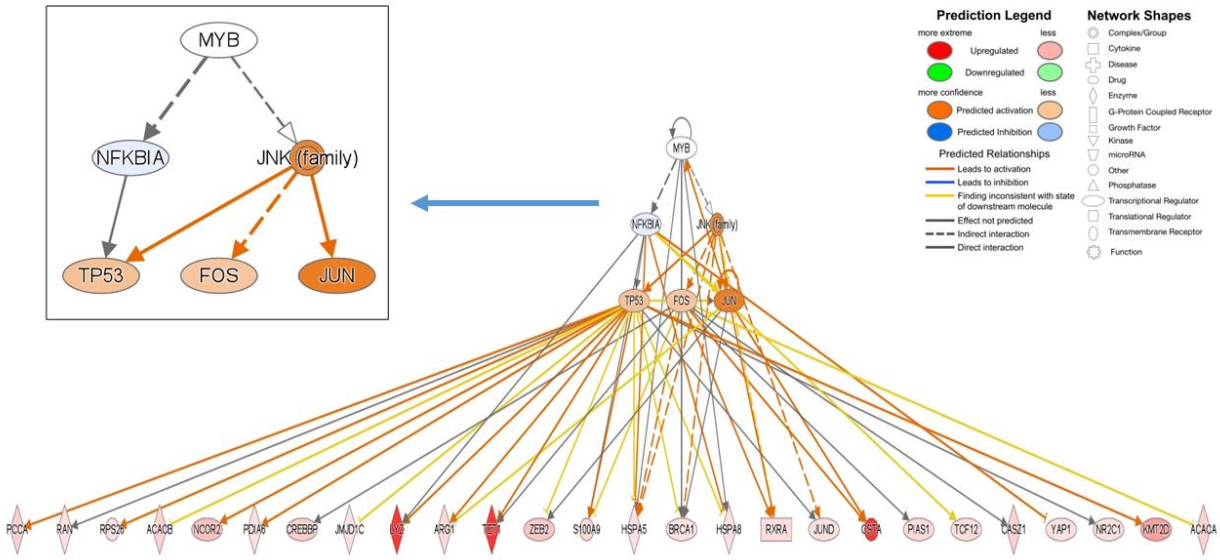


Figure 4.22 A schematic representation of the MYB complex mechanistic network.
The network illustrates the regulatory interactions and predicted activation/inhibition relationships in a sample group displaying PLAG1 expression. At the top of the network, MYB serves as a central transcriptional regulator, influencing key downstream targets such as TP53, FOS, and JUN through intermediates like NFKBIA and the JNK (family) signalling pathway. The colour coding indicates molecular expression changes, with upregulated nodes in red/orange and downregulated nodes in green, while the arrows represent predicted activation or inhibition relationships. The activation of JUN and FOS, alongside JNK signalling, suggests a proliferative or stress-response pathway, while TP53 modulation hints at potential tumour suppressor interactions. Given the presence of PLAG1 expression, MYB's activation may

contribute to transcriptional programs linked to cell proliferation, differentiation, or apoptosis resistance, depending on the cellular context.

The mechanistic network of MYB, detailing protein expression changes and predicted activation or inhibition relationships is outlined in Appendix Table 8

JUN Mechanistic Network

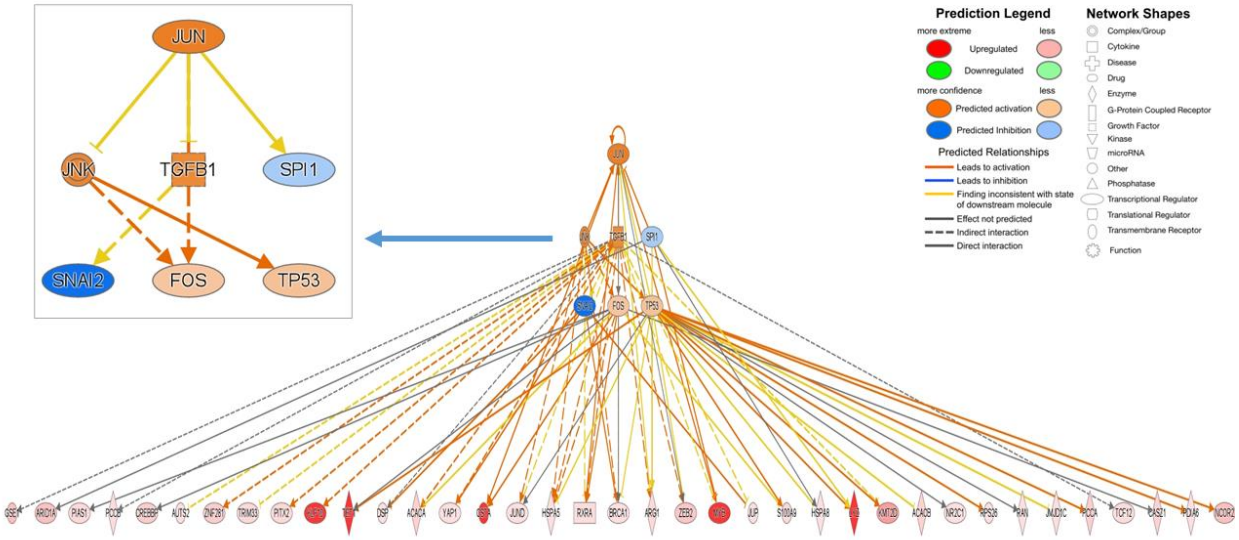


Figure 4.23 Schematic illustration of the JUN mechanistic network.

The JUN Mechanistic Network illustrates the predicted regulatory interactions and activation/inhibition relationships in a sample group where JUN plays a central role as a transcriptional regulator. Positioned at the top of the network, JUN influences key downstream targets, including FOS, TP53, and TGFB1, through intermediate regulators such as JNK and SPI1. The network colour coding indicates expression changes, with upregulated molecules shown in red/orange and downregulated ones in green, while different arrow types denote activation, inhibition, or indirect interactions. The activation of TGFB1 and JNK suggests involvement in signalling pathways related to proliferation, differentiation, or stress responses. Additionally, the inhibition of SNAI2 and modulation of SPI1 indicate a regulatory balance between transcriptional activation and repression. This network highlights JUN's role in transcriptional control, potentially driving cell fate decisions and stress-adaptive responses depending on the biological context.

Appendix Table 8.7 provides an overview of the JUN mechanistic network, including expression data and the predicted regulatory interactions of JUN with key proteins.

4.4.5.5 Affected Diseases and Biological Functions

A Downstream Effects Analysis was conducted to better understand biological trends and predict how molecular changes in the experimental dataset may impact disease and biological functions. This analysis enabled us to visualise these trends quickly and facilitated a better understanding of these effects. The heatmap in Figure 4.24 provides a general overview of the affected diseases and functional categories. The study of the DEPs showed that the most significant enrichment in diseases and biological functional categories is cancer and gene expression respectively. Tables 4.8 and 4.9 summarise the top five most significant diseases and functions.

Table 4.8 Top 5 diseases and disorders.

Diseases and disorders	p-value range	Number of molecules
Cancer	3.20E-03 - 5.69E-18	107
Organismal Injury and Abnormalities	3.20E-03 - 5.69E-18	107
Developmental Disorders	2.80E-03 - 1.48E-17	62
Hereditary Disorders	3.19E-03 - 1.48E-17	55
Neurological Diseases	3.08E-03 - 1.48E-17	91

Table 4.9 Top Molecular and Cellular Functions.

Molecular and Cellular Functions	p-value range	Number of Molecules
Gene Expression	2.26E-03 - 3.49E-34	76
Cell Cycle	3.19E-03 - 1.99E-16	55
Cellular Assembly and Organisation	2.94E-03 - 2.58E-12	24
DNA Replication, Recombination and Repair	2.94E-03 - 2.58E-12	27
Cell Death and Survival	2.01E-03 - 1.31E-10	66

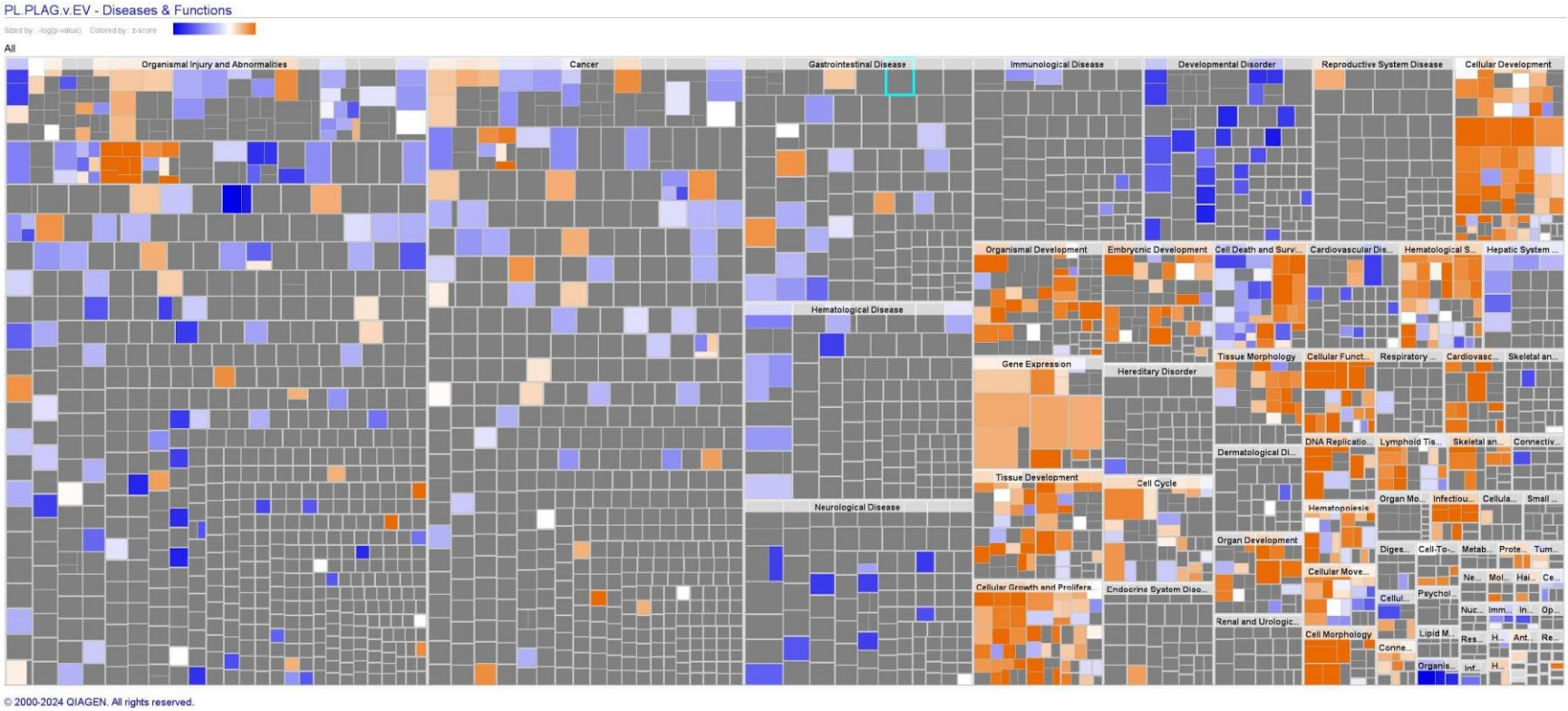


Figure 4.24 Diseases and Biological Functions Heatmap.
The treemap (hierarchical heat map) illustrates the affected downstream functional groups based on differentially expressed proteins, with each main box denoting a specific category of diseases and functions associated with PLAG1 ectopic expression. The size of the rectangles corresponds to the increasing overlap significance, either up or down as a group, as determined by the Fisher exact test (FET) p-value, while the intensity of colour reflects higher absolute Z-scores

Functions known to be affected in pleomorphic adenoma

The downstream effects analysis revealed an increase (activation) of functional categories such as Cellular Development and Cellular Growth and Proliferation, as depicted in Figures 4.25 and 4.26. These categories encompass cellular processes related to cell proliferation, differentiation, and colony formation. Additionally, a notable enrichment in the cell death and survival category was observed, supported by statistical significance (Figure 4.27). It is important to note that this category comprises cellular processes with opposing activation states; while cell viability, survival, and self-renewal were enhanced, apoptosis was decreased.

The bubble plot shown in Figure 4.28 summarises the most notably increased and decreased biological functions that align with the biological context of pleomorphic adenoma documented in the literature.

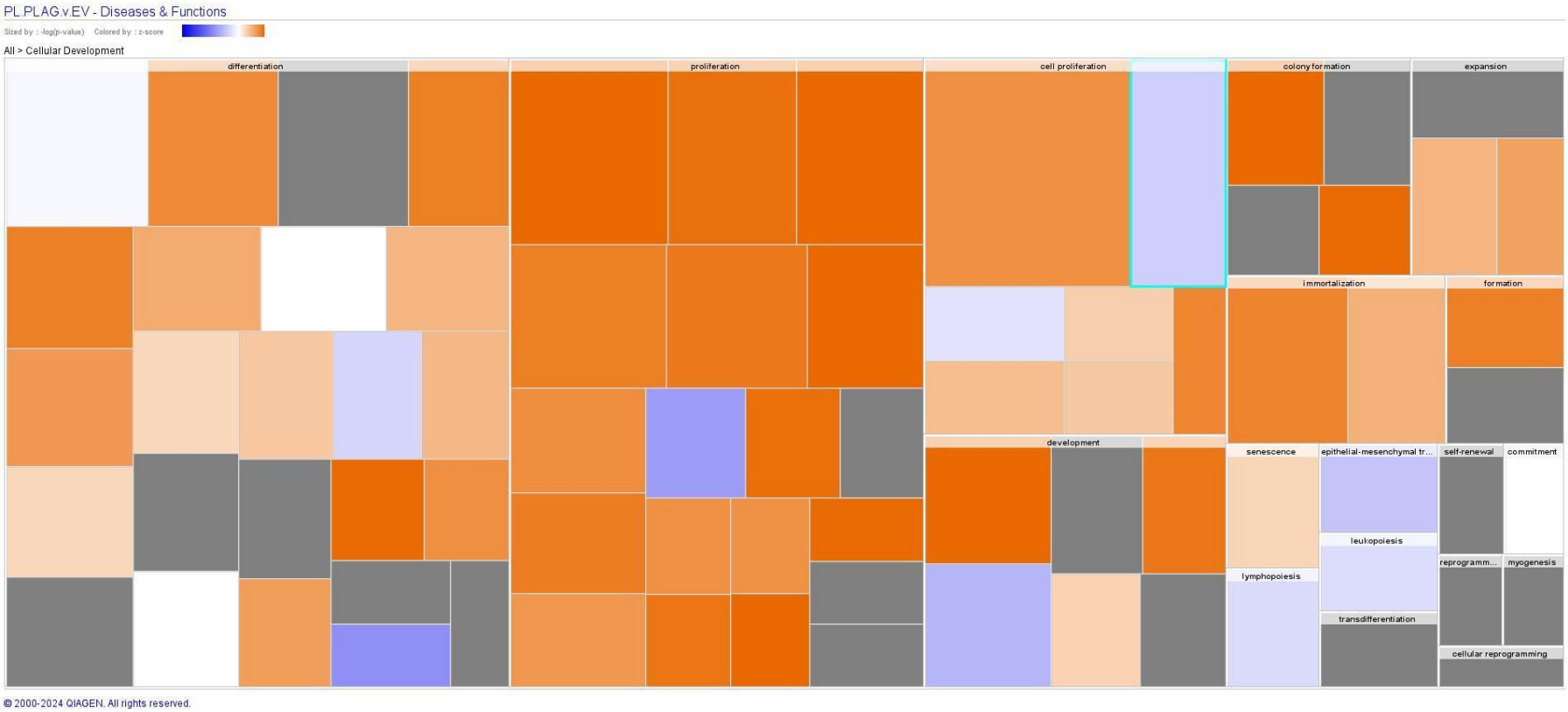


Figure 4.25 Heatmap of cellular development.
This heatmap represents cellular development functions, highlighting key processes such as cell proliferation, differentiation, and colony formation. The colour gradient reflects the z-score, with orange shades indicating activation and blue shades suggesting inhibition of specific cellular functions. Larger blocks correspond to higher -log (p-value), representing stronger statistical significance of the functional associations. Notably, proliferation and differentiation-related processes are predominantly activated, suggesting a strong regulatory influence in promoting cell growth and specialisation. The activation of colony formation aligns with other proliferation-related functions, reinforcing the overall theme of enhanced cell growth and expansion. This trend may indicate a biological environment favouring cell survival, anchorage-independent growth, or tumorigenic potential.

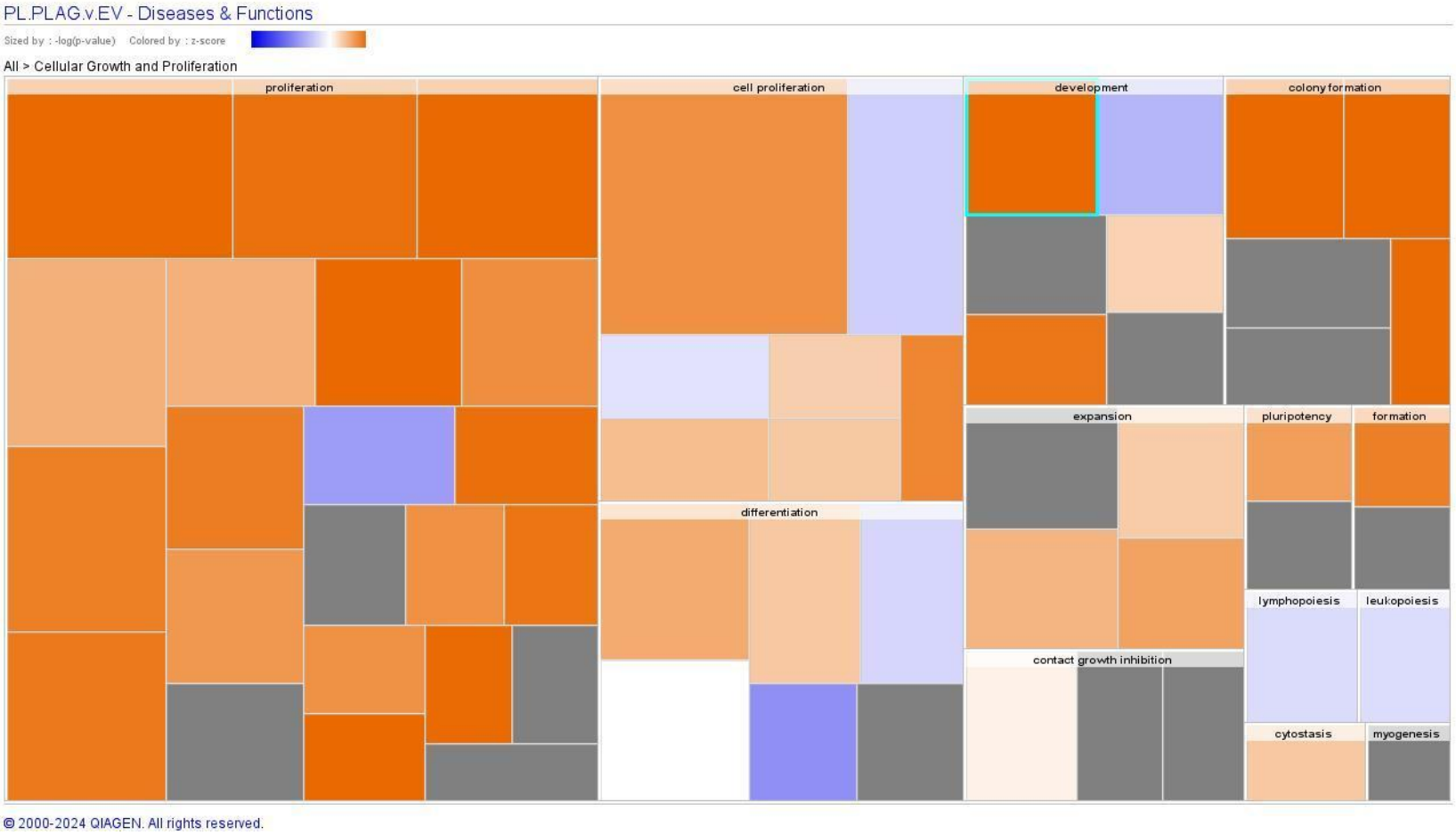


Figure 4.26 Cell Growth and Proliferation Map.

The Cell Growth and Proliferation Treemap visualises diseases and functions associated with PLAG1 vs EV analysis, categorising biological processes based on statistical significance and effect direction. Key biological categories include Proliferation & Growth (proliferation, cell proliferation, colony formation, expansion), Development & Differentiation (development, differentiation, pluripotency), and Hematopoiesis & Blood Cell Processes (lymphopoiesis, leukopoiesis).

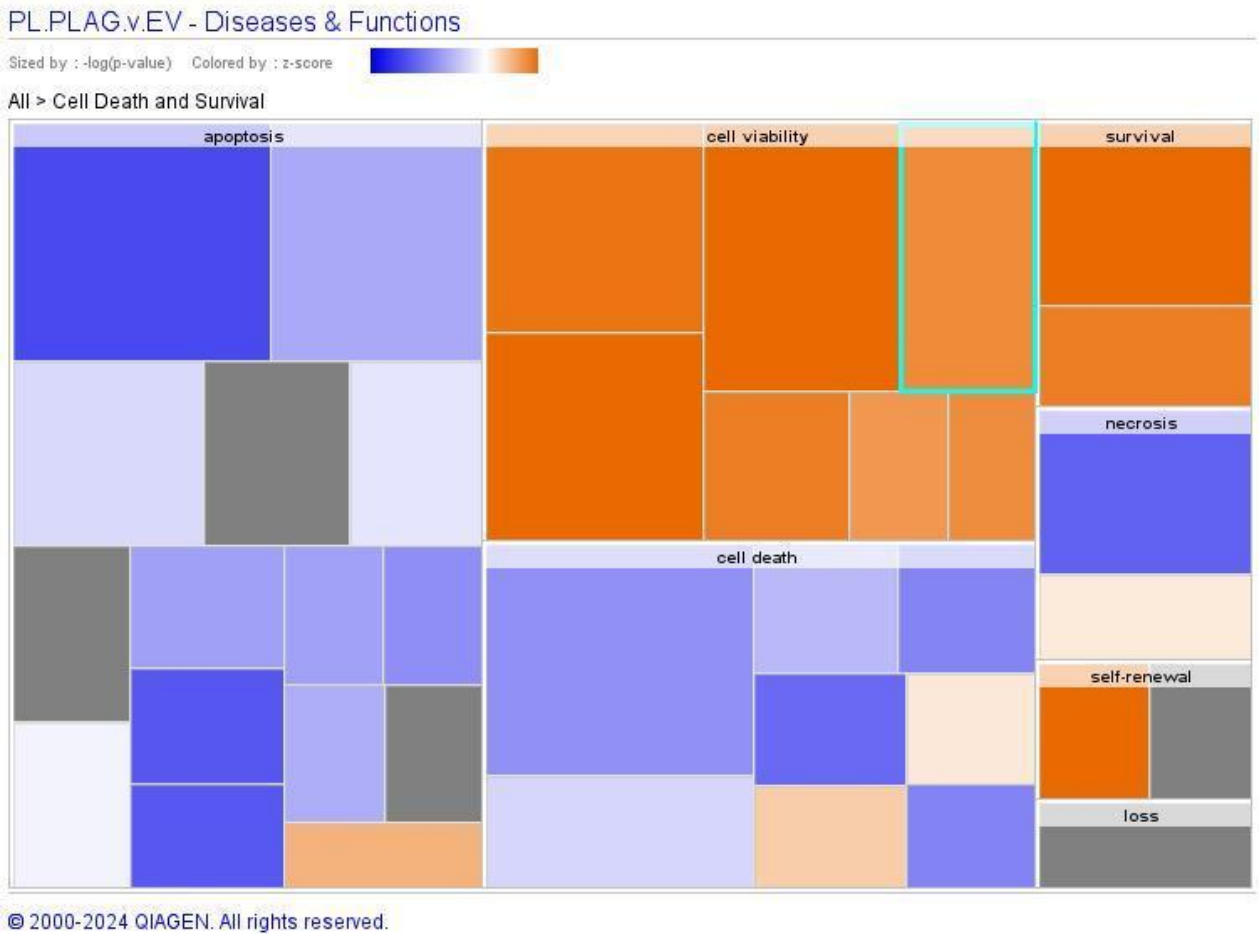


Figure 4.27 Cell Death and Survival heatmap.
The treemap visualisation highlights key biological functions related to cell death and survival, with colour-coded activation (orange) and inhibition (blue) based on z-scores. Activated functions include cell viability, survival, and self-renewal, indicating a shift toward cellular maintenance and proliferation. The upregulation of these processes suggests a protective or proliferative response, potentially linked to enhanced cell growth, tissue regeneration, or resistance to stress-induced damage. Conversely, inhibited functions such as apoptosis, necrosis, and cell death reflect suppression of programmed and uncontrolled cell death mechanisms.

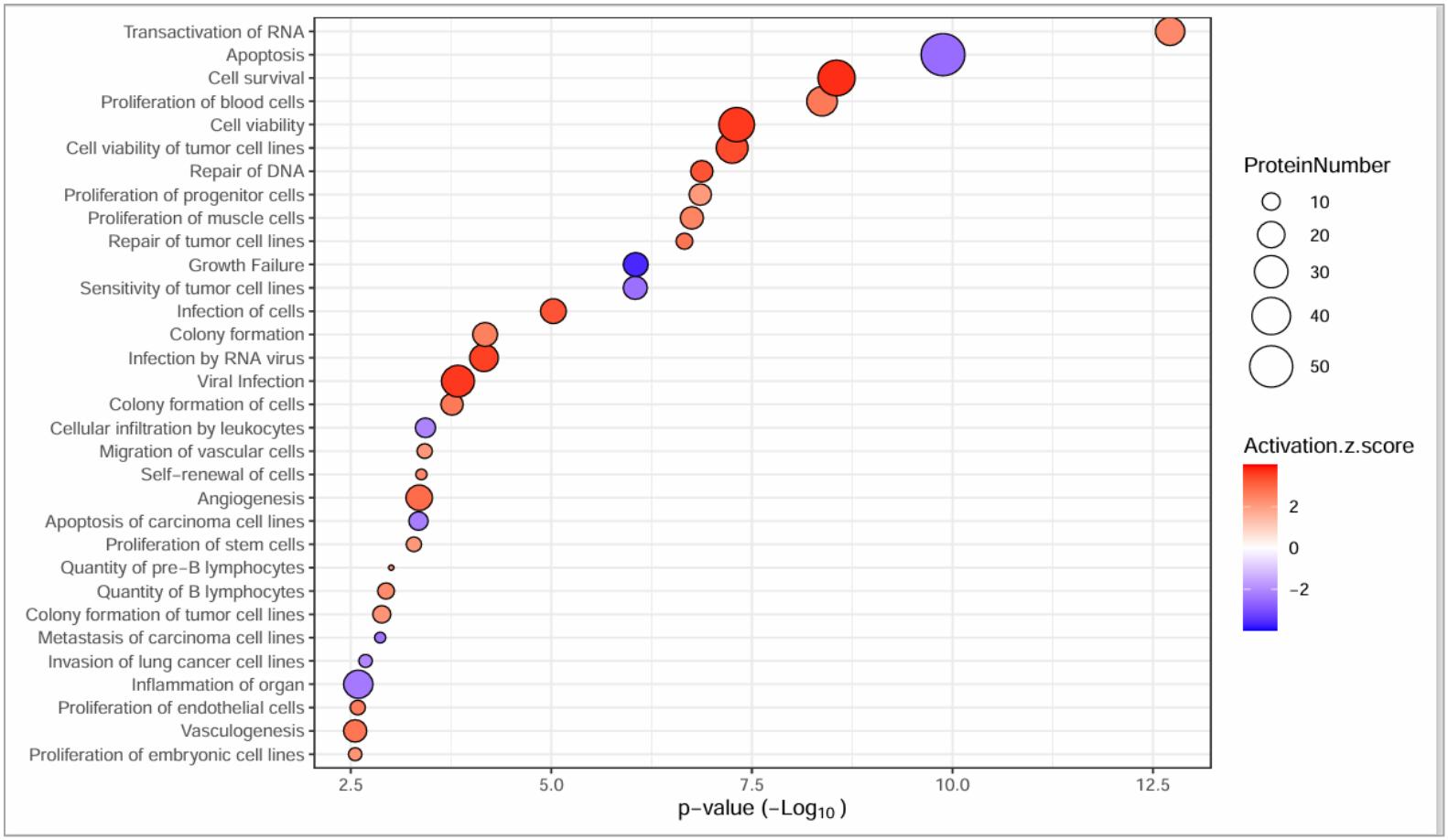


Figure 4.28 Most significantly activated and inhibited biological functions.
The bubble plot illustrates the most significantly upregulated functions (red bubbles) and downregulated or suppressed functions (blue bubbles) based on their activation Z-score. The x-axis represents the statistical significance of each function, measured as $-Log_{10}$(p-value), where higher values indicate stronger significance. The y-axis lists the analysed biological functions. The size of each bubble corresponds to the number of genes involved in the respective function, with larger bubbles indicating a greater number of associated genes.

This analysis reveals a shift toward cell survival, proliferation, and repair, while apoptosis and infection-related processes are suppressed. Among the most activated functions are cell viability, proliferation of blood and progenitor cells, and DNA repair, suggesting enhanced regenerative capacity

or potential tumour progression. Additionally, angiogenesis and vasculogenesis are upregulated, indicating increased blood vessel formation, which is crucial for tissue growth and metastasis. Conversely, apoptosis is strongly inhibited, reflecting a reduction in programmed cell death, which may contribute to enhanced cell survival, particularly in cancerous cells. The suppression of growth failure and infection-related pathways, including viral infection and immune infiltration, suggests a possible shift toward immune evasion or reduced susceptibility to external stimuli.

Overall, these findings highlight a pro-survival and proliferative cellular environment with potential implications for cancer progression, tissue regeneration, and therapy resistance.

4.4.5.6 Regulator Effect Analysis

This analysis was conducted to elucidate the biological implications of upstream molecules (whether activated, inhibited or overexpressed) on downstream phenotypic or functional outcomes. This was accomplished by establishing connections between the upstream regulators and their targets in the dataset, as well as the consequential downstream affected functions. The analysis enables the formulation of a causal hypothesis to delineate the mechanisms through which the upstream regulators influence downstream functions.

The IPA identified several molecules as top (master) upstream regulators that affect the DEPs in the dataset. It then created regulator effect networks by integrating the results of the upstream regulators with their respective downstream effects. These master upstream regulators include growth factors (IGF1 and EGF), cytokines (IL6, IL33), enzymes (SNCA), compounds (ERK1/2), transcription regulators (HNF4A), and microRNA (mir-8). Based on their activation Z-scores, the analysis suggests that the genes SNCA, HNF4A, and mir-8 are expected to be inhibited, while the other upstream regulators are anticipated to be activated.

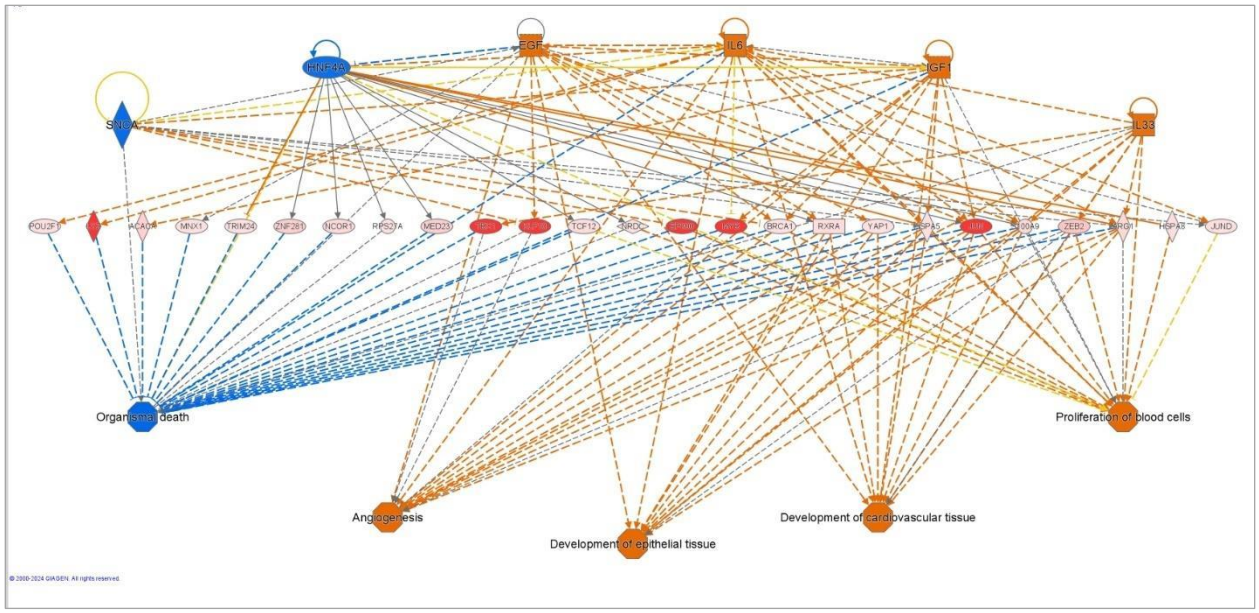


Figure 4.29 Regulator effect network of SNA, HNF4A, EGF, IL6, IFG1 and IL6.
The figure illustrates a complex network of interactions involving six master regulators (SNCA, HNF4A, EGF, IL6, IFG1, and IL6) and their respective target proteins in the dataset. This interaction activated several functions while also inhibiting organismal death, likely by promoting cell survival. Red lines represent activation effects, while blue lines indicate inhibition effects. The network highlights key regulatory connections, showing how these master regulators influence downstream targets to modulate biological functions.

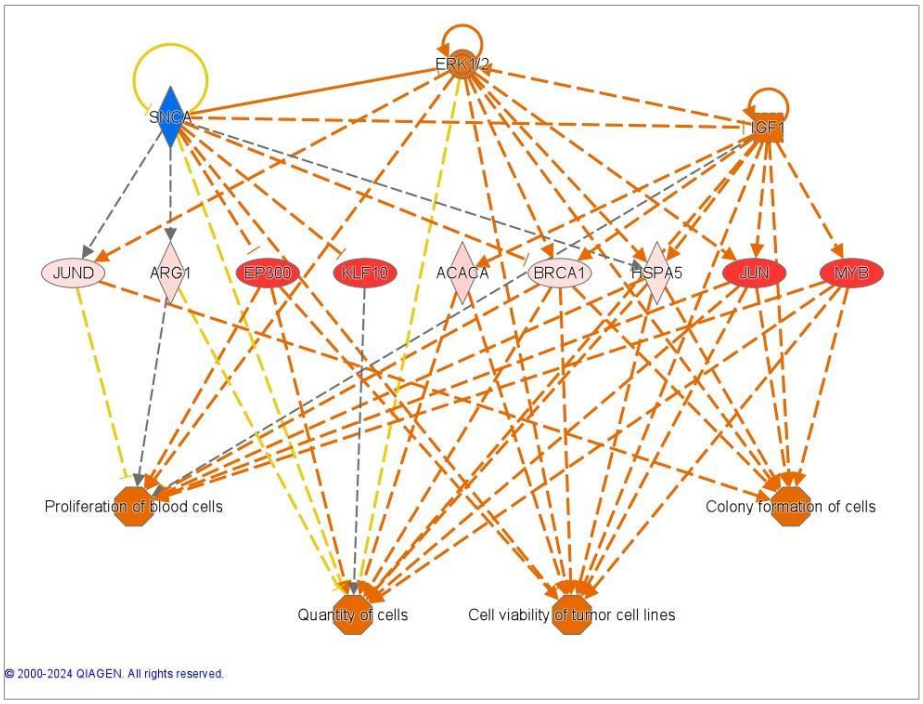


Figure 4.30 Regulator effect network of SNCA, ERK1/2 and IF1.
The interaction and overlap of these three upstream regulators influence nine target proteins, leading to observed changes in four tumour-related functions.

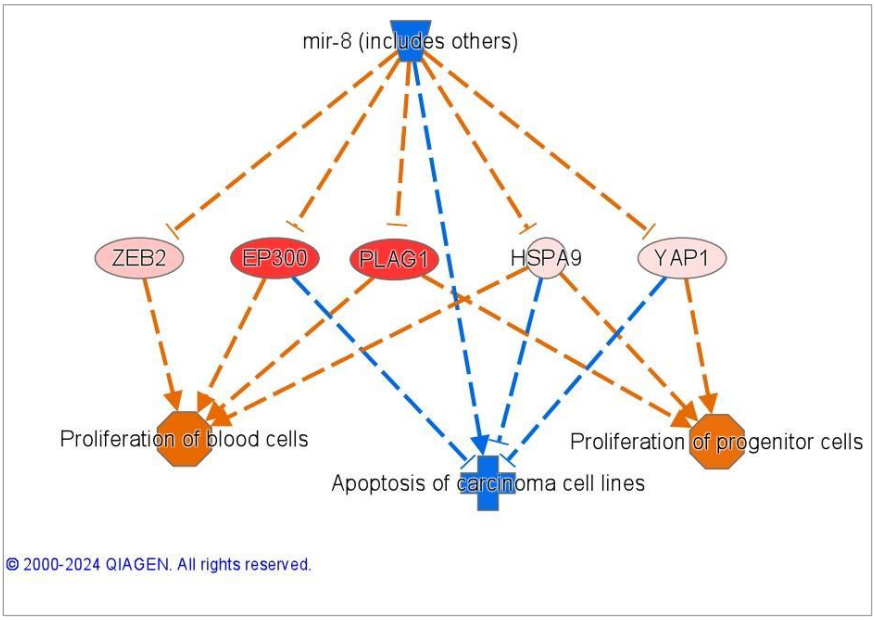


Figure 4.31 Regulator effect Network of mir-8 upstream molecule.
This network demonstrates the inhibitory effect of mir-8 on downstream targets in the dataset, including PLAG1 and their respective affected functions.

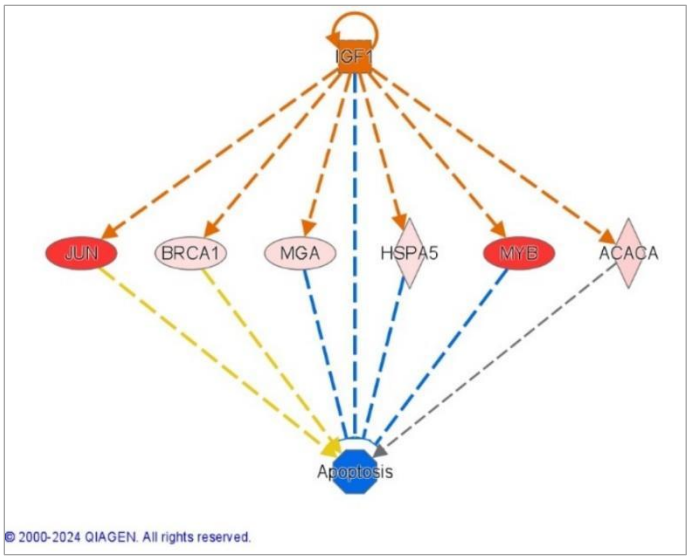


Figure 4.32 IGF 1 regulator effect network.
The network diagram visually represents the impact of IFG1 on the inhibition of apoptosis.

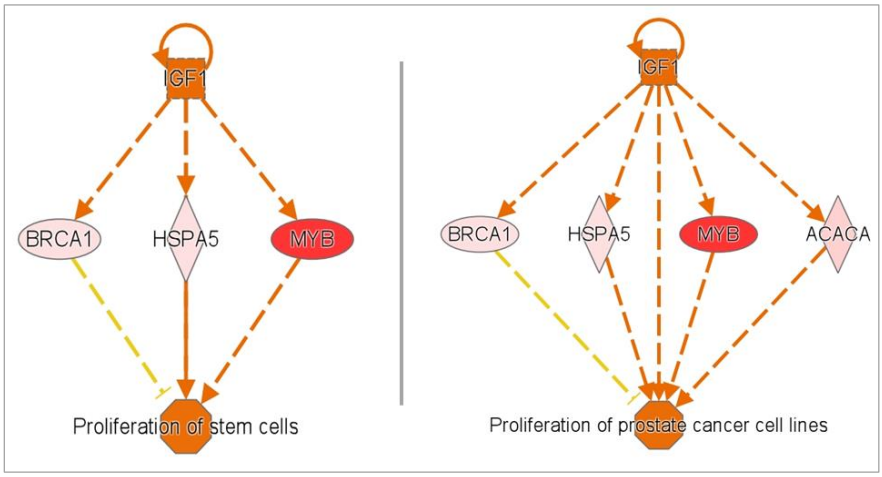


Figure 4.33 Regulator effect networks of IGF 1.
The activation of IGF1 promotes the proliferation of stem cells (left) and the proliferation of cancer cell lines (right) by targeting target proteins in the dataset.

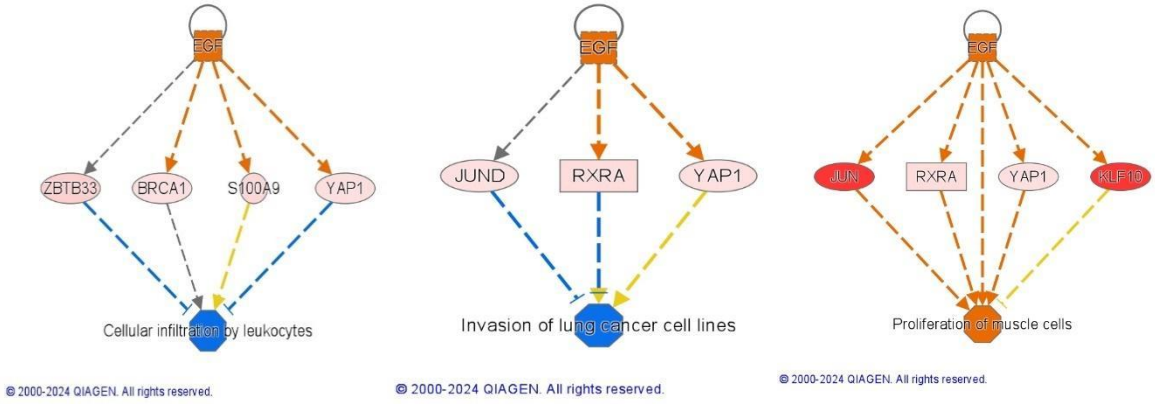


Figure 4.34 EGF regulator effect networks.
The networks portray the individual pathways driven by EGF, which inhibit cellular infiltration by leukocytes and invasion of cancer cell lines while promoting the proliferation of muscle cells.

4.4.6 Validation of the candidate proteins using open-source tools

To further validate and visualise the network of PLAG1-proximal interactors identified as upstream regulators, we used web-based free online tools to confirm the relationship between these molecules. Networks of protein-protein interactions provide the fundamentals for understanding biological processes and the molecular mechanisms involved in disease pathogenesis. Figures 4.35 and 4.36 illustrate the interconnections between the most significant biological upstream regulators as well as the four potential candidate genes for validation obtained from the STRING (Search Tool for the Retrieval of Interacting Genes/Proteins) database. Notably, no direct interactions with PLAG1 have been identified. This suggests that previous experimental studies may not have examined PLAG1 proximal interactions in the context of pleomorphic adenoma.

Additionally, we used SIGNOR (the SIGnaling Network Open Resource) to further investigate and capture any documented interactions between our proteins of interest. We were able to detect two direct interactions among the set of proteins (Figure 4.37).

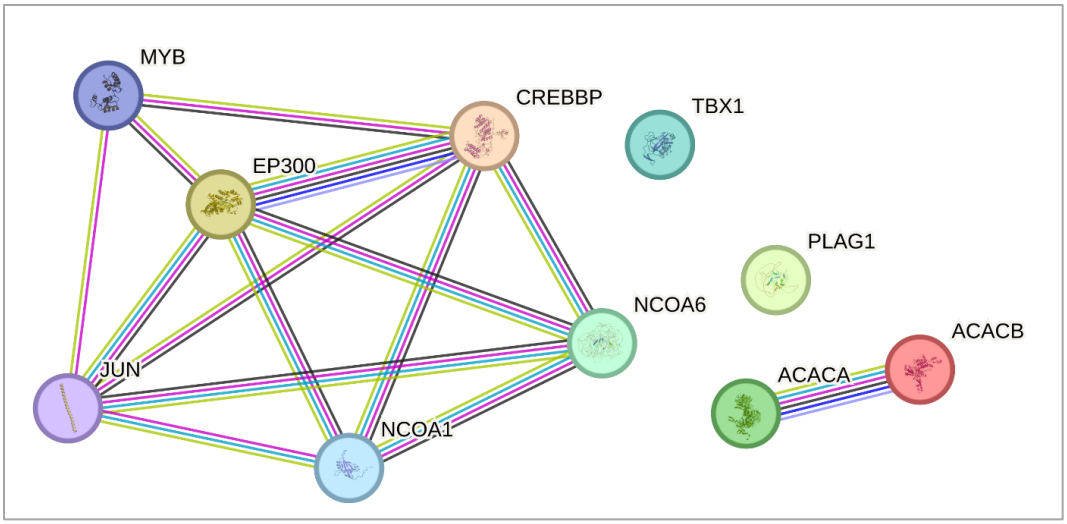


Figure 4.35 Protein-protein interactions of the biological upstream regulators.
The network depicts the relationships among key biological upstream regulators, highlighting distinct connectivity patterns. MYB, JUN, EP300, CREBBP, NCOA1, and NCOA6 exhibit strong interactions, indicating their cooperative roles in transcriptional regulation and chromatin remodelling. These interactions, supported by multiple evidence sources, suggest functional coordination in gene expression control. In contrast, TBX1 shows no detected interactions, suggesting it may function independently or within a different regulatory context. No direct interactions with PLAG1 have been identified, which suggests that previous experimental studies may not have examined PLAG1 proximal interactions in the context of pleomorphic adenoma. ACACA and ACACB display a strong interaction with each other but lack connections to the core regulatory network, implying a more specific role in metabolic pathways rather than transcriptional regulation. The distinct separation of these molecular groups emphasises the functional divergence between transcriptional regulators and metabolic enzymes in cellular processes.

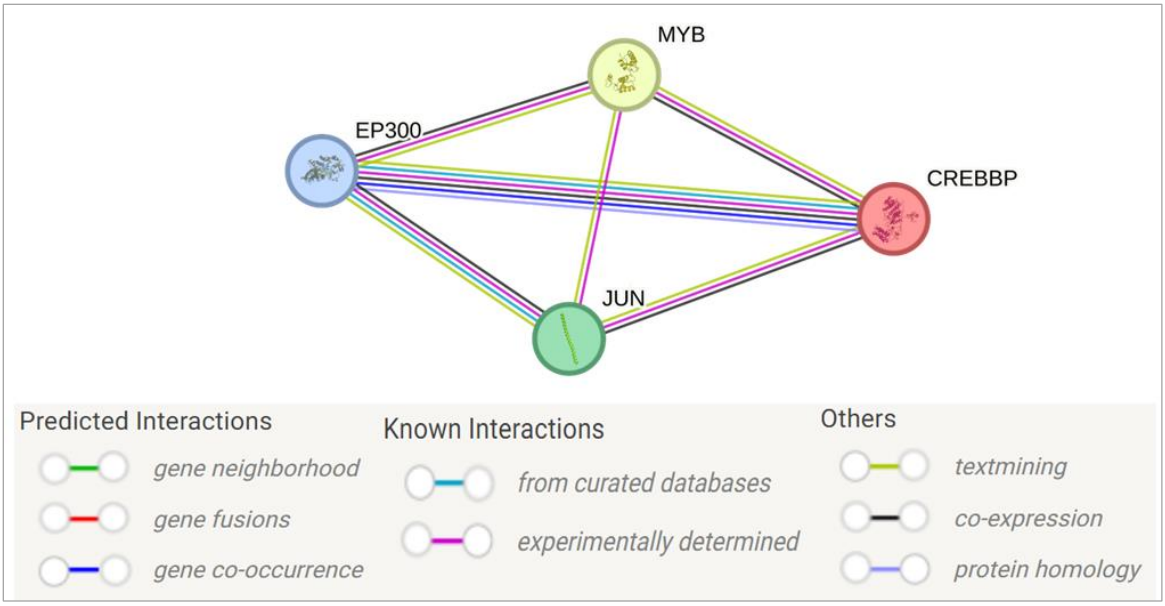


Figure 4.36 Interaction network of candidate proteins MYB, JUN, EP300 and CREBBP.
This interaction network illustrates the predicted and known relationships among MYB, JUN, EP300, and CREBBP, key regulators involved in transcriptional control and chromatin remodelling. The edges between nodes represent different types of interactions: predicted interactions include gene neighbourhood (green), gene fusions (red), and gene co-occurrence (blue), while known interactions are derived from curated databases (light pink) and experimental evidence (dark pink). Additional associations include text mining (yellow), co-expression (black), and protein homology (light blue). The strong connectivity among these proteins suggests their cooperative role in gene regulation, with MYB and JUN functioning as transcription factors and EP300 and CREBBP acting as histone acetyltransferases to modulate chromatin structure. These interactions emphasize their potential involvement in key biological processes, including cell proliferation, differentiation, and oncogenic signalling.

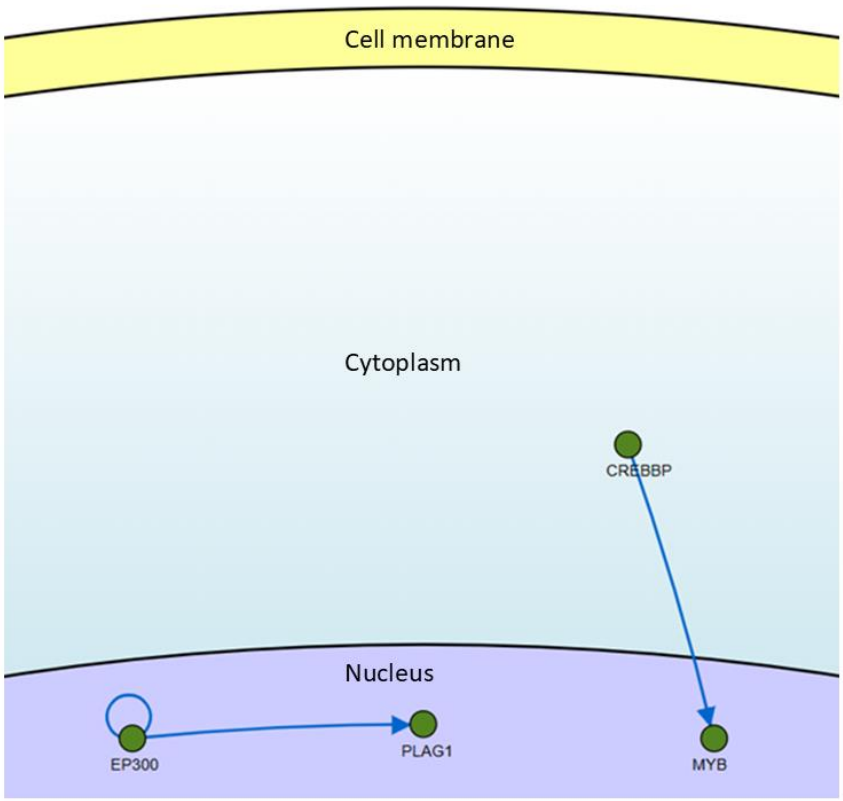


Figure 4.37 Direct protein interactions.
This diagram illustrates the predicted physical and functional interactions between EP300, PLAG1, MYB, and CREBBP, as identified using the SIGNOR (Signaling Network Open Resource) database. According to SIGNOR, MYB and CREBBP interact directly, suggesting a functional relationship in transcriptional regulation. Additionally, EP300 is capable of activating PLAG1 through acetylation, potentially influencing its transcriptional activity. Given the roles of these proteins in gene regulation, their interactions suggest potential regulatory mechanisms that may impact cellular processes such as proliferation, differentiation, or oncogenesis. The localisation of these proteins within the nucleus and cytoplasm highlights their compartment-specific functions in gene expression modulation.

4.5 Discussion

Biotin-based proximity labelling proteomic analysis was performed to gain a deeper insight into PLAG1's mechanisms for developing pleomorphic adenoma of the salivary glands.

Following mass spectrometry and raw data processing, multiple testing correction (FDR) was applied to determine the statistical significance of the quantitative data. However, only nine proteins were found to be FDR significant, indicating perhaps that this statistical method is too stringent. In the field of proteomics, there is an ongoing debate about the appropriate use of multiple testing corrections. While these corrections are important for avoiding false positives (type I errors), they can also make the analysis prone to false negatives (type II errors) if not applied effectively. Some challenges in using multiple corrections in proteomic studies include the small sample sizes due to cost constraints and the impact of TMT tagging, which compresses ratios and reduces effect sizes (not relevant for this current analysis). These issues are not as prevalent in other types of studies, such as microarray experiments, where multiple correction analysis is commonly and effectively used. The analysis of simulated proteomic datasets has shown that these characteristics can lead to multiple testing corrections producing a large number of false negative results. This indicates that while multiple testing corrections are a useful tool, they should not be considered an automatic requirement and should be viewed as one of several available options (Pascovici et al., 2016).

In determining whether to employ multiple correction techniques, our choice largely relies on whether we prioritise avoiding type I or type II errors. According to some researchers, there has been an overemphasis on minimising false positives, and they advocate for giving greater consideration to false negatives. They argue that while false positives can be further investigated and possibly eliminated later, a false negative may go undetected if an experiment is not repeated, leading to missed true effects. In the type of study, we present in this thesis, where we are conducting an initial broad exploration of proteomics data to pinpoint areas for more detailed investigation in future work, we believe it is more crucial to minimise false negatives. This viewpoint is not uncommon in the field, as only a minority of published proteomics papers currently apply multiple-testing correction (Diz et al., 2011; Handler and Haynes, 2020).

Therefore, the statistical significance of the quantitative data was adjusted using a significance cut-off of p-value < 0.05 to compare the conditions of interest in the experiments.

Only proteins demonstrating upregulation with real values (proteins that show an actual increase in expression or enrichment with quantifiable, non-missing values) and meeting the specified cut-off criteria were considered for the IPA analysis. In order for a protein to be included in the analysis, it must be expressed in a minimum of three PLAG1-expressing samples as compared to no expression in any EV samples: 3 v 0, 4 v 1 or 4 v 0. The latter condition represents the ideal inclusion criteria, indicating that the protein is expressed in all PLAG1-expressing samples while showing no expression in any of the control samples. These proteins were exclusively quantified in the condition of interest were given a p-value of 0, and a LogFC of 10, and all LogFCs below 0 were removed from the dataset, prior to upload to IPA.

Canonical Pathway Analysis

IPA Canonical pathway analysis was conducted to identify the most significantly affected pathways across the entire dataset and to predict their activation or inhibition. The software evaluates the statistical significance of pathway overlap using the p-value derived from the right-tailed Fisher's Exact Test. This approach enables the identification of pathways that exhibit significant associations. Furthermore, the software leverages the Z-score to predict the likely state of activation or inhibition within these pathways. This dual methodology fosters a comprehensive understanding of pathway dynamics, thereby allowing for more informed analyses and decision-making based on the activation or inhibition status of the pathways. The analysis also facilitated the identification of overlapping pathways based on shared molecules.

Our study analysis revealed that the Activation of Anterior HOX genes and Transcriptional Regulation by RUNX1 emerge as the most notably enriched pathways across the dataset, displaying positive activity patterns. In contrast, the PPARα/RXRα Activation and Sumoylation Pathways showed clear signs of inhibition, as evidenced by their negative z-scores of -1.89 and -1.63, respectively. Notably, while the Glucocorticoid Receptor Signalling is among the significantly enriched pathways, the software did not provide an activity prediction. These pathways, particularly HOX genes, engage in intricate interaction with various signalling pathways, including

Wnt, Notch, Hedgehog, Wnt/β-catenin signalling, and the AKT pathway (Shenoy et al., 2022; Gonzalez and Medici, 2014). These complex interactions have a discernible impact on the behaviour of cancer cells and their response to therapeutic interventions, underscoring the capacity of the pathways to regulate a diverse array of targets within the context of cancer. Furthermore, the analysis of overlapping canonical pathways revealed the intricate interplay observed among the most significantly altered pathways identified in this study. It is noteworthy that HOX genes demonstrate the most prominent crosstalk within these overlapping pathways, with CREBBP and EP300 being the most commonly shared proteins across the pathways. However, there have been limited studies examining the correlation between these affected pathways and PLAG1, the causative factor of pleomorphic adenoma, thus rendering this study pioneering.

Activation of anterior HOX genes

The HOX genes, a complex group of genes, are initially expressed in the developing embryo during the gastrulation stage (Boncinelli and Mallamaci, 1995). These genes function as transcription factors, inducing the expression of genes involved in various cellular processes, such as proliferation, adhesion, differentiation, migration, and programmed cell death, within the context of normal physiological conditions through their homeobox domain. (Hombría and Lovegrove, 2003; Moens and Selleri, 2006). The deregulation of HOX genes at various stages of development leads to abnormalities in skeletal morphology such as hand-foot-genital syndrome, syndactyly, Guttmacher syndrome, congenital heart diseases, retinal degenerative diseases, and cancer (Lescroart and Zaffran, 2018; Quinonez and Innis, 2014; Shah and Sukumar, 2010; Zagozewski et al., 2014). Accumulating evidence indicates that homeobox (HOX) genes, serving as master regulators during embryogenesis, are involved in both organogenesis and oncogenesis (Smith et al., 2019). The ectopic expression of HOX genes disrupts the normal intracellular signalling process, thereby inducing cellular dysfunction and compromising tumour immune surveillance. Sustained activation of key signalling pathways controlling cell growth, the cell cycle, and cell proliferation establishes an environment conducive to tumour development and progression. (Giancotti, 2014). Researchers have conducted extensive reviews on the involvement of HOX genes in both normal and tumour microenvironments (Bhatlekar et al., 2018). Notably, the

regulation of HOX genes is known to be tissue-specific, and their dysregulation exhibits variation across different cancer types. These genes demonstrate dual functionality as either oncogenes or tumour suppressors, contingent upon the specific context. HOX genes significantly influence cancer cells' behaviour by regulating various cellular processes, encompassing cell proliferation, differentiation, epithelial-mesenchymal transition (EMT), metastasis, invasion, angiogenesis, apoptosis, drug resistance, and stem cell properties.

In a study conducted in 2024, Padam et al. investigated the regulatory role of HOX and mutated cancer driver genes in the progression of head and neck cancers, specifically squamous cell carcinoma. The study aimed to demonstrate the crucial regulation of mutation-driven effects on homeobox genes, expanding the current understanding of this cancer's progression. The investigation revealed that the expression of HOX genes is influenced by various transcription factors (TF), including histone acetyltransferase (EP300), cAMP response element-binding protein (CREBBP), nuclear receptor corepressor (NCOR1), and lysine demethylase 6A (KDM6A), all of which exhibit significant enrichment within the dataset encompassed by this study.

Furthermore, Cantile et al. (2009) conducted a study on the expression of the HOX gene, specifically HOXD13, using immunohistochemistry. The research compared the expression of HOXD13 in normal tissues and its potential deregulation during the development of specific tumour types. They employed tissue microarrays (TMAs) containing over 4,000 different normal tissue and cancer samples from 79 different tumour categories. HOXD13 expression was further confirmed at the mRNA level using real-time quantification. The analysis showed that HOXD13 expression tested positive in 1,584 out of 2,743 (57.7%) of the valuable neoplastic tissue samples included in the study. Within the context of salivary gland tumours, the study revealed a significant increase in HOXD13 expression in pleomorphic adenomas and adenoid cystic carcinomas (ACC), with percentages of 97% and 90%, respectively, compared to normal submandibular gland tissue samples. This finding was consistent with the HOXD13 mRNA expression in the two groups being compared.

Notably, HOX genes interact with various signalling pathways, including Wnt, Notch, Hedgehog, Wnt/β-catenin signalling, and the AKT pathway. These interactions have a discernible impact on

the behaviour of cancer cells and their response to therapeutic interventions. The oncogenic implications of HOX genes are demonstrated by their coordination of multiple signalling pathways, underscoring their capacity to regulate a diverse array of targets within the context of cancer (Shenoy et al., 2022). Figure 4.16 illustrates the intricate interplay observed among the most significantly altered canonical pathways identified in this study. Notably, HOX genes exhibit the most prominent crosstalk within these overlapping pathways.

Transcriptional regulation by RUNX1

The Runt-related transcription factor1 (RUNX1) is a member of the RUNX family of transcription factors and is known to influence cancer pathogenesis. It modulates various biological processes, including cell proliferation, apoptosis, differentiation, and lineage determination (Chuang et al., 2013; Ito et al., 2015). RUNX1 has the capacity to activate gene expression by recruiting co-activators, RNA polymerase, and other transcriptional machinery to the gene promoter. Conversely, it can repress gene expression by competing with other transcription factors or recruiting co-repressors. The impact of RUNX1 on gene expression is contingent upon the specific context. Depending on the cellular environment and interacting partners, it can function as an activator or repressor (Guan et al., 2023).

The transcription factor RUNX1 is notably regarded as a master regulator of haematopoiesis (Sroczynska et al., 2009). It plays a pivotal role in developing hematopoietic stem cells, megakaryocytes, and various blood cell lineages. Its significance extends to both normal and malignant hematopoietic processes. In the context of cancer, the dysregulation of RUNX1 is closely linked to the activation of oncogenic signalling pathways such as TGF-β, Wnt/β-catenin and Notch signalling pathways (Gonzalez and Medici, 2014), indicating its potential as a target for therapy (Tuo et al., 2022). The prevalence of RUNX1 overexpression is widespread across various human malignancies, including breast carcinomas (Ferrari et al., 2014), acute myeloid leukaemia AML (Chuang et al., 2013) and ovarian carcinomas (Keita et al., 2013) and correlates with patient prognosis. RUNX1 is also recognised as an oncogene in epithelial skin cells owing to its regulatory role in tissue stem cells. (Hoi et al., 2010). The intricate involvement of RUNX proteins in both

leukaemia and solid cancers underscores their complexity and importance in the field of cancer research.

Dysregulation of RUNX can result in genomic instability by impairing DNA repair mechanisms (Chen et al., 2024). Additionally, the expression of RUNX1 shows a clear correlation with the levels of immune infiltrates of cancer-associated fibroblasts (CAFs) in multiple tumours (Tuo et al., 2022). The dual functionality of RUNX1 as both a prognostic biomarker and an indicator of the immune microenvironment within tumours highlights its potential significance in the context of cancer treatment.

RUNX1 interacts with chromatin modifiers such as histone acetyltransferases, including EP300, which is significantly represented in our analysis, and regulates chromatin structure. Furthermore, a collaborative relationship between the transcription factor RUNX1 and MYB has been observed in normal hematopoietic stem and progenitor cells and T-cell acute lymphoblastic leukaemia (T-ALL) (Guan et al., 2023). Additionally, Gao et al. (2014) conducted a study to identify potential genes that may interact with the MYB pathway to facilitate ACC tumorigenesis. Their findings suggested a potential involvement of RUNX1 in regulating MYB transcriptional activation. Our experimental data analysis identifies MYB as a potential regulator, indicating a probable interaction between these two molecules. All these findings imply a potential role of RUNX1 in regulating the expression of MYB.

In normal salivary glands, Ono Minagi et al. (2017) demonstrated that the gene RUNX1 is implicated in the development of granular convoluted tubules within the submandibular glands of mice. Furthermore, in lacrimal glands, RUNX1 is identified as a gland morphogenesis and regeneration regulator, potentially exerting regulatory control over the cell cycle (Voronov et al., 2013). In the context of head and neck adenoid cystic carcinoma, Soares et al. (2021) conducted a study to explore the effect of RUNX1 expression on the prognostic outcomes. In their study, it was revealed that RUNX1 exhibited a significant correlation with several clinicopathological parameters in ACC highlighting the potential implications for the clinical behaviour of ACC, notably concerning tumour size, lymph node status, clinical stage, presence of metastasis, neural and vascular invasion, surgical margins and recurrence. This aligns with the findings of Gao et al.

(2014), which identified RUNX1 as a regulator of MYB, the key driver in ACC. Thus, the interaction between RUNX1 and MYB likely contributes to ACC progression and clinical behaviour.

In 2023, Jun-shui et al. explored the regulatory role of the long non-coding RNA (LncRNA) RUNX1-IT1 on microRNA-195 (miR-195) and its downstream target CyclinD1 in carcinoma ex pleomorphic adenoma (malignant pleomorphic adenoma). Their findings indicated elevated levels of LncRNA RUNX1-IT1 and CyclinD1 but reduced levels of miR-195 in tumour tissues compared to non-tumour tissues. Remarkably, the expression of LncRNA RUNX1-IT1 exhibited an inverse correlation with miR-195 and a direct association with CyclinD1 levels. Moreover, increased levels of LncRNA RUNX1-IT1 and CyclinD1, along with decreased miR-195 expression, were found to correlate with larger tumours, recurrent disease, and distant metastasis. The study concluded that the regulatory influence of LncRNA RUNX1-IT1 on miR-195/CyclinD1 may contribute to carcinoma ex pleomorphic adenoma progression, suggesting its potential as a therapeutic target for intervention.

Inhibition of 'PPARα/RXRα Activation pathway'

Peroxisome proliferator-activated receptor alpha (PPARs) are a specific type of nuclear receptor (NRs) that function as ligand-activated transcription factors (TFs) (Mirza et al., 2019). When a ligand binds, PPARs form dimeric complexes with retinoid-X-receptors (RXRs), which then bind response elements in target genes to carry out crucial regulatory functions (Amber-Vitos et al., 2016). PPARs are particularly known for their significant roles in lipid and glucose homeostasis, nutrient sensing, inflammation, cellular differentiation, and development (Feige et al., 2006). The PPAR family includes three isoforms: PPAR-alpha (NR1C1), PPAR-beta/delta (NR1C2), and PPAR-gamma (NR1C3).

As stated above, PPARs are not limited to a single function. Their versatility is evident in their anti-fibrotic and anti-inflammatory effects, which are beneficial in a wide range of conditions, including cancer, autoimmune diseases, liver steatosis, and type 2 diabetes (T2D) (Fougerat et al., 2020; Liu et al., 2020; Font-Díaz et al., 2021). PPARs promote the expression of anti-inflammatory molecules while simultaneously inhibiting the production of extracellular matrix proteins and pro-

inflammatory cytokines. Furthermore, they modulate the response and phenotype of immune cells, including macrophages and lymphocytes (Liu et al., 2020).

When it comes to cancer, the role of PPARs is of utmost importance as they have a well-established track record of demonstrating anti-tumorigenic effects (Font-Díaz et al., 2021). Their activation can trigger apoptosis and induce cell death in tumours, thereby halting tumour growth and inflammation. The influence of PPARs on tumour development is primarily linked to the regulation of cell-cycle blockade genes, such as p18, p21, and p27, which in turn leads to apoptosis by suppressing B-cell lymphoma 2 (Bcl-2) and reducing angiogenesis through the inhibition of vascular endothelial growth factor (VEGF) (Koga et al., 2001; Tianet al., 2016).

Pozzi et al. (2007) and Panigrahy et al. (2008) investigated the effects of PPARα on tumour growth and vascularisation using mice as experimental models. Their findings suggested that activating PPARα could be beneficial in preventing or treating various types of cancer. When administered orally, different PPARα agonists have been shown to inhibit the growth of tumours derived from melanoma, Lewis lung carcinoma, glioblastoma, and fibrosarcoma cell lines. They have also been found to inhibit angiogenesis in these models. These inhibitory effects occur through the PPARα-dependent inhibition of endothelial cell proliferation and the PPARα-dependent downregulation of an enzyme called cytochrome CYP2C. This enzyme plays a role in promoting angiogenesis by catalysing the epoxidation of arachidonic acid to epoxyeicosatrienoic acids. Thus, inhibiting tumour vascularisation disrupts the tumour microenvironment, affecting cell survival and growth, and may potentially slow tumour growth.

Our IPA canonical pathway analysis revealed that the DEPs enriched in the PPARα/RXRα pathway, manifest an inhibitory effect, resulting in the silencing of this pathway. The inhibition of the PPARα/RXRα pathway engenders detrimental consequences for cancer development and progression. This is evidenced by the observed activation of crucial cancer-related biological processes and functions, including angiogenesis, vasculogenesis, and endothelial cell proliferation, along with the inhibition of apoptosis.

Inhibition of Sumoylation pathway

Sumoylation (Small Ubiquitin-like Modifier) is a post-translational modification process that involves the attachment of SUMO proteins to target proteins. This modification can alter the function, localisation, stability, and interactions of proteins, impacting a wide range of cellular processes, including transcription regulation, DNA repair, cell cycle progression, and stress responses (Flotho & Melchior, 2013).

Sumoylation is crucial for maintaining normal cellular functions and has been implicated in various diseases, particularly cancer. The process is tightly regulated, and dysregulation of Sumoylation pathways can lead to aberrant cell proliferation, apoptosis, and genomic instability, which are hallmarks of cancer (Seeler & Dejean, 2017).

According to Yu et al., (2014), Sumoylation acts as an inhibitory modification on PLAG1's transactivation activity. This is in contrast to acetylation, which enhances the transcriptional activity of PLAG1. The inhibition of the Sumoylation pathway, as observed in this study, suggests that Sumoylation normally suppresses PLAG1 activity. If Sumoylation is inhibited, PLAG1 might become more active than usual. This can potentially disrupt the normal regulatory mechanisms in cells, leading to abnormal cell growth and differentiation. In pleomorphic adenoma, where PLAG1 is frequently overexpressed, inhibition of Sumoylation could exacerbate tumour development by allowing PLAG1 to activate its target genes more effectively, thereby promoting tumour growth.

Upstream Regulator Analysis

The upstream regulator analysis was conducted to identify potential upstream regulators that may have contributed to the observed changes in the differential protein expression. The use of IPA facilitated the detection of various molecules, including transcription factors, miRNA, chemicals, drugs, or compounds, that could have influenced the dataset identifiers. These molecules may not have been overexpressed in the dataset (no Exp Log ratio provided) but could be activated by post-translation modifications. However, the study aimed to initially delineate the

potential biological upstream regulators expressed in the dataset, representing the PLAG1 proximal interactors. This approach focused on biologically validating only the expressed molecules, taking into account statistical significance and biological context. Identification of proteins shared among the most significant pathways or connecting multiple downstream genes indicated their importance. Based on these selection criteria, MYB, JUN, EP300 and CREBBP have been identified as our prime candidate proteins for in-depth research and biological validation of their co-localisation with PLAG1 in pleomorphic adenoma tissue samples using immunohistochemistry.

The gene EP300 (E1A binding protein p300) encodes a protein that plays a key role in regulating gene expression. It achieves this by chemically modifying histones, which act as spools for DNA that package our genes (Eckner et al., 1994). This modification, known as acetylation, loosens the packaging and enables genes to be more readily turned on or off. EP300 is crucial in regulating cell proliferation and differentiation (Gayther et al., 2000).

In the context of cancer, the EP300 protein exerts influence over essential biological functions pertinent to tumour initiation and progression. In its capacity as a transcription factor, EP300 can promote cancer by orchestrating the regulation of a diverse set of genes implicated in cancer, including but not limited to proliferation, survival, and migration (Ding et al., 2023). EP300 can undergo modifications such as mutations, deletions, and amplifications or alterations in expression level (overexpression), which can lead to alterations in its normal function (Kim et al., 2022).

The gene EP300 demonstrates a unique duality in cancer. Depending on the specific genetic alteration and the type of cancer, it can function as either a tumour suppressor or an oncogene, a gene that promotes cancer. In certain instances, mutations within EP300 may compromise its capacity to regulate gene expression, resulting in the suppression of genes that govern cell growth and differentiation. This process can impede tumour development. EP300 has the potential to impact various DNA repair mechanisms by regulating genes involved in DNA repair or via interaction with repair proteins, thereby facilitating the repair process (Manickavinayaham et al., 2019). Alternatively, mutations within specific domains of EP300 may elicit a functional gain,

thereby promoting the proliferation and viability of cancer cells. This occurs through the acetylation-induced activation of genes such as cyclin D1 and cyclin E, which play crucial roles in propelling cells through the cell cycle. The EP300 protein is postulated to possess the capacity to regulate genes associated with cell survival, thereby potentially contributing to the development of resistance to cancer therapies. Additionally, EP300 may play a key role in angiogenesis, a fundamental process for tumour growth. Furthermore, its influence on cell migration and invasion may be intricately linked to its facilitation of metastasis. (Mahmud et al., 2019).

EP300 also plays a pivotal role in regulating immune cell function and differentiation, thus exerting a substantial influence on the tumour microenvironment and immune responses. This interplay ultimately affects the progression of cancer and the effectiveness of immunotherapy. (Chen et al., 2021).

Altered expression levels and mutations in EP300 have been associated with a variety of cancer types, including colon adenocarcinoma, lung adenocarcinoma, oesophageal carcinoma, bladder carcinoma, cutaneous melanoma, hepatocellular carcinoma, breast invasive carcinoma, and head and neck squamous cell carcinoma (Martin et al., 2014; Chen et al., 2022).

The function of EP300 in salivary gland tumours, specifically ACC, is currently under investigation and not completely understood. Research has indicated that the expression of EP300 seems to be increased in certain ACC tumours compared to healthy tissue, implying a possible role in the formation or progression of ACC (Lin et al., 2023). Additionally, the NOTCH signalling pathway is recognised for its involvement in ACC development. Notably, research indicates that the activation of this pathway may be associated with the upregulation of genes implicated in epigenetic regulation, such as EP300 and its closely related counterpart, CREBBP (Mat Lazim et al., 2023).

In their study, Zheng and Yang (2005) investigated the regulatory functions of acetylation in the transcriptional control of PLAG1 and PLAGL2. Their findings suggested that both genes can be activated by EP300-mediated acetylation and repressed by deacetylation. These findings imply a significant biological impact of EP300 in the regulation and activation of PLAG1, indicating its potential as a therapeutic target in associated tumours.

While there has been significant research interest in the role of EP300 in salivary gland tumours, the majority of studies have concentrated, as mentioned above, on ACC. However, there remains a notable gap in the literature regarding the specific impact of EP300 on pleomorphic adenoma (PA). At present, there are no conclusive findings or definitive evidence regarding the direct impact of EP300 on PA. This emphasises the necessity for further comprehensive investigation in this particular area to gain a clearer understanding of any potential influence.

Our study identified EP300 as one of the most significantly upregulated genes, representing a PLAG1 proximal interactor. The IPA analysis has also acknowledged the EP300 protein as an upstream regulator, indicating direct and indirect impacts on the differentially expressed proteins and affected downstream functions. As per the data presented in Table 4.6, EP300 has the ability to directly activate 42 downstream targets within the dataset. Additionally, it exerts an indirect influence by modulating the expression of other intermediate targets, such as JUN and FOXP, thereby impacting the expression of target genes. Furthermore, in conjunction with its closely related CREBBP protein, EP300 was found to be one of the most commonly shared proteins among the significantly affected pathways, underscoring its importance in regulating canonical pathways (Table 4.2 overlapping pathways). The downstream effects analysis of the affected functions has revealed the significant impact of EP300 on various tumour-related functions. This encompasses the activation of biological processes related to cell viability and survival, DNA repair, blood cell proliferation, RNA transactivation, viral infection, and B lymphocyte quantity. Conversely, EP300 has been shown to suppress functions associated with apoptosis, growth failure, and tumour cell sensitivity. These findings align with previously reported literature indicating the potential role of EP300 in promoting cancer.

It is important to note that EP300 represents a complex protein with multifaceted functions and is currently the subject of active research. Consequently, gaining a comprehensive understanding of its involvement holds the potential for the development of innovative therapeutic strategies directed towards EP300 or its downstream pathways.

CREBBP (CREB-binding protein) is a transcription factor and typically acts as a histone acetyltransferase (HAT) comparable to EP300. This protein modifies histones through the addition

of acetyl groups, thereby inducing relaxation of chromatin structure. Consequently, this process allows for the expression of genes crucial for regulating cell growth and differentiation (Shiama, 1997).

Mutations or altered expression can disrupt CREBBP's regulatory function, inhibiting genes that regulate cellular proliferation and differentiation and thus hindering tumour development (Zhang et al., 2017). Conversely, aberrant CREBBP activity can potentiate tumorigenesis by dysregulating cellular growth and viability genes, thereby contributing to uncontrolled cell division. Some studies suggest that CREBBP alterations may also impact DNA repair mechanisms, resulting in genomic instability and promoting tumorigenesis (Mat Lazim et al., 2023). The protein CREBBP is also implicated in angiogenesis, a process vital for the proliferation and metastasis of tumours (Arany et al., 1996). Its influence on cell migration and invasion may be correlated with its role in advancing metastasis. Additionally, CREBBP may regulate genes associated with cell survival, potentially contributing to resistance to cancer therapies. These diverse roles of CREBBP in cancer make it a potential biomarker, providing valuable information for cancer diagnosis and prognosis. Like EP300, the specific function of CREBBP in cancer is contingent upon the cancer type and the specific mutations present within the CREBBP gene. Depending on the context, it can serve as both a tumour suppressor and an oncogene.

Mutations in the CREBBP gene, while not the most prevalent drivers of cancer development, are present in a notable percentage of certain cancer types. The specific prevalence varies depending on the type of cancer, underscoring the intricacy of this area of research. CREBBP mutations are relatively common in hematologic malignancies such as acute myeloid leukaemia, acute lymphoblastic leukaemia, diffuse large B-cell lymphoma, follicular lymphoma, and some myelodysplastic syndromes (Zhang et al., 2017; Zhu et al., 2023). They are also identified in various solid tumours, including adenocarcinoma of the lung, bladder cancers, and certain gastrointestinal cancers (Li et al., 2021). Additionally, CREBBP mutations have been documented less frequently in other cancers, such as head and neck squamous cell carcinoma (HNSCC) (Martin et al., 2014), however, research into the involvement of CREBBP/EP300 in the development and metastasis of HNSCC is currently ongoing. Notably, the presence of CREBBP/EP300 mutations has

been correlated with unfavourable prognoses in HNSCC, as well as treatment resistance subsequent to radiation therapy (Kumar et al., 2020).

Similar to EP300, the majority of research studies have established a correlation between CREBBP alterations and ACC. There is evidence to suggest that gene mutations, specifically CREBBP mutations, exhibit increased prevalence in recurrent and metastatic ACC when compared to the primary tumour (Ho et al., 2019). Furthermore, the literature indicates a more aggressive biological manifestation of ACC in the recurrent/solid type, characterised by enriched CREBBP and EP300 mutations. It is noted that CREBBP and EP300 function as co-activators of MYB, contributing to the tumour's malignancy (Lin et al., 2023).

To date, no previous research has looked into the role of the CREBBP protein in relation to pleomorphic adenoma. However, in this study, we demonstrated a link between the conditional overexpression of PLAG1 and the CREBBP protein. Our analysis showed that CREBBP was significantly upregulated in the PLAG1-expressing cells compared to the mock samples.

As a biological upstream regulator, CREBBP can directly control the expression of 27 target genes within the dataset and influence the expression of other downstream proteins by intricately affecting other intermediary transcription factors and enzymes such as JUN, HSPA5, BRCA1, and HSPA8. (Figure 4.21 and Table 4.7). The biological significance of this molecule is underscored by its essential participation in key signalling pathways. This is clearly demonstrated by the extensive overlap of canonical pathways, indicating its multifaceted role in regulating vital cellular processes. CREBBP was found to be involved in 16 out of the top 20 significant pathways. (Table 4.2).

The biological significance of the molecule CREBBP, is underscored by its essential involvement in several critical pathways. These pathways include the activation of the anterior HOX genes pathway, transcriptional regulation by RUNX1, and transcriptional regulation of white adipocyte differentiation. This multifaceted role highlights the extensive impact of CREBBP on vital biological functions, as evidenced by its implication in 16 out of the top 20 significant pathways.

Furthermore, the downstream effects analysis points to CREBBP's active participation in promoting the proliferation of blood cells, angiogenesis, vasculogenesis, cell viability, cell survival,

RNA transactivation, viral infection, and B lymphocyte quantity. Notably, the protein's ability to inhibit apoptosis and reduce the sensitivity of tumour cell lines parallels that of its close relative, EP300. This comprehensive involvement underscores the potential significance of CREBBP as a promising target for therapeutic interventions and as a biomarker in diverse biological processes and diseases.

As mentioned earlier, the proteins CREBBP and EP300 are acknowledged as co-activators of the transcription factor MYB. They exert their regulatory influence by binding to MYB's transactivation domain (TAD), thereby initiating the transcriptional process. This molecular mechanism ultimately contributes to the progression of tumour malignancy driven by MYB (George et al., 2014).

The MYB, proto-oncogene, encodes a transcription factor that exerts a pivotal influence on the maintenance of undifferentiated states in stem and progenitor cells across various tissues. Its predominant expression is localised in hematopoietic tissues, colonic crypts, and neural stem cells. MYB's regulatory role is essential for preserving the undifferentiated state in these cell populations. (Ramsay et al., 2008).

In healthy tissues, the expression of MYB family proteins, such as MYB, MYBL1, and MYBL2, is tightly controlled. However, human malignancies exhibit a range of alterations, including gene amplification, mutations, and structural rearrangements due to chromosomal translocation or gene fusion. These changes result in heightened biological activity, playing a role in driving different aspects of tumour formation. MYB proteins are believed to have a central role in a wide range of cancer types by influencing multiple aspects of cancer development. This includes their impact on cell proliferation, angiogenesis, ability to promote cell survival, role in maintaining stemness, invasion of surrounding tissues, and remodelling of the supportive stromal environment.

From the earliest research conducted on MYB, it was noted that this gene exhibits a highly selective expression pattern within stem cells. Subsequent findings unveiled its significant role in maintaining the undifferentiated, or stem cell, state (Kuehl et al., 1988; Dyson et al., 1989; Introna et al., 1994). The overexpression of MYB family genes in multiple malignancies indicates their

potential role in supporting the continued progression of cancer. This suggests that these genes may contribute to enabling cancer cells to adapt to the dynamic microenvironmental changes, thereby displaying essential adaptive capabilities crucial for their survival. Consequently, cellular plasticity emerges as a pivotal attribute facilitating the resilience of cancer cells when confronted with challenging environmental conditions.

The uncontrolled proliferation of cells, a crucial biological process, facilitates the onset of cancer at the primary site and its establishment at secondary metastatic sites. A preliminary study on MYB function, which plays a role in cell cycle progression, revealed a transient upsurge in MYB transcript levels through post-transcriptional mechanisms in various cell types, thereby establishing a correlation between MYB function and uncontrolled cell proliferation (Thompson et al., 1986). The significance of MYB function was later validated by suppressing its expression, leading to a notable reduction in the growth of myeloid leukaemia cells (Anfossi et al., 1989). Additionally, Drabsch et al. (2007) have provided evidence indicating that the activity of the MYB gene is subject to regulation by oestrogen/ER signalling. Their findings suggest that MYB contributes to the proliferation and survival of oestrogen receptor-positive (ER+) breast cancer cells.

Several reports have extensively detailed the pivotal role of MYB in the behavioural characteristics that promote the metastatic dissemination of cancer cells. MYB has been identified as a participant in the interaction with the Wnt effector β-catenin, leading to the co-activation of downstream target genes implicated in the process of invasion and metastasis in breast cancer cells (Li et al., 2016). It has also been proven that MYB plays a crucial role in promoting the ability of pancreatic cancer cells to invade surrounding tissues and spread to the liver, lungs, and spleen (Srivastava et al., 2015).

Several studies have highlighted the crucial role of MYB proteins in bolstering the survival of cancer cells by promoting cell growth and DNA damage repair, which can present a significant challenge to the effectiveness of anticancer medications (Wang et al., 2015; Yang et al., 2019).

A significant number of ACC tumours, which are mostly found in the salivary glands but also in the breast and other tissues, exhibit a recurring translocation (Persson et al., 2009; Brill et al.,

2011; Roden et al., 2015) involving the fusion of the MYB proto-oncogene on chromosome 6q with the NFIB gene on chromosome 9p, potentially leading to the expression of novel MYB–NFIB fusion oncogenes. Consequently, this fusion generates a chimeric protein with a modified function compared to normal MYB. (Mitani et al., 2010; Brill et al., 2011; Mitani et al., 2011). The chromosomal translocation frequently results in the overexpression of the MYB protein or its aberrant fusion product. This increased expression leads to uncontrolled cellular proliferation, thus contributing to the pathogenesis of tumours.

The involvement of MYB in salivary gland tumours beyond ACC remains inadequately established. Although there are indications in some studies that MYB alterations may manifest in other types of tumours, conclusive evidence is currently lacking. In a study conducted by Lee et al. (2019), the co-expression of MYB and PLAG1 in human salivary gland neoplasm samples, particularly pleomorphic adenoma, was examined using immunohistochemistry. Results indicated the absence of Myb and PLAG1 expression in normal salivary gland tissue but among the 48 pleomorphic adenoma cases studied, PLAG1 protein expression was detected in 29 cases (60.4%), while 16 out of the 48 pleomorphic adenomas (33.3%) exhibited MYB positivity. The co-expression of PLAG1 and MYB genes implies a potential relationship between their functions. However, it does not conclusively prove a direct interaction. It is plausible that MYB and PLAG1 may affect each other's functions through downstream pathways. For example, MYB may have the ability to regulate the expression of genes involved in cell proliferation, which could be highly relevant to PLAG1's involvement in tumorigenesis.

Our research has provided compelling evidence that the activation of PLAG1 expression in vitro leads to the biotinylation and enrichment of the MYB protein. According to the IPA analysis, MYB was involved in three main canonical pathways: transcriptional regulation by RUNX1, ESR-mediated signalling, and the sumoylation pathway, which reflects its significance in initiating the oncogenic process. The data presented in Figure 4.18 and detailed in Table 4.4 demonstrate the intricate interplay of MYB, the upstream molecule, and downstream targets, highlighting MYB's ability to target 27 genes selectively. Furthermore, MYB's influence extends to the modulation of downstream targets, wherein it can either upregulate or inhibit (affect) other genes, such as HSPA5, BRCA1, and HSPA8, or impact other upstream regulators like JUN. Furthermore, our

findings highlighted the key role of MYB in various downstream functions, underscoring its significant oncogenic potential. We observed that MYB activation is associated with critical functions such as the proliferation of blood cells, angiogenesis, cell viability, colony formation, proliferation of progenitor cells and stem cells, increased quantity of B lymphocytes and pre-B lymphocytes, cell survival, and transactivation of RNA. Conversely, MYB also demonstrated inhibitory effects on the metastasis process, indicating its potential dual role as an oncogene or tumour suppressor, contingent on the specific cancer context. This intricate duality in MYB's functions contributes to the complexity of our understanding of cancer biology, which resonates with the benign nature of pleomorphic adenoma, which exhibits a subtle inclination towards malignant transformation into carcinoma.

The JUN gene encodes a protein, referred to as c-JUN, which is involved in the regulation of gene expression within normal tissue. In conjunction with the protein c-Fos, C-Jun forms AP-1, an early response transcription factor. Together, they directly interact with specific DNA sequences, influencing how genes are expressed. Both c-Jun and its partnering proteins are influenced by external triggers such as growth factors, cytokines, stress, and UV irradiation. Additionally, the production of c-Jun is regulated by its own product, creating a mechanism that prolongs the effects of external triggers on gene expression (Shaulian et al., 2002).

JUN mutation or dysregulation has been identified in several cancer types, such as breast cancer (Vleugel et al., 2006) and colorectal cancer (Wang et al., 2000), fibrosarcoma (Bossy-Wetzel et al., 1992), glioma (Blau et al., 2012) and lung carcinoma (Szabo et al., 1996).

JUN can promote uncontrolled cell growth, a hallmark of cancer, by binding to specific DNA sequences in the promoters of genes critical for cell proliferation, such as cyclin D1 (Schwabe et al., 2003). This interaction activates gene expression, propelling cells through the cell cycle. Furthermore, JUN can suppress genes involved in programmed cell death (apoptosis), thereby enabling the survival of damaged or abnormal cells (Katiyar et al., 2010). Nevertheless, JUN can demonstrate a context-dependent role in the context of cancer as its functionality is frequently modulated by its counterpart protein, FOS, within the AP-1 complex. The specific JUN to FOS ratio can dictate whether the complex facilitates cellular proliferation or differentiation (maturation).

In specific contexts, a well-balanced AP-1 complex with sufficient FOS may promote differentiation, potentially impeding tumour development.

The role of the JUN gene can differ based on the type of cell and the particular mutations found in the tumour. In certain types of cancer, mutations in the JUN gene can result in a loss of its ability to suppress tumours and can instead promote uncontrolled growth and it has also been shown to contribute to the formation of new blood vessels (angiogenesis) required for tumour growth. Moreover, several studies have indicated that JUN may have a regulatory role in Matrix metalloproteinases (MMPs), which are enzymes responsible for the degradation of extracellular matrix components and promoting invasion (Cheung et al., 2006). JUN could also influence genes that code for cell adhesion molecules, thereby affecting cell migration and metastasis (Zhang et al., 2007).

Extensive research has been conducted on the role of the JUN gene in various cancers, yet its precise involvement in salivary gland tumours remains unclear, highlighting the need for further research in this area. This lack of specific research sets salivary gland cancer apart from other cancer types, such as lung cancer, where the role of the JUN gene has been more extensively explored.

Our proteomics data analysis has demonstrated that following the overexpression of PLAG1, JUN emerges as a particularly influential upstream regulator. In its role as a transcription factor, JUN appears to directly influence a remarkable 84 genes, making it one of the most potent regulators among its counterparts. Furthermore, it can indirectly modulate other genes through intermediary proteins, including the upstream regulator MYB, as shown in Figure 4.19 and Table 4.5. JUN plays a crucial role in regulating multiple canonical pathways, notably contributing to the activation of the anterior HOX gene, glucocorticoid receptor signalling, oestrogen receptor (ESR)-mediated signalling, and retinoic acid receptor (RAR) activation. Conversely, it has been observed to have inhibitory effects on the PPARα/RXRα activation and sumoylation pathways. These findings underline the multifaceted role of JUN in cellular signalling and highlight its potential significance in various physiological and pathological processes. The data indicates that the changes in biological functions align with the JUN protein's expected role, as described in existing

scientific literature. This not only validates our research but also underscores its potential to inspire future studies. The changes observed include increased cell viability and survival, DNA repair, angiogenesis, vasculogenesis, transactivation of RNA, and the proliferation of several cell types critical for tumour growth, such as embryonic cells, endothelial cells, blood cells, and muscle cells. Conversely, JUN appears to suppress apoptosis-related functions, inflammatory response, and cellular movement and invasion.

Mechanistic network analysis

Upstream regulators within a biological system are not isolated entities but rather intricately interconnected. This revelation, brought to light by the rigorous research methodology of Mechanistic Network Analysis, unveils potential connections or interactions among multiple upstream regulators. The statistical enrichment method employed in this analysis has uncovered significant associations and overlaps between these regulators and the targets present within the experimental dataset, sparking a new level of intrigue and engagement in our research.

Following an extensive analysis, the IPA (Ingenuity Pathway Analysis) has identified, MYB and JUN as pivotal upstream regulators within the sample group expressing PLAG1. Accordingly, a comprehensive mechanistic network analysis has been formulated to facilitate a detailed understanding of the interacting upstream regulators that drive altered gene expression through downstream effector molecules. Figures 4.22 and 4.23 portray the complex interplay between the upstream regulators MYB and JUN and their corresponding target genes within the dataset. In this intricate interplay, both proteins impact shared intermediary regulators, such as JNK, FOS, and TP53. These intermediary regulators, in turn, exert influence on the target genes within the experimental dataset located at the bottom of the networks, indicating the substantial functional overlap between MYB and JUN.

In a study conducted by Quintana et al. (2011) to ascertain MYB target genes in human MCF-7 breast cancer cells, the JUN gene was identified as a novel downstream target of MYB. However, our data analysis revealed an intricate relationship between these two molecules. Specifically, our findings indicate that JUN possesses the capacity to directly trigger the activation of MYB. In contrast, MYB's impact on JUN appears to be indirect, thereby introducing an additional layer of

complexity to their interaction, signifying the need to further explore their regulatory mechanisms. It is important to note that the two upstream regulators, EP300 and CREBBP, have been reported in the literature to function as potential modulators of MYB and JUN through the modification of histones in close proximity to the binding sites of MYB and JUN on DNA. This modification renders these regions more accessible to transcription factors, thereby facilitating the binding and subsequent activation of gene expression. Furthermore, the utilisation of online protein-protein interaction tools such as STRING and Signor not only served to validate these interactions but also uncovered a well-documented association between PLAG1 and EP300. It has been reported that EP300 activates PLAG1 expression through acetylation (Zheng and Yang, 2005). The interaction between all these influential regulators in the context of human cancer opens up new possibilities for targeted therapies in tumours expressing these molecules.

Diseases and functions

Following the identification of potential upstream regulators, we performed the downstream effector analysis. This crucial step identifies the functions and biological processes affected by the deregulation of the DEPs in response to PLAG1 overexpression. It also predicts their activation state based on their Z score. A z-score greater than zero indicates likely activation of the biological function, while a z-score less than zero suggests likely inhibition. Z-scores exceeding two or falling below -2 are typically deemed statistically significant. The analysis relies on established connections between specific proteins and their corresponding biological effects, as documented in the literature.

Based on our results, the most significant diseases and biological functions affected by PLAG1 induction were cancer and gene expression, as represented by their significant p-values (Tables 4.8 and 4.9). The majority of the enhanced biological processes align with the established effects of PLAG1 in pleomorphic adenoma and are consistent with its documented role in the literature. However, most of the inhibited functions are associated with PLAG1's role in development and extend beyond the biological context of pleomorphic adenoma, with the exception of biological functions linked to apoptosis, invasion, and metastasis, which are known to be inhibited by PLAG1 in this context. Using this analysis, we were able to correlate the upstream regulators and their

targets in the dataset with the downstream effects on the altered biological processes and functions as discussed in the upstream regulator section.

Interestingly, this analysis revealed the activation of functional categories previously demonstrated to be affected by PLAG1 ectopic expression in pleomorphic adenoma (Van Dyck et al., 2007; Wang et al., 2013; Wang et al., 2022). These categories include Cellular Development and Cellular Growth and Proliferation, which encompass cellular processes related to cell proliferation, differentiation, migration, and colony formation (Figures 4.25 and 4.26). Moreover, a notable enrichment in the cell death and survival category was observed, supported by statistical significance (Figure 4.27). This category is particularly important as it comprises cellular processes with opposing activation states; while cell viability, survival, and self-renewal were enhanced, apoptosis was decreased.

The PLAG1 protein and its proximal interactors have been demonstrated to play a role in the development of tumours by promoting several crucial functions essential for tumour growth including the formation of new blood vessels (angiogenesis) and blood vessel formation (vasculogenesis), as well as promoting cell viability, survival, and the ability to form colonies. PLAG1 also contributes to the proliferation of a variety of cells while inhibiting apoptosis, the natural process of cell death. Interestingly, PLAG1 is also known to have a suppressing effect on the invasion and spread of tumour cells to other parts of the body (metastasis). All these counteracting functions reflect the dual function of PLAG1 as being an oncogene and tumour suppressor. It is important to note that the balance between these opposing functions determines the non-cancerous (benign) nature of the tumour. Any alterations to this delicate balance could potentially lead to more aggressive behaviour and, in rare instances, a transformation into a malignant tumour (carcinoma ex PA). These findings underscore the complex role of PLAG1 in cellular processes and open up new avenues for further research.

Regulator effect analysis

Following the identification of upstream regulators and their downstream effects on biological processes and functions, it was fundamental to generate a wider view of the entire process of tumorigenesis. Linking the upstream regulators with their targets in the dataset and the

downstream affected functions has been instrumental in generating a holistic understanding of the entire process of tumorigenesis, thereby expanding our knowledge of the mechanisms driven by PLAG1 in the development of pleomorphic adenoma.

Using the IPA, a total of 16 regulator effect networks were predicted, however, it is important to note that several of these networks delineate the role of PLAG1 in the context of embryonic development, which lies beyond the scope of our current focus. Within the biological context of pleomorphic adenoma, the IPA recognised several molecules as being top (master) upstream regulators affecting the DEPs in the dataset and constructed regulator effect networks by integrating the upstream regulator results with their respective downstream effects. These master upstream regulators include growth factors (IGF1 and EGF), cytokines (IL6, IL33) enzymes (SNCA), compounds (ERK1/2), transcription regulators (HNF4A) and microRNA (mir-8), all of which have been evidenced to be implicated in cancer (Schetter et al., 2008; Yan et al., 2018; Hua et al., 2020; Sugiura et al., 2021). Based on their activation Z-scores, the analysis suggests that the genes SNCA, HNF4A, and mir-8 would be inhibited, while the other upstream regulators are expected to be activated.

Despite not being expressed in our samples, the presence of their reportedly commonly affected target molecules in the dataset strongly indicates the implication of these master regulators in the observed changes. This opens up intriguing possibilities for further research. As mentioned earlier, those upstream regulators may be activated, but not overexpressed, by post-transcriptional or post-translational modifications. In the regulator effect networks, the master upstream regulators are located at the top, while their targets in the dataset, primarily transcription factors and enzymes, are in the middle. Diseases and functions are positioned at the bottom of the network, representing the downstream outcome of the observed changes initially attributed to PLAG1 overexpression (Figures 4.29 to 4.34).

The network illustrated in Figure 4.29 reveals a complex web of interactions involving six significant upstream regulators: SNCA, HNF4A, EGF, IL6, IGF, and IL33. These interactions have a profound impact on 25 target proteins in the dataset, leading to the activation of downstream functions such as angiogenesis, development of cardiovascular tissue, development of epithelial

tissue, and proliferation of blood cells while at the same time inhibiting organismal death. Similarly, SNCA and IGF1 interact and overlap with ERK1/2 (Figure 4.30), affecting nine target proteins in the dataset. This interaction activates cell viability in tumour cell lines, colony formation of cells, proliferation of blood cells, and cell quantity. All of these functions have been reported to be affected in pleomorphic adenoma.

MicroRNA (mir-8), a mono regulator, has the capacity to drive a cascade of protein interaction in its regulatory network. It is predicted to exert a compound inhibitory effect on five downstream targets, including our protein of interest, PLAG1 (Figure 4.31). This inhibition promotes the proliferation of blood cells and progenitor cells while suppressing cell apoptosis, a crucial function in the context of tumorigenesis.

Insulin-like Growth Factor 1 (IGF1) and Epidermal Growth Factor (EGF) can function autonomously or as constituents of a broader network to instigate downstream targets, resulting in modifications to biological functions, as depicted in Figures 4.32, 4.33, and 4.34. IGF1 serves as a potent oncogene, as evidenced by its comprehensive impact on multiple affected functions. Notably, it has been suggested that MYB, along with BRCA1, are implicated in all IGF effect networks, underscoring their significance in the IGF1 pathway. Analogously, EGF facilitates the activation of tumour-related functions; however, it is expected to preserve the benign nature of the tumour by impeding or blocking the cell's ability to invade surrounding structures.

The signalling pathways of both EGF and IGF1 have been reported to interact and potentially enhance each other's impact on the progression of tumours. Specifically, the crosstalk between Epidermal Growth Factor Receptor (EGFR) and Insulin-like Growth Factor 1 Receptor (IGF1R) signalling pathways may give rise to more aggressive cancer phenotypes and resistance to therapies targeting these pathways (Hua et al., 2020).

The networks of regulatory effects represent custom-designed pathways predicted by the IPA. These pathways are generated by identifying the best matches (relationships) between regulators, as well as between regulators and diseases/functions. This approach allows for the discovery of potential new relationships within the specific scope of the study. It's important to

mention that certain targets in the dataset, such as MYB, JUN, and EP300, have been identified as potential upstream regulators. Therefore, we refer to them as biological upstream regulators.

The numerous predicted networks and pathways, whether acting autonomously or in complex overlapping patterns, offer potential targets for future investigation into their involvement in pleomorphic adenoma development. A comprehensive review of the existing literature and experimental data is essential to further explore the nature and function of each identified master upstream regulator and its downstream target genes/ biological function in the broader context of tumorigenesis, with specific emphasis on salivary gland tumours and pleomorphic adenoma. Consequently, we have designated this area as a focal point for our upcoming research to gain a more comprehensive understanding of our findings.

In summary, the proximity proteomic analysis and subsequent IPA analysis unveiled a complex regulatory network downstream of PLAG1 overexpression. PLAG1 appears to exert its oncogenic potential by interacting with proteins involved in transcriptional regulation, epigenetic modifications, and cellular metabolism. The observed activation of oncogenic signalling pathways, including HOX and RUNX1, coupled with the inhibition of tumour suppressive pathways such as PPARα/RXRα and sumoylation, suggests a multifaceted mechanism of action for PLAG1. Significantly, the downstream effector analysis revealed a pro-tumorigenic phenotype characterised by increased cell proliferation, survival, and colony formation, while concurrently suppressing apoptosis, invasion, and metastasis. These findings underscore PLAG1's critical role in driving tumorigenesis and progression in pleomorphic adenoma.

These findings suggest that PLAG1 may exerts its oncogenic potential in pleomorphic adenoma by disrupting essential cellular processes through protein-protein interactions and transcriptional regulation. The identified pathways and proteins are implicated in tumorigenesis, suggesting their potential roles in driving tumour growth.

However, it is crucial to note that this is a preliminary analysis, and further investigation is required to fully elucidate the precise mechanisms underlying PLAG1's oncogenic function in this specific tumour context.

5 Chapter 5 Investigating the role of PLAG1 on cell behaviour

5.1 Aims and Objectives.

Following the successful overexpression of PLAG1, it was crucial to investigate its impact on cell behaviour through the execution of functional assays as these serve not only to establish a connection between genetic and proteomic data and observable biological functions but also help to validate and substantiate the research findings presented in Chapter 4 section 4.4.5.5 (Affected Diseases and Biological Functions) with experimental evidence. The primary objective of this chapter was, therefore, to determine the influence of PLAG1 expression on cell behaviour and further to validate the results of the IPA downstream effect analysis. The work outlined in this chapter aimed to:

- Investigate and confirm the influence of PLAG1 de novo expression on cell proliferation via a cell proliferation assay.
- Perform a clonogenic assay to evaluate the effect of PLAG1 on cell survival and colony formation.
- Investigate the downstream effect of PLAG1 in modulating apoptosis using the TACS Annexin V-FITC Apoptosis detection kit.
- Conduct a transwell migration assay to investigate the effect of PLAG1 on cell migration.
- Examine the invasive potential of PLAG1 via a transwell invasion assay.

5.2 Materials and Methods

Following the evaluation of the timeline of de novo PLAG1 expression, transfection was performed as described in section 2.2.3. Briefly, HEK293 cells were cultured and transiently transfected with the two distinct constructs: PLAG1 and EV. To ensure consistent and reliable results from the functional assays, the transfection was performed in a T75 flask. Upon confirmation of successful transfection by fluorescence imaging, the cells were trypsinised and seeded according to the specific requirements of the respective assays, which included proliferation, migration, invasion, and colony formation. However, due to technical challenges in conducting the apoptosis assay simultaneously with the other assays, the transfection of cells for this assay was performed in the same 6-well plate as the assay. Detailed protocols for these assays are described in sections 2.6.

5.3 Results

5.3.1 Evaluation of PLAG1 protein expression level timeframe

Before proceeding with any functional experiments, it was important to evaluate the timeline of de novo PLAG1 protein expression. Transfection was carried out on two sets of samples, including a negative control, to enable monitoring and analysis at different time points. One set of samples was lysed 24 hours after transfection, while the other set was incubated and monitored for green fluorescence over the course of 7 days. On the seventh day, the samples were also lysed and PLAG1 protein expression, for both sets of samples, was evaluated using western blotting. Transfected HEK293 cells continued to exhibit high levels of green fluorescent protein expression for up to seven days, which was in line with the consistent presence of clear PLAG1 protein bands in samples collected on both day one and day seven post-transfection (Figures 5.1 and 5.2).

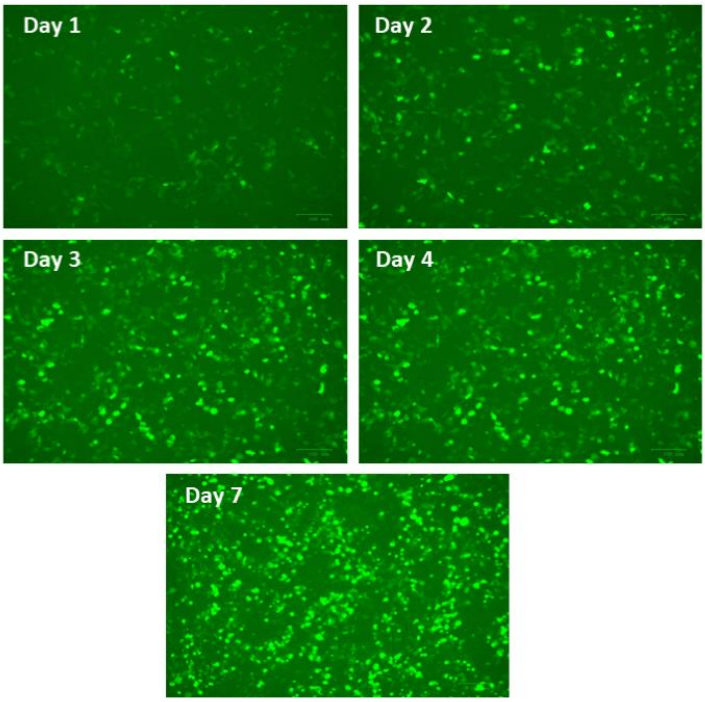


Figure 5.1 The expression pattern of PLAG1 in transiently transfected HEK293 cells.
Fluorescent microscopy images depict HEK293 cells transiently transfected with the FLAG-tagged PLAG1 pCDNA3.1+IREs GFP construct at five distinct time points, revealing sustained levels of GFP expression over the observed duration.

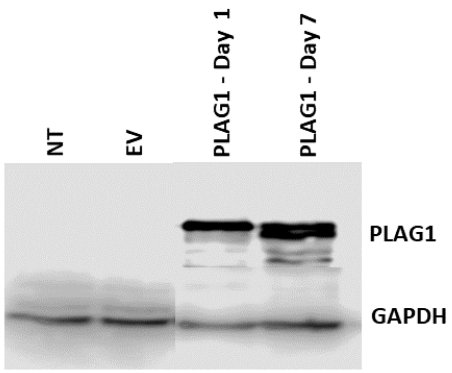


Figure 5.2 Western blotting analysis of PLAG1 protein expression.
PLAG1 protein expression was assessed via western blot in transiently transfected HEK293 cells at 1 and 7 days post-transfection using an anti-PLAG1 antibody. Consistent protein expression levels were observed in samples collected on day 1 and day 7. Control groups comprised non-transfected (NT) and cells mock-transfected with an empty vector (EV).

5.3.2 The effect of PLAG1 on cell proliferation and survival

The data outlined in section 4.4.5.5 (Affected Diseases and Biological Functions) demonstrates the downstream impact of PLAG1 overexpression on biological functions, indicating a considerable increase in cell proliferation alongside enhancements in cell viability and survival. In this chapter, two functional assays were performed to experimentally investigate the role of PLAG1 in relation to these specific biological functions and to corroborate our IPA proteomics data.

The CellTiter 96® AQueous One Solution Cell Proliferation Assay (MTS) was utilised to assess viable cell proliferation. The results clearly demonstrate that the overexpression of PLAG1 led to a significant increase in cell proliferation (**$p < 0.01$) with, notably,a 3-fold increase in cell proliferation compared to the EV, as shown in Figure 5.3. The data was analysed using GraphPad Prism with Student's t-tests.

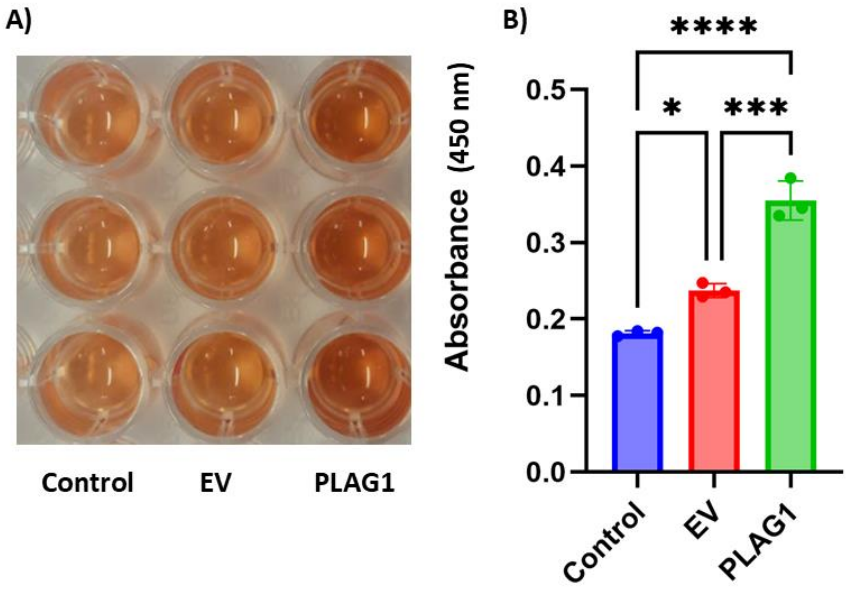


Figure 5.3. Cell proliferation assay.
A) The difference in the colour intensity between the two groups of samples (EV and PLAG1) is proportional to the number of viable proliferating cells. B) Values were analysed using GraphPad prism by student's t-test; data is presented as mean ± SD (N=3) **p < 0.01.

The clonogenic assay allowed us to determine the downstream effect of de novo PLAG1 expression on cell viability as this allows the assessment of a cell's capability to undergo clonal expansion from a single cell while maintaining its unlimited proliferation and differentiation capacity (Rajendran and Jain, 2017). After a 10-day incubation in normal growth media, the number of colonies formed was determined by scanning the plates and quantifying the number of colonies using ImageJ software.

Our data demonstrated a 35% (1.35-fold) increase in the number of colonies formed following PLAG1-transfection of HEK293 cells compared to the EV group. The results indicated that the PLAG1 gene significantly induced cell survival and expansion with a p-value of 0.0019, as shown in Figure 5.4.

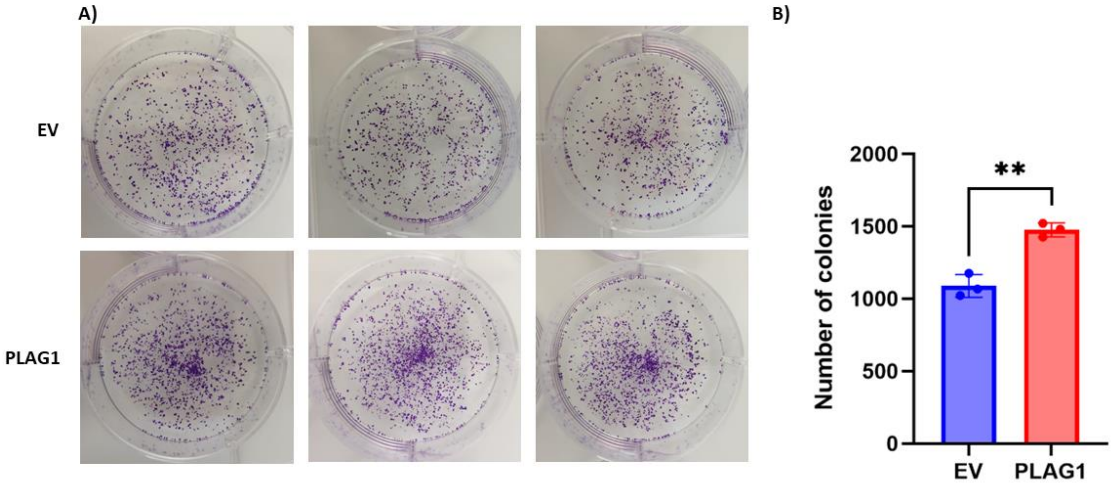


Figure 5.4 Colony formation assay.
A) Representative images of colonies of HEK293 cells transfected with empty vector EV (top) and PLAG1 (bottom) and cultured in normal growth media. After ten days, cells were fixed and stained with 0.2% Crystal Violet solution. B) Data quantification demonstrated a statistically significant (P= 0 0.0019) increase in colony formation and survival of PLAG1 transfected HEK293 cells in comparison to the control. Values were analysed using GraphPad prism by Student's t-test; data is presented as mean ± SD (N=3).

5.3.3 The effect of PLAG1 on programmed cell death (apoptosis)

In order to investigate the effect of PLAG1 on cell apoptosis and to validate our IPA analysis which suggested that PLAG1 might act as an inhibitor of apoptosis (chapter 4), we conducted the Annexin V-FITC Apoptosis assay, which. involves the use of flow cytometry following annexin V and PI staining. Annexin V is detectable in both early and late apoptosis, while PI stain is indicative of late apoptosis or necrosis. Early apoptotic cells were identified as annexin V-positive and PI-negative (upper left quadrant), while late apoptotic cells were dual-positive for annexin V and PI (upper right quadrant). Analysis of the findings in Figure 5.5 revealed that the exogenous expression of PLAG1 did not precipitate a statistically significant alteration in the proportion of either early or late apoptotic cells compared to the control group (NS $P > 0.05$).

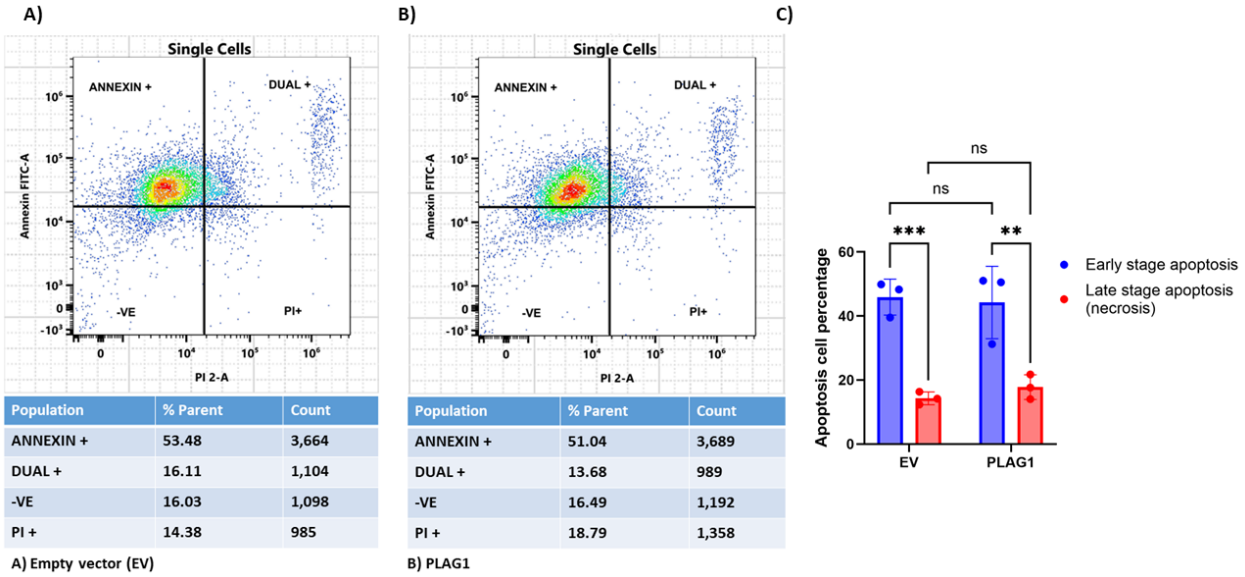


Figure 5.5 Results of the Annexin V-FITC apoptosis assay.
Apoptosis assay using flow cytometry after staining with annexin V-FITC/propidium iodide (PI). Representative scatter plots of Annexin FITC (y-axis) vs PI (x-axis) of A) Cells transfected with EV (control) and B) Cells transfected with PLAG1. The early-stage apoptosis, as indicated by the Annexin+ percentage (upper left quadrant) of the control group and PLAG1, at 53.48% and 51.04%, respectively, reveals no significant difference between the two groups. C) Values were analysed using GraphPad prism by Two-way ANOVA; data is presented as mean ± SD (N=3) NS $P > 0.05$.

5.3.4 The effect of PLAG1 on cell migration and invasion

The effect of PLAG1 expression on cell migration and invasion, was assessed after seeding of transfected cells into transwells in low-serum growth media, with the bottom well being filled with 10% FBS growth media to function as a chemoattractant. After the incubation period specified for each assay (see section 2.6), the membranes were fixed and stained with haematoxylin and eosin (H&E) and then mounted onto glass slides for quantification and data analysis. One key difference between the two assays lies in the design of the insert used. In order for the cells to migrate, they only need to penetrate the pores and move downwards toward the serum-rich media. Conversely, invasion necessitates the dissolution of the cell basement membrane and thus the inserts used for the invasion assay are pre-coated with Matrigel, which acts as a substitute extracellular matrix (ECM).

The membranes (on glass slides) were scanned using a Panoramic 250 Slide Scanner and ImageJ software was used for cell counting. The experiments were conducted at least three times in triplicate, however, no migration or invasion was detected across the timecourse of our experiment (Figure 5.6).

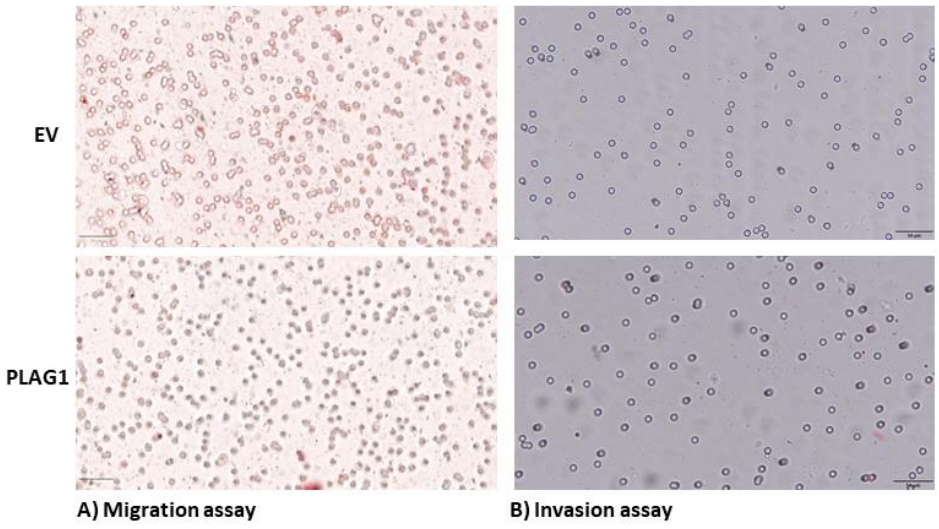


Figure 5.6 Transwell migration and invasion assays.
Transwell migration assay (A) and invasion assay (B). There were no observed indications of cell migration or invasion in cells transfected with EV (top) and PLAG1 (bottom) within the specified experimental timeframes. Specifically, the migration assay was conducted over 48 hours, while the invasion assay was extended to 120 hours**.**

5.4 Discussion

Studying cellular behaviour is fundamental to the investigation of protein function and its role in cellular pathways and biological processes. The primary objective of the research detailed in this chapter was to ascertain the influence of PLAG1 protein on cellular behaviour and to provide evidence to experimentally validate our IPA downstream effect analysis detailed in section 4.4.5.5. To accomplish this, gene transfection experiments were performed using HEK293 cells lacking endogenous PLAG1 expression, followed by downstream functional assays. Prior to commencing the in vitro functional assay, the PLAG1 protein expression level was assessed over a seven-day period. This evaluation was undertaken to validate the consistent and sustained expression of PLAG1 throughout the duration of any functional assays.

The initial in vitro findings suggest that PLAG1 protein significantly influences cell proliferation, survival, and colony formation but we saw no discernible involvement in cell migration or invasion. Furthermore, our investigation into the impact of PLAG1 on apoptosis modulation yielded inconclusive results within the time of our experiment.

The influence of PLAG1 on cell proliferation is not in isolation rather, it is frequently shaped by interactions with other genes and signalling pathways. PLAG1 exerts its influence by binding to specific DNA sequences and regulating the expression of target genes. There are several mechanisms through which PLAG1 promotes cell proliferation. It can influence the expression or activity of growth factors and their receptors, particularly IGF2, a potent growth factor involved in cell proliferation, differentiation, and cell death. PLAG1 can also interact with pivotal cell cycle proteins such as cyclins, cyclin-dependent kinases (CDKs), and their inhibitors, thereby intricately regulating the progression of the cell cycle. Additionally, it might prevent programmed cell death (apoptosis) by activating pathways that promote cell survival indirectly contributing to increased cell numbers by preventing cell death (Juma et al., 2016). The PLAG1 gene exhibits the capacity to collaborate with other oncogenes or to antagonise tumour suppressors, thereby facilitating cellular proliferation.

The MTS assay represents a widely employed colourimetric technique utilised for assessing cellular viability and proliferation, commonly employed within both research and clinical

contexts. Specifically, this method involves the introduction of a tetrazolium compound known as MTS into the cell culture medium as cells that are metabolically active express enzymes capable of converting the MTS substrate into a soluble, coloured formazan product. This enzymatic conversion serves as a reliable indicator of cellular activity as the abundance of formazan generated is directly proportional to the population of viable cells. Subsequently, the optical density of the resultant colour is measured using a microplate reader.

The results of this assay conclusively indicate that the overexpression of PLAG1 elicited a threefold increase in cell viability and proliferation with a **p-value less than 0.0001 compared to the EV group. This is in agreement with the previously described proliferative capacity of PLAG1 and is consistent with my results described in Chapter 4. The IPA analysis of the affected diseases and biological functions (section 4.4.5.5) suggested that PLAG1 has the potential to stimulate or activate cell proliferation through the significant p-values and a Z-score activation (Z-score > 2) of proliferation-related functions. These activated functions include, but are not limited to, the proliferation of blood cells, progenitor cells, stem cells, and endothelial cells. Additionally, the IPA data suggested a significant increase in cell viability, with 37 different proteins involved in this biological function.

An in vitro clonogenic cell survival assay was conducted to further investigate the capacity of a single cell to proliferate into a colony. This assay rigorously examines each cell within the population for its ability to undergo persistent division.

HEK293-293 cells are well-known for their robust growth and capacity to form colonies when plated at low density. This characteristic is evident in the control group, which comprises EV, mock-transfected cells. However, our findings indicate that the overexpression of PLAG1 in HEK293 cells resulted in a 1.35-fold increase in the cell survival and their ability to form colonies, as supported by the P-value of P= 0 0.0019 when compared to the EV group. This result not only stands on its own but also correlates with the findings of the PLAG1 downstream affected functions of the IPA analysis in chapter 4 (section 4.4.5.5). In this analysis, colony formation and cell survival functions were predicted to be significantly activated (Z-score > 2) in PLAG1-expressing cells compared to the control group, as shown in Figure 5.4. Our findings, therefore,

emphasise the fundamental role of PLAG1 as a principal driver in cell proliferation, viability and survival.

Apoptosis, also referred to as programmed cell death, is a well- orchestrated process during which a cell autonomously triggers a self-destruct sequence to systematically eliminate itself. This substantially benefits the organism by effectively removing undesirable or dysfunctional cells, preserving tissue homeostasis and impeding unregulated cellular proliferation.

The dysregulation of apoptosis has been implicated in various pathological conditions, including cancer, neurodegenerative diseases, and immune disorders (Sjöström and Bergh, 2001). The interplay between PLAG1 and apoptosis in pleomorphic adenoma presents a multifaceted scenario. PLAG1, identified as an oncogene, is recognised for its role in promoting tumorigenesis by stimulating cell proliferation, however, a previous study has suggested that PLAG1 may also exert an inhibitory influence on apoptosis in specific cellular contexts (Wang et al., 2020). This research was, however, performed using bovine muscle cells, and so further studies are needed to fully understand the effects of PLAG1 on apoptosis within the context of pleomorphic adenoma.

We analysed the effect of PLAG1 protein on apoptosis after staining the cells with annexin V and PI and then subjecting to flow cytometry. Annexin V binds to phosphatidylserine, a molecule that relocates from the inner to the outer layer of the plasma membrane during apoptosis. This makes annexin V a marker for the early and late stages of apoptosis. On the other hand, PI is a fluorescent intercalating agent that is excluded by viable cells but can penetrate the plasma membrane of cells that have lost membrane integrity, such as late apoptotic or necrotic cells. When used together, annexin V and PI staining can distinguish different stages of apoptosis as early apoptotic cells are identified as annexin V-positive and PI-negative while cells in late apoptosis are positive for both annexin V and PI.

Our data demonstrated that the exogenous expression of PLAG1 did not yield a statistically significant change in the proportion of early and late apoptotic cells in comparison to the control group (NS $P > 0.05$)indicating that the introduction of PLAG1 does not induce apoptosis.

Due to time constraints, the apoptosis assay was performed 24 hours after transfection and thus the apparently similar apoptotic effect in both the PLAG1 and mock transfection groups might indicate the initial stages of PLAG1-associated inhibition of apoptosis. Tumorigenesis is a time-dependent, multistep process, and thus capturing a snapshot of events, particularly at an early stage, may not accurately represent the true cellular changes.

To address this, future studies replicating this assay for a longer period of time would potentially allow any effects of PLAG1 on cell apoptosis to be fully elucidated.

Despite the benign nature of pleomorphic adenomas they can grow into large tumours and display more aggressive behaviour with a slight tendency for malignant transformation. This makes our investigations into the impact of PLAG1 on the ability of cells to undergo migration and/or invasion essential in order to elucidate the molecular mechanisms implicated in the process of tumorigenesis.

A transwell migration assay was conducted to investigate the downstream effect of PLAG1 overexpression on cell migration. However, our data indicated that there was no significant difference in the migratory potential of PLAG1-expressing cells versus the control group; no migration was detected in either group. This is in line with the study conducted by (Dalin et al., 2017), which examined the influence of PLAG1 on the migration of myoepithelial carcinoma cells. Their findings indicated that expression of PLAG1 alone did not induce changes in cell migration, however, when combined or co-localized with another growth factor, such as TGFBR3, cell motility was positively impacted. These findings, however, were contradicted by the study of Goto et al. (2020) in which PLAG1 was seen to promote transwell migration and invasion in immortalised normal human salivary gland acinar (NS-SV-AC) and ductal (NS-SV-DC) cell lines.

The IPA downstream effect analysis section 4.4.5.5 (affected diseases and functions) indicated that numerous cell migration functions were statistically significant but remained inactivated (Z-score <2). This indicates that there is a significant overlap between the genes/proteins in our dataset and the genes associated with cell migration. Nevertheless, the direction of change (activation or inhibition) is not strong enough to confidently predict the activation state of cell migration. Only one function annotation, migration of vascular endothelial cells (VECs), was

significantly increased with a Z-score of activation equal to 2.138. As mentioned earlier, Pleomorphic adenoma is a benign tumour that can grow large. Increased migration of VECs promotes tumour rapid growth by supporting angiogenesis (Dudley, 2012; Leone et al., 2024). In the case of Ca Ex PA, migrating VECs play a role in forming new blood vessels at distant sites, which helps spread tumour cells to other parts of the body (Suzuki et al., 2021).

A Transwell invasion assay was conducted to assess the impact of PLAG1 protein on cell invasion, however, no invasive cells were detected in either of the comparison groups (PLAG1 versus mock-transfected cells). This is in agreement with our previous findings (section 4.4.5.5), in which PLAG1 would be expected to significantly inhibit cellular functions related to cell invasion and metastasis with a Z-score<2. Based on these findings, it can be postulated that PLAG1 does not act as the exclusive regulator of cell invasion, however, its influence on other biological processes, particularly cell proliferation, suggests a broader functional scope. This is in agreement with a study by Huang et al., (2020) which examined the involvement of PLAG1 in ovarian cancer and found that the silencing of PLAG1 led to a reduction in the expression of IGF2, IGF1 receptor, and insulin receptor substrate 1, along with impeded proliferation, migration, and invasion of ovarian cells. Furthermore, the study revealed that the impact of PLAG1 was reliant on IGF2.

Currently, there is a lack of direct and conclusive evidence supporting PLAG1 as a definitive driver of cell migration and invasion and so any correlation between PLAG1 and cell migration and invasion likely operates through intermediary pathways.

In conclusion, while PLAG1 may not exclusively govern cell migration, invasion, and apoptosis, its impact on cell proliferation implies a broader significance.

6 Chapter 6 General Discussion

Salivary gland tumours represent a set of rare tumours, but their effects can significantly impact a patients' quality of life. Some types of salivary gland tumours are very aggressive and may have detrimental effect on patients' lives (Young et al., 2023).

In Europe and the UK, rare cancers are defined as those with an incidence of fewer than 6 per 100,000 per year and collectively, they make up about 24% of all cancer cases diagnosed annually Cancer Research UK (2024).

Rare tumours often have unique genetic and molecular characteristics that can provide insights into cancer biology as discoveries made in investigations into rare tumours can reveal mechanisms that are also relevant to more common cancers. Additionally, studying rare tumours contributes to the broader field of personalised medicine as through understanding the molecular profile of these tumours, treatments can be tailored to individual patients, enhancing efficiency and reducing side effects.

PA is the most common salivary gland tumour, representing two thirds of all salivary gland tumours. Despite being benign, they can become large and rarely will metastasise, a condition known as metastasising PA; this occurs even though the primary tumour is histologically benign. If left untreated, benign PAs can undergo malignant transformation to Ca ex PA, which is an aggressive malignancy with poor prognosis (Matsumiya-Matsumoto et al., 2022).

PLAG1 is a transcriptional factor that is frequently rearranged and overexpressed in PA. It is involved in various cellular processes such as proliferation and differentiation, and it has been proposed that dysregulation of PLAG1 may contribute to tumour development and progression (Martins et al., 2005). The exact mechanisms and pathways by which PLAG1 affect the biology and behaviour of PA are, however, not yet fully understood. The molecular landscape of PA and Ca ex PA is complex and evolving and whilst PLAG1 overexpression is considered a hallmark of PA, the acquisition of additional genetic and epigenetic alterations is thought to be crucial for malignant transformation. A deeper understanding of the molecular mechanisms underlying these tumours is essential for developing targeted therapies and improving patients' outcomes.

The aims of this study were to provide a comprehensive overview of the molecular role of PLAG1 in PA. Transcriptomic analysis facilitated an investigation into the impact of PLAG1 on downstream target genes to identify the altered functions and pathways Additionally, using biotin-based proximity labelling, my study aimed to identify proteins that directly interact with PLAG1, offering insight into the molecular interactions and pathways influenced by PLAG1 from a different perspective. Furthermore, functional assays were performed to investigate the effect of PLAG1 on cell behaviour. This approach also helps relate the findings from the biotin-based proximity labelling, providing a comprehensive understanding of the functions altered by PLAG1 and its interacting proteins.

The transcriptomic analysis outlined in Chapter 3 revealed a comprehensive picture of the role of PLAG1 in tumour development; specifically the gene set enrichment highlighted several key areas:

Firstly, **extracellular matrix and structural components** were prominently enriched. Terms related to collagen and extracellular matrix components such as cellular responses to retinoic acid, suggest that PLAG1 significantly impacts the tumour's structural and environmental context.

Secondly, **cell signalling and growth regulation** pathways were also notably enriched including growth factor binding and activity, Notch binding, and key signalling pathways such as Hedgehog, Wnt and cell adhesion molecules. The enrichment in cell adhesion molecules is significant as it suggests that PLAG1 may influence how tumour cells adhere to each other and to the extracellular matrix, impacting tumour cohesion and potentially affecting metastasis. Overall, these results indicate that PLAG1 may play a critical role in regulating essential signalling pathways that control cell growth, differentiation and communication within the tumour.

The oncogenic signature gene set analysis reinforces the benign nature of the tumour, showing reduced stem-like properties and low KRAS activity, which aligns with its slow-growing, differentiated phenotype. The downregulation of **ESC_V6.5_UP_EARLY.V1_DN** suggests a shift away from pluripotency, while the enrichment of **KRAS.600_UP.V1_DN** indicates that KRAS signalling is not a key driver, with potential compensation from pathways like Wnt/β-catenin or PI3K/AKT. Additionally, enrichment in DNA repair and immune response pathways, particularly

TGFB, highlights a balance between genomic stability and potential tumour progression. While PA remains largely non-invasive, factors such as ECM remodelling, Wnt/LEF1 activation, and TGFB signalling may contribute to recurrence and warrant further investigation.

The IPA analysis of the biotin-based proximity labelling outlined in Chapter 4 provided valuable insights into the role of PLAG1 in PA by examining various molecular aspects. **Canonical pathways** revealed crucial information about the regulatory environment influenced by PLAG1. The activation of the **HOX gene and RUNX1 pathways** indicates that PLAG1 plays a significant role in cellular development and differentiation. The anterior HOX pathway with its extensive crosstalk with other pathways, suggests it is central in regulating various molecular processes in the tumour. Meanwhile, the inhibition of the **PPARα/RXRα and Sumoylation pathways** highlights how PLAG1 may alter metabolic and stress response pathways, potentially affecting tumour metabolism and resilience.

Upstream regulator analysis focused on regulators that had expression data available, as they represent proteins interacting with PLAG1. These upstream regulators were biotinylated due to their proximity to PLAG1, making them detectable by the proximity labelling technique. Their significance lies in their role in mediating the effects of PLAG1, shaping the downstream pathways and cellular processes influenced by PLAG1, thereby contributing to the overall regulatory network within the tumour.

Upstream regulators such as **MYB, JUN, EP300 and CREBBP** were identified as crucial factors that regulate the transcriptional landscape in PA. MYB and JUN, known for their roles in cell proliferation and differentiation, interact with each other and share common downstream effectors, suggesting they coordinate key processes in tumour growth. EP300 and CREBBP further modulate these processes, influencing gene expression related to cell survival and proliferation.

Mechanistic network analysis underscored the interaction between MYB and JUN, highlighting their role in regulating downstream targets related to tumour development. This network provides insight into how PLAG1 might influence tumour biology by modulating these key transcription factors and their effect on cellular processes.

The **downstream effect analysis** (affected functions) revealed that PLAG1 significantly impacts processes related to cellular development and growth, particularly those associated with endothelial cells, blood cells and blood cell formation. Key functions include angiogenesis and vasculogenesis which are crucial for supporting tumour growth and expansion. Additionally, PLAG1 influences cellular processes related to cell viability and survival, which support tumour persistence and resistance to therapies. However, the data also show that while PLAG1 enhances cell survival, it inhibits metastasis and apoptosis, indicating a complex role in balancing tumour growth and spread.

Regulator effect analysis identified several master upstream regulators such as IGF1, EGF, IL6, IL33, SENCA, ERK1/2, HNF4A, and mircoRNA (mir 8). Although these regulators were activated but not directly expressed in the data, their significance lies in their potential impact on tumour biology. Growth factors such as IGF1 and EGF suggest a role in promoting cell growth, while cytokines like IL6 and IL33 may influence tumour microenvironment. Enzymes and signalling pathways like SCNA and ERK1/2 further indicate a complex regulatory network involving PLAG1. MicroRNA (mir8) and HNF4A highlight additional layer of regulation that could affect tumour progression.

Overall, IPA analysis provides a comprehensive view of how PLAG1 influences PA, highlighting its role in activating pathways and impacting cellular processes related to growth and survival.

In studying the effect of PLAG1 on cell behaviour (chapter 5), several assays provided insights into its role in PA. The proliferation assay revealed a significant increase in cell survival and proliferation, aligning with the biotin-based proximity labelling results that PLAG1 promotes cell growth. Similarly, the colony formation assay showed increased colony formation, supporting the PLAG1in enhancing tumour cells proliferation.

However, migration and invasion assays were negative, suggesting that early stage PA driven by PLAG1 may not support these processes. It is possible to speculate that PLAG1's influence on cell migration and invasion might be more relevant during later stages of tumour progression or transformation into a more aggressive form such as Ca ex PA. At these later stages, altered signalling pathways or additional mutations might possibly activate migration and invasion.

The apoptosis assay, conducted only 24 hours after transfection, did not yield conclusive results, likely because the short timeframe may not have been sufficient to observe the effects of PLAG1 on cell death. It is possible that PLAG1 requires more time to influence apoptotic pathways. Therefore, performing the assay at multiple time points over a longer period could provide a more comprehensive understanding on how PLAG1 affects apoptosis, offering better alignment with the proximity labelling results.

Based on the comprehensive data from transcriptomic analysis, IPA analysis, and functional assays, PLAG1 demonstrates a **dual role** in PA. Initially, PLAG1 acts primarily as an oncogene, driving tumour growth by enhancing cell proliferation, survival, and the formation of the extracellular matrix. This role supports the benign nature of PA by promoting localised tumour growth while maintaining encapsulation, a characteristic that prevents the tumour from invading surrounding tissues. For instance, the upregulation of pathways related to cell growth, extracellular matrix components and extracellular encapsulating structures helps maintain the tumour's non-invasive, well-defined boundary.

In the early stages of PA, PLAG1's influence is primarily on pathways that support benign characteristics. The absence of significant effects on cell migration and invasion in the functional assays aligns with this benign behaviour. The encapsulation of the tumour, supported by PLAG1's regulation of extracellular matrix components and cell adhesion molecules, helps confine the tumour and prevent metastasis.

However, this dual role suggests that PLAG1's function may evolve as the tumour progresses. In rare instances, additional genetic changes or disruptions in regulatory networks can lead to a shift in PLAG1's role. For example, changes in the balance of signalling pathways, interactions with other regulatory factors, or loss of suppressive mechanisms could enable the tumour to acquire invasive and metastatic capabilities. Although PLAG1 initially supports benign growth, these alterations might eventually allow the tumour to overcome its encapsulated nature and spread.

In summary, while PLAG1 primarily supports the benign nature of PA through its early-stage effects on growth and encapsulation, its potential for malignancy in rare cases highlights the complexity of its role. This dual function underscores the importance of understanding how

PLAG1's regulatory mechanisms can shift from benign to malignant states and emphasises the need for further research into the conditions that trigger this transition.

6.1 Conclusion and summary of key findings

The integration of transcriptomic analysis, IPA analysis of biotin-based proximity labelling, and functional assays provides a comprehensive overview of the molecular landscape downstream of PLAG1 overexpression in PA. This approach elucidates PLAG1's complex role in tumour behaviour, supporting benign growth and potentially contributing to malignancy in certain cases.

Dual Role of PLAG1: Functions as an oncogene promoting benign growth through enhanced cell proliferation and extracellular matrix formation, with potential for malignant behaviour in rare cases.

Affected Functions and Pathways: Significant effects on cell proliferation and extracellular matrix formation. Key pathways activated include Hedgehog, Wnt, and anterior Hox genes, with inhibition of PPAR alpha and RXR alpha pathways.

Upstream Regulators: MYB, JUN, EP300, and CREBBP are crucial regulators interacting with PLAG1, impacting its downstream effects and potentially influencing tumour progression.

Mechanistic Insights: MYB and JUN interact in a network affecting cellular development and growth, highlighting their role in regulating tumour behaviour.

These findings offer crucial insights into how PLAG1 influences PA and underscore the need for further research into its role in tumour progression. This study provides a general overview of the mechanisms of PLAG1, paving the way for future studies to explore its role in tumour progression and transformation.

6.2 Future work

To build on the findings from the transcriptomic analysis, IPA data, and functional assays regarding PLAG1 in PA, several avenues for future research are recommended:

1. Co-localisation Studies: Using immunohistochemistry to co-localise PLAG1 with its proximal interacting proteins in patient samples of pleomorphic adenoma and carcinoma. This will help visualise the spatial distribution and expression patterns of these proteins in relation to PLAG1.
2. Immunoprecipitation: Employing immunoprecipitation to further study protein-protein interactions. This technique involves pulling down PLAG1 and any interacting proteins from cell lysates using specific antibodies. The co-precipitated proteins can then be analysed to identify and characterise additional interactors.
3. Validation of Key Findings: Performing additional validation experiments to confirm the observed effects of PLAG1 on cell proliferation, extracellular matrix formation, and key signalling pathways. This could involve using different cell lines or animal models.
4. Extended Functional Assays: Conducting long-term functional assays, including migration and invasion assays at various time points, to better understand the role of PLAG1 in tumour progression and potential malignant transformation.
5. Detailed Mechanistic Studies: Exploring in greater depth the interaction between PLAG1 and identified upstream regulators like MYB and JUN. Investigating how these interactions influence downstream signalling pathways and tumour behaviour.
6. Broader Molecular Profiling: Expanding the molecular profiling to include proteomics and metabolomics to gain a more comprehensive understanding of how PLAG1 affects tumour biology.
7. Clinical Correlation: Correlating the findings with clinical data to assess the relevance of PLAG1 expression and associated pathways in patient outcomes and disease progression.
8. Investigate Dual Role of PLAG1: Examining the conditions under which PLAG1 shifts from supporting benign tumour growth to potentially driving malignancy. This includes studying the transition from benign to malignant states and identifying additional genetic or environmental factors involved.

These steps will help further elucidate the role of PLAG1 in pleomorphic adenoma and provide insights into its potential as a therapeutic target or biomarker for tumour progression.

6.3 Study strengths:

Comprehensive Approach: The integration of transcriptomic analysis, biotin-based proximity labelling, and functional assays provides a thorough understanding of PLAG1's role in pleomorphic adenoma. By combining these methods, the study offers a broad perspective on gene expression, protein interactions, and cellular behaviour.

Detailed Pathway Analysis: The use of IPA to identify canonical pathways and upstream regulators allows for in-depth insights into how PLAG1 influences tumour biology. This includes understanding specific signalling pathways like Hedgehog and Wnt, which are crucial in tumour progression.

Functional Validation: Functional assays such as proliferation, colony formation, and apoptosis provide direct evidence of PLAG1's effects on cellular processes. These assays help validate the findings from transcriptomic data and protein interaction studies, confirming their biological relevance.

Advanced Techniques: The use of biotin-based proximity labelling technique enables the identification of PLAG1 interacting proteins. This approach enhances the understanding of PLAG1's molecular interactions and its impact on tumorigenesis.

6.4 Study limitations:

Limited Long-Term Data: Functional assays like apoptosis were only performed 24 hours after transfection, which may not capture the full extent of PLAG1's effects over time. Longer-term studies are needed to assess the full impact of PLAG1 on cell behaviour.

Cell Line and Model Variability: Results obtained from cell lines may not fully represent the complexity of human tumours. Differences in cell line characteristics and the lack of human tissue diversity might limit the generalisability of the findings.

Incomplete Interaction Data: While biotin-based proximity labelling provides valuable information on protein interactions, it may not capture all potential interactors or the dynamic nature of these interactions in different cellular contexts.

Lack of Clinical Correlation: The study does not include direct analysis of patient samples beyond initial observations. Correlating findings with clinical outcomes and tissue samples would strengthen the relevance and applicability of the results.

These strengths and limitations highlight the study's comprehensive approach to understanding PLAG1 in pleomorphic adenoma while also pointing out areas where further research and refinement are needed.

7 References

Abdel-Hafiz, M., Abdel-Wahab, M. F., El-Naggar, A. K., & El-Naggar, S. K. (2012). PLAG1 overexpression in pleomorphic adenomas of the salivary glands: A potential diagnostic and prognostic marker. Modern Pathology, 25(1), 118-127.

Adini, A., Kornaga, T., Firoozbakht, F. and Benjamin, L.E., 2002. Placental growth factor is a survival factor for tumor endothelial cells and macrophages. Cancer research, 62(10), pp.2749-2752.

Albonici, L., Giganti, M.G., Modesti, A., Manzari, V. and Bei, R., 2019. Multifaceted role of the placental growth factor (PlGF) in the antitumor immune response and cancer progression. International Journal of Molecular Sciences, 20(12), p.2970.

Amber-Vitos, O., Chaturvedi, N., Nachliel, E., Gutman, M. and Tsfadia, Y., 2016. The effect of regulating molecules on the structure of the PPAR-RXR complex. *Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids*, *1861*(11), pp.1852-1863.

Andrae, J., Gallini, R. and Betsholtz, C., 2008. Role of platelet-derived growth factors in physiology and medicine. *Genes & development*, *22*(10), pp.1276-1312.

Andreadis, D., Poulopoulos, A., Epivatianos, A., Nomikos, A., Parlitsis, D., Christidis, K., Barbatis, C., Kavvadas, D., Toskas, A., Papamitsou, T. and Antoniades, D., 2020. Cell adhesion molecules' altered profile in benign and malignant salivary gland tumors. The paradigm of beta4-integrin, desmoglein-2, ICAM-1 and CD44s. *Journal of Biological Research-Thessaloniki*, *27*, pp.1-10.

Anfossi, G., Gewirtz, A.M. and Calabretta, B., 1989. An oligomer complementary to c-myb-encoded mRNA inhibits proliferation of human myeloid leukemia cell lines. Proceedings of the National Academy of Sciences, 86(9), pp.3379-3383.

Antony, J., Gopalan, V., Smith, R.A. and Lam, A.K., 2012. Carcinoma ex pleomorphic adenoma: a comprehensive review of clinical, pathological and molecular data. *Head and neck pathology*, *6*, pp.1-9.

Antony, J., Gopalan, V., Smith, R.A. and Lam, A.K., 2012. Carcinoma ex pleomorphic adenoma: a comprehensive review of clinical, pathological and molecular data. *Head and neck pathology*, *6*, pp.1-9.

Arany, Z., Huang, L.E., Eckner, R., Bhattacharya, S., Jiang, C., Goldberg, M.A., Bunn, H.F. and Livingston, D.M., 1996. An essential role for p300/CBP in the cellular response to hypoxia. *Proceedings of the National Academy of Sciences*, *93*(23), pp.12969-12973.

Åström A, d'Amore ES, Sainati L, Panarello C, Morerio C, Mark J & Stenman G 2000 Evidence of involvement of the PLAG1 gene in lipoblastomas. International Journal of Oncology 16 1107–1117.

Åström A-K, Voz ML, Kas K, Röijer E, Wedell B, Mandahl N, Van de Ven W, Mark J & Stenman G 1999 Conserved mechanism of PLAG1 activation in salivary gland tumors with and without chromosome 8q12 abnormalities identification of SII as a new fusion partner gene. Cancer Research 59 918–923.

Auclair, P.L., 1991. Salivary gland neoplasms: general considerations. Surgical pathology of the salivary glands, pp.135-164.

Baker, J., Liu, J.P., Robertson, E.J. and Efstratiadis, A., 1993. Role of insulin-like growth factors in embryonic and postnatal growth. *Cell*, *75*(1), pp.73-82.

Baredes, S., Ludwin, D.B., Troublefield, Y.L., Langer, P.D. and Mirani, N., 2003. Adenocarcinoma ex-pleomorphic adenoma of the lacrimal sac and nasolacrimal duct: a case report. The Laryngoscope, 113(6), pp.940-942.

Barnes, L., 2005. Pathology and genetics of head and neck tumours. *World Health Organization classification of tumours/IARC Press*.

Berkovitz, B.K., Holland, G.R. and Moxham, B.J., 1992. Color atlas & textbook of oral anatomy, histology, and embryology. Mosby Inc.

Bhatlekar, S., Fields, J.Z. and Boman, B.M., 2018. Role of HOX genes in stem cell differentiation and cancer. *Stem Cells International*, *2018*(1), p.3569493.

Blau, L., Knirsh, R., Ben-Dror, I., Oren, S., Kuphal, S., Hau, P., Proescholdt, M., Bosserhoff, A.K. and Vardimon, L., 2012. Aberrant expression of c-Jun in glioblastoma by internal ribosome entry site (IRES)-mediated translational activation. Proceedings of the National Academy of Sciences, 109(42), pp.E2875-E2884.

Boncinelli, E. and Mallamaci, A., 1995. Homeobox genes in vertebrate gastrulation. *Current opinion in genetics & development*, *5*(5), pp.619-627.

Bossy-Wetzel, E., Bravo, R. and Hanahan, D., 1992. Transcription factors junB and c-jun are selectively up-regulated and functionally implicated in fibrosarcoma development. Genes & development, 6(12a), pp.2340-2351.

Boveri, T. (1914). “Zur Frage der Entstehung maligner Tumoren.” *Jena: Gustav Fischer*.

Brill, L.B., Kanner, W.A., Fehr, A., Andrén, Y., Moskaluk, C.A., Löning, T., Stenman, G. and Frierson, H.F., 2011. Analysis of MYB expression and MYB-NFIB gene fusions in adenoid cystic carcinoma and other salivary neoplasms. *Modern pathology*, *24*(9), pp.1169-1176.

Burns, J.L. and Hassan, A.B., 2001. Cell survival and proliferation are modified by insulin-like growth factor 2 between days 9 and 10 of mouse gestation.

Cancer Research UK. (2022). Cancer Statistics 2022. Cancer Research UK. London, UK.

Carlsson, P., Borg, Å., Påhlman, S., & Isaksson, M. (2009). PLAG1 is a key oncogene in pleomorphic adenoma of the salivary glands. Cancer Research, 69(11), 4666-4674.

Caro, I. and Low, J.A., 2010. The role of the hedgehog signaling pathway in the development of basal cell carcinoma and opportunities for treatment. *Clinical Cancer Research*, *16*(13), pp.3335-3339.

Ceci, C., Atzori, M.G., Lacal, P.M. and Graziani, G., 2020. Role of VEGFs/VEGFR-1 signaling and its inhibition in modulating tumor invasion: experimental evidence in different metastatic cancer models. International journal of molecular sciences, 21(4), p.1388.

Chaudhry, A.P., Leifer, C., Cutler, L.S., Satchidanand, S., Labay, G.R. and Yamane, G.M., 1986. Histogenesis of adenoid cystic carcinoma of the salivary glands: Light and electron microscopic study. Cancer, 58(1), pp.72-82.

Chen, W., Gu, T., Chen, Q., Qu, C., Zhang, C., Hu, Y., Xia, R., Zhang, Y., Wang, M., Huang, X. and Li, J., 2023. Extracellular matrix remodelling and stiffening contributes to tumorigenesis of salivary carcinoma ex pleomorphic adenoma - A study based on patient-derived organoids. *Cell & Bioscience*, *13*(1), p.122.

Chen, X., Wang, L., Yang, M., Zhao, W., Tu, J., Liu, B. and Yuan, X., 2024. RUNX transcription factors: biological functions and implications in cancer. *Clinical and Experimental Medicine*, *24*(1), pp.1-25.

Chen, Z., Chen, C., Li, L., Zhang, T. and Wang, X., 2021. Pan-cancer analysis reveals that E1A binding protein p300 mutations increase genome instability and antitumor immunity. Frontiers in Cell and Developmental Biology, 9, p.729927.

Chen, Z., Chen, J., Gu, Y., Hu, C., Li, J.L., Lin, S., Shen, H., Cao, C., Gao, R., Li, J. and Ha, P.K., 2014. Aberrantly activated AREG–EGFR signaling is required for the growth and survival of CRTC1–MAML2 fusion-positive mucoepidermoid carcinoma cells. Oncogene, 33(29), pp.3869-3877.

Cheung, L.W., Leung, P.C. and Wong, A.S., 2006. Gonadotropin-releasing hormone promotes ovarian cancer cell invasiveness through c-Jun NH2-terminal kinase–mediated activation of matrix metalloproteinase (MMP)-2 and MMP-9. *Cancer research*, *66*(22), pp.10902-10910.

Chicoteka, E. A., Silva, A. M., Soares, J., & Soares, J. P. (2019). Secretory carcinoma: A review of clinical, pathological, and molecular features. Pathology-Research and Practice, 215(1), 151447.

Chitturi RT, Veeravarmal V, Nirmal RM, Reddy VRB. 2015. Myoepithelial cells (MEC) of the salivary glands in health and tumours. J Clin Diagn Res 9:ZE14–ZE18.

Cho, K.J., El-Naggar, A.K., Mahanupab, P., Luna, M.A. and Batsakis, J.G., 1995. Carcinoma ex pleomorphic adenoma of the nasal cavity: a report of two cases. The Journal of Laryngology & Otology, 109(7), pp.677-679.

Chuang, L.S.H., Ito, K. and Ito, Y., 2013. RUNX family: Regulation and diversification of roles through interacting proteins. *International journal of cancer*, *132*(6), pp.1260-1271.

Clevers, H. and Nusse, R., 2012. Wnt/β-catenin signalling and disease. Cell, 149(6), pp.1192-1205.

Corio, R.L., Sciubba, J.J., Brannon, R.B. and Batsakis, J.G., 1982. Epithelial-myoepithelial carcinoma of intercalated duct origin: a clinicopathologic and ultrastructural assessment of sixteen cases. Oral Surgery, Oral Medicine, Oral Pathology, 53(3), pp.280-287.

Crissman, J.D., Wirman, J.A. and Harris, A., 1977. Malignant myoepithelioma of the parotid gland. Cancer, 40(6), pp.3042-3049.

Daftary, G.S. and Taylor, H.S., 2006. Endocrine regulation of HOX genes. *Endocrine reviews*, *27*(4), pp.331-355.

Dalin, M.G., Katabi, N., Persson, M., Lee, K.W., Makarov, V., Desrichard, A., Walsh, L.A., West, L., Nadeem, Z., Ramaswami, D. and Havel, J.J., 2017. Multi-dimensional genomic analysis of myoepithelial carcinoma identifies prevalent oncogenic gene fusions. Nature communications, 8(1), pp.1-13.

Dalin, M.G., Katabi, N., Persson, M., Lee, K.W., Makarov, V., Desrichard, A., Walsh, L.A., West, L., Nadeem, Z., Ramaswami, D. and Havel, J.J., 2017. Multi-dimensional genomic analysis of myoepithelial carcinoma identifies prevalent oncogenic gene fusions. *Nature communications*, *8*(1), p.1197.

Damm, D.D. and Fantasia, J.E., 2001. Large palatal mass. Carcinoma ex-pleomorphic adenoma. General dentistry, 49(6), pp.574-658.

Darabi, S., Zuazo, C.E., Braxton, D.R., Eisenberg, B.L. and Demeure, M.J., 2023. Precision medicine in a community cancer center: pan-cancer DNA/RNA sequencing of tumors reveals clinically relevant gene fusions. *Biologics*, *3*(3), pp.198-208.

Dardick, I. and Van Nostrand, A.P., 1985. Myoepithelial cells in salivary gland tumours— revisited. Head & neck surgery, 7(5), pp.395-408.

Dardick, I., Daya, D., Hardie, J. and Van Nostrand, A.W.P., 1984. Mucoepidermoid carcinoma: ultrastructural and histogenetic aspects. Journal of Oral Pathology & Medicine, 13(4), pp.342 358.

Dardick, I., Van Nostrand, A.P., Jeans, M.D., Rippstein, P. and Edwards, V., 1983. Pleomorphic adenoma, I: ultrastructural organization of “epithelial” regions. Human pathology, 14(9), pp.780-797.

de Brito, B.S., Giovanelli, N., Egal, E.S., Sánchez-Romero, C., do Nascimento, J.D.S., Martins, A.S., Tincani, Á.J., Del Negro, A., de Oliveira Gondak, R., de Almeida, O.P. and Kowalski, L.P., 2016. Loss of expression of Plag1 in malignant transformation from pleomorphic adenoma to carcinoma ex pleomorphic adenoma. *Human pathology*, *57*, pp.152-159.

de Lima-Souza, R.A., Bělohlávková, K., Michal, M., Altemani, A., Mariano, F.V. and Skálová, A., 2025. Atypical and worrisome histological features in pleomorphic adenoma: challenging and potentially significant diagnostic pitfall. Virchows Archiv, pp.1-13.

DeChiara, T.M., Efstratiadis, A. and Robertsen, E.J., 1990. A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. *Nature*, *345*(6270), pp.78-80.

di Magliano, M.P. and Hebrok, M., 2003. Hedgehog signalling in cancer formation and maintenance. *Nature reviews cancer*, *3*(12), pp.903-911.

Ding, N., You, A.B., Yang, H., Hu, G.S., Lai, C.P., Liu, W. and Ye, F., 2023. A tumor-Suppressive molecular axis EP300/circRERE/miR-6837-3p/MAVS activates type I IFN pathway and antitumor immunity to suppress colorectal cancer. Clinical Cancer Research, 29(11), pp.2095-2109.

Diz, A.P., Carvajal-Rodríguez, A. and Skibinski, D.O., 2011. Multiple hypothesis testing in proteomics: a strategy for experimental work. *Molecular & Cellular Proteomics*, *10*(3).

Dobrotkova, V., Chlapek, P., Mazanek, P., Sterba, J. and Veselska, R., 2018. Traffic lights for retinoids in oncology: molecular markers of retinoid resistance and sensitivity and their use in the management of cancer differentiation therapy. *BMC cancer*, *18*, pp.1-13.

Drabsch, Y., Hugo, H., Zhang, R., Dowhan, D.H., Miao, Y.R., Gewirtz, A.M., Barry, S.C., Ramsay, R.G. and Gonda, T.J., 2007. Mechanism of and requirement for estrogen-regulated MYB expression in estrogen-receptor-positive breast cancer cells. Proceedings of the National Academy of Sciences, 104(34), pp.13762-13767.

Dudley, A.C., 2012. Tumor endothelial cells. *Cold Spring Harbor perspectives in medicine*, *2*(3), p.a006536.

Dulguerov, P., Todic, J., Pusztaszeri, M. and Alotaibi, N.H., 2017. Why do parotid pleomorphic adenomas recur? A systematic review of pathological and surgical variables. Frontiers in surgery, 4, p.26.

Dyson, P.J., Poirier, F. and Watson, R.J., 1989. Expression of c-myb in embryonal carcinoma cells and embryonal stem cells. *Differentiation*, *42*(1), pp.24-27.

Eberly, H.W., Sciscent, B.Y., Lorenz, F.J., Rettig, E.M. and Goyal, N., 2024. Current and Emerging Diagnostic, Prognostic, and Predictive Biomarkers in Head and Neck Cancer. *Biomedicines*, *12*(2), p.415.

Eckner, R., Ewen, M.E., Newsome, D., Gerdes, M., DeCaprio, J.A., Lawrence, J.B. and Livingston, D.M., 1994. Molecular cloning and functional analysis of the adenovirus E1A-associated 300-kD protein (p300) reveals a protein with properties of a transcriptional adaptor. Genes & development, 8(8), pp.869-884.

Egeblad, M., Rasch, M.G. and Weaver, V.M., 2010. Dynamic interplay between the collagen scaffold and tumor evolution. *Current opinion in cell biology*, *22*(5), pp.697-706.

Ellis GL, Auclair PL. 2008. AFIP atlas of tumour pathology: The normal salivary glands, 4th ed. Maryland: ARP Press.

El-Naggar, A.K., Callender, D., Coombes, M.M., Hurr, K., Luna, M.A. and Batsakis, J.G., 2000. Molecular genetic alterations in carcinoma ex-pleomorphic adenoma: A putative progression model?. Genes, Chromosomes and Cancer, 27(2), pp.162-168.

El-Naggar, A.K., Chan, J.K., Grandis, J.R., Takata, T. and Slootweg, P.J. eds., 2017. WHO classification of head and neck tumours. International Agency for Research on Cancer (IARC).

El-Tanani, M., Rabbani, S.A., Satyam, S.M., Rangraze, I.R., Wali, A.F., El-Tanani, Y. and Aljabali, A.A., 2025. Deciphering the role of cancer stem cells: Drivers of tumour evolution, therapeutic resistance, and precision medicine strategies. Cancers, 17(3), p.382.

Eneroth, C.M., 1971. Salivary gland tumors in the parotid gland, submandibular gland, and the palate region. Cancer, 27(6), pp.1415-1418.

Enescu, A.Ş., Enescu, A., Bălăşoiu, M.A.R.I.A., Ciolofan, M.S. and Căpitănescu, A.N., 2014. Histopathological study of pleomorphic adenoma of salivary glands. *Romanian Journal of Morphology and Embryology= Revue Roumaine de Morphologie et Embryologie*, *55*(3 Suppl), pp.1149-1153.

Enescu, A.Ş., Enescu, A., Bălăşoiu, M.A.R.I.A., Ciolofan, M.S. and Căpitănescu, A.N., 2014. Histopathological study of pleomorphic adenoma of salivary glands. *Romanian Journal of*

Morphology and Embryology= Revue Roumaine de Morphologie et Embryologie, *55*(3 Suppl), pp.1149-1153.

Feige, J.N., Gelman, L., Michalik, L., Desvergne, B. and Wahli, W., 2006. From molecular action to physiological outputs: peroxisome proliferator-activated receptors are nuclear receptors at the crossroads of key cellular functions. *Progress in lipid research*, *45*(2), pp.120-159.

Ferrari, N., Mohammed, Z. M. A., Nixon, C., Mason, S. M., Mallon, E., McMillan, D. C., … Blyth, K. (2014). Expression of RUNX1 correlates with poor patient prognosis in triple negative breast cancer. PLoS One, 9(6), e100759. https://doi.org/10.1371/journ al.pone.0100759

Flotho, A. and Melchior, F., 2013. Sumoylation: a regulatory protein modification in health and disease. *Annual review of biochemistry*, *82*(1), pp.357-385.

Fonseca, D., Arya, S.S., Kodandapani, S., Chandini, A., Kurapati, S., Rao, C., Gadepalli, T. and Ali, Z., 2022. Metastasising Pleomorphic Adenoma: A Rare Entity. *Indian Journal of Otolaryngology and Head & Neck Surgery*, pp.1-3.

Font-Díaz, J., Jiménez-Panizo, A., Caelles, C., dM Vivanco, M., Pérez, P., Aranda, A., Estébanez-Perpiñá, E., Castrillo, A., Ricote, M. and Valledor, A.F., 2021, August. Nuclear receptors: Lipid and hormone sensors with essential roles in the control of cancer development. In *Seminars in cancer biology* (Vol. 73, pp. 58-75). Academic Press.

Fougerat, A., Montagner, A., Loiseau, N., Guillou, H. and Wahli, W., 2020. Peroxisome proliferator-activated receptors and their novel ligands as candidates for the treatment of non-alcoholic fatty liver disease. *Cells*, *9*(7), p.1638.

Frantz, C., Stewart, K.M. and Weaver, V.M., 2010. The extracellular matrix at a glance. *Journal of cell science*, *123*(24), pp.4195-4200.

Furuta, S., Wang, J.M., Wei, S., Jeng, Y.M., Jiang, X., Gu, B., Chen, P.L., Eva, Y.H.L. and Lee, W.H., 2006. Removal of BRCA1/CtIP/ZBRK1 repressor complex on ANG1 promoter leads to accelerated mammary tumor growth contributed by prominent vasculature. *Cancer cell*, *10*(1), pp.13-24.

Futreal, P.A., Coin, L., Marshall, M., Down, T., Hubbard, T., Wooster, R., Rahman, N. and Stratton, M.R., 2004. A census of human cancer genes. Nature reviews cancer, 4(3), pp.177 183.

Gao, R., Cao, C., Zhang, M., Lopez, M.-C., Yan, Y., Chen, Z., … Kaye, F. J. (2014). A unifying gene signature for adenoid cystic cancer identifies parallel MYB-dependent and MYB-independent therapeutic targets. Oncotarget, 5(24), 12528–12542. https://doi.org/10.18632/ oncot arget.2985

Gayther, S.A., Batley, S.J., Linger, L., Bannister, A., Thorpe, K., Chin, S.F., Daigo, Y., Russell, P., Wilson, A., Sowter, H.M. and Delhanty, J.D., 2000. Mutations truncating the EP300 acetylase in human cancers. Nature genetics, 24(3), pp.300-303.

George, O.L. and Ness, S.A., 2014. Situational awareness: regulation of the myb transcription factor in differentiation, the cell cycle and oncogenesis. *Cancers*, *6*(4), pp.2049-2071.

Giancotti, F.G., 2014. Deregulation of cell signalling in cancer. *FEBS letters*, *588*(16), pp.2558-2570.

Gnepp, D.R., 1993. Malignant mixed tumors of the salivary glands: a review. Pathology annual, 28, pp.279-328.

Gonzalez, D.M. and Medici, D., 2014. Signaling mechanisms of the epithelial-mesenchymal transition. *Science signaling*, *7*(344), pp.re8-re8.

Goto, Y., Ibi, M., Sato, H., Tanaka, J., Yasuhara, R., Aota, K., Azuma, M., Fukada, T., Mishima, K. and Irié, T., 2020. PLAG1 enhances the stemness profiles of acinar cells in normal human salivary glands in a cell type-specific manner. *Journal of Oral Biosciences*, *62*(1), pp.99-106.

Guan, L., Del Carpio-Cano, F., Voora, D., Myers, R. and Rao, A.K., 2023. RUNX1 Isoforms Regulate RUNX1 and Target-Genes Differentially in Platelets/Megakaryocytes: Association with Clinical Cardiovascular Events. *Blood*, *142*, p.1359.

Guo, W., Lasky, J.L. and Wu, H., 2006. Cancer stem cells. Pediatric research, 59(4), pp.59-64.

Guzzo, M., Locati, L.D., Prott, F.J., Gatta, G., McGurk, M. and Licitra, L., 2010. Major and minor salivary gland tumors. *Critical reviews in oncology/hematology*, *74*(2), pp.134-148.

Handler, D.C. and Haynes, P.A., 2020. Statistics in proteomics: A meta-analysis of 100 proteomics papers published in 2019. *Journal of the American Society for Mass Spectrometry*, *31*(7), pp.1337-1343.

Hao, Y., Baker, D. and Ten Dijke, P., 2019. TGF-β-mediated epithelial-mesenchymal transition and cancer metastasis. International journal of molecular sciences, 20(11), p.2767.

Harjunpää, H., Llort Asens, M., Guenther, C. and Fagerholm, S.C., 2019. Cell adhesion molecules and their roles and regulation in the immune and tumor microenvironment. Front Immunol. 2019; 10: 1078.

Harmsen, M.C., Getova, V., Zhang, M., Zhao, F., van Dongen, J., Martinez Garcia, F.D., Nizamoglu, M. and Burgess, J.K., 2024. Organ-Derived Extracellular Matrix (ECM) Hydrogels: Versatile Systems to Investigate the Impact of Biomechanics and Biochemistry on Cells in Disease Pathology. In *Handbook of the Extracellular Matrix: Biologically-Derived Materials* (pp. 875-901). Cham: Springer International Publishing.

Hayes, M.M., Lesack, D., Girardet, C., Del Vecchio, M. and Eusebi, V., 2005. Carcinoma ex pleomorphic adenoma of the breast. Report of three cases suggesting a relationship to metaplastic carcinoma of matrix-producing type. Virchows Archiv, 446(2), pp.142-149.

Ho, A.S., Ochoa, A., Jayakumaran, G., Zehir, A., Mayor, C.V., Tepe, J., Makarov, V., Dalin, M.G., He, J., Bailey, M. and Montesion, M., 2019. Genetic hallmarks of recurrent/metastatic adenoid cystic carcinoma. *The Journal of clinical investigation*, *129*(10), pp.4276-4289.

Hoi, C. S. L., Lee, S. E., Lu, S.-Y., McDermitt, D. J., Osorio, K. M., Piskun, C. M., … Tumbar, T. (2010). Runx1 directly promotes proliferation of hair follicle stem cells and epithelial tumor formation in mouse skin. Molecular and Cellular Biology, 30(10), 2518–2536. https://doi.org/10.1128/MCB.01308-09

Holsinger, F.C. and Bui, D.T., 2007. Anatomy, function, and evaluation of the salivary glands. *Salivary gland disorders*, pp.1-16.

Hombría JC-G, Lovegrove B. Beyond homeosis-HOX function in morphogenesis and organogenesis. Differentiation. 2003;71(8):461–76.

Hu, Y.H., Zhang, C.Y., Xia, R.H., Tian, Z., Wang, L.Z. and Li, J., 2016. Prognostic factors of carcinoma ex pleomorphic adenoma of the salivary glands, with emphasis on the widely invasive carcinoma: a clinicopathologic analysis of 361 cases in a Chinese population. *Oral Surgery, Oral Medicine, Oral Pathology and Oral Radiology*, *122*(5), pp.598-608.

Huang, H.W., Chang, C.C., Wang, C.S. and Lin, K.H., 2021. Association between inflammation and function of cell adhesion molecules influence on gastrointestinal cancer development. *Cells*, *10*(1), p.67.

Huang, W., Li, B.R. and Feng, H., 2020. PLAG1 silencing promotes cell chemosensitivity in ovarian cancer via the IGF2 signaling pathway. International Journal of Molecular Medicine, 45(3), pp.703-714.

Humphrey, S.P. and Williamson, R.T., 2001. A review of saliva: normal composition, flow, and function. *The Journal of prosthetic dentistry*, *85*(2), pp.162-169.

Ianez RF, Buim ME, Coutinho-Camillo CM, Schultz R, Soares FA, Lourenc¸o SV. 2010. Human salivary gland morphogenesis: Myoepithelial cell maturation assessed by immunohistochemical markers. Histopathology 57:410–417.

Introna, M., Luchetti, M., Castellano, M., Arsura, M. and Golay, J., 1994, April. The myb oncogene family of transcription factors: potent regulators of hematopoietic cell proliferation and differentiation. In *Seminars in Cancer Biology* (Vol. 5, No. 2, pp. 113-124).

Ito, Y., Bae, S.C. and Chuang, L.S.H., 2015. The RUNX family: developmental regulators in cancer. *Nature Reviews Cancer*, *15*(2), pp.81-95.

Jiménez-Lara, A.M., Aranda, A. and Gronemeyer, H., 2010. Retinoic acid protects human breast cancer cells against etoposide-induced apoptosis by NF-kappaB-dependent but cIAP2-independent mechanisms. *Molecular cancer*, *9*, pp.1-16.

Johnson, N.W., Warnakulasuriya, S., Gupta, P.C., Dimba, E., Chindia, M., Otoh, E.C., Sankaranarayanan, R., Califano, J. and Kowalski, L., 2011. Global oral health inequalities in incidence and outcomes for oral cancer: causes and solutions. *Advances in dental research*, *23*(2), pp.237-246.

Juma, A.R., Damdimopoulou, P.E., Grommen, S.V., Van de Ven, W.J. and De Groef, B., 2016. Emerging role of PLAG1 as a regulator of growth and reproduction. Journal of Endocrinology, 228(2), pp.R45-R56.

Jun-shui, W.E.I., Xin, S.U.N., Jin-biao, X.U. and Yi-da, C.H.E.N., 2023. LncRNA RUNX1-IT1 regulating malignant pleomorphic adenoma via mir-195/CyclinD1. *Shanghai Journal of Stomatology*, *32*(1), p.85.

Kamil, M., Shinsato, Y., Higa, N., Hirano, T., Idogawa, M., Takajo, T., Minami, K., Shimokawa, M., Yamamoto, M., Kawahara, K. and Yonezawa, H., 2019. High filamin-C expression predicts enhanced invasiveness and poor outcome in glioblastoma multiforme. *British journal of cancer*, *120*(8), pp.819-826.

Kanatas, A., Ho, M.W.S. and Mücke, T., 2018. Current thinking about the management of recurrent pleomorphic adenoma of the parotid: a structured review. *British Journal of Oral and Maxillofacial Surgery*, *56*(4), pp.243-248.

Kanda, S., Mitsuyasu, T., Nakao, Y., Kawano, S., Goto, Y., Matsubara, R. and Nakamura, S., 2013. Anti-apoptotic role of the sonic hedgehog signaling pathway in the proliferation of ameloblastoma. *International journal of oncology*, *43*(3), pp.695-702.

Kandasamy, J., Smith, A., Diaz, S., Rose, B. and O'Brien, C., 2007. Heterogeneity of PLAG1 gene rearrangements in pleomorphic adenoma. Cancer genetics and cytogenetics, 177(1), pp.1-5.

Kanehisa, M., Furumichi, M., Sato, Y., Kawashima, M. and Ishiguro-Watanabe, M., 2023. KEGG for taxonomy-based analysis of pathways and genomes. *Nucleic acids research*, *51*(D1), pp.D587-D592.

Kas, K., Voz, M.L., Röijer, E., Åström, A.K., Meyen, E., Stenman, G. and Van de Ven, W.J., 1997. Promoter swapping between the genes for a novel zinc finger protein and β-catenin in

pleiomorphic adenomas with t (3; 8)(p21; q12) translocations. Nature genetics, 15(2), pp.170 174.

Katiyar, S., Casimiro, M.C., Dettin, L., Ju, X., Wagner, E.F., Tanaka, H. and Pestell, R.G., 2010. C-jun inhibits mammary apoptosis in vivo. *Molecular Biology of the Cell*, *21*(23), pp.4264-4274.

Katoh, M. and Katoh, M., 2007. WNT signaling pathway and stem cell signaling network. Clinical cancer research, 13(14), pp.4042-4045.

Kawczak, P., Feszak, I., Brzeziński, P. and Bączek, T., 2024. Structure–Activity Relationships and Therapeutic Applications of Retinoids in View of Potential Benefits from Drug Repurposing Process. *Biomedicines*, *12*(5), p.1059.

Keita, M., Bachvarova, M., Morin, C., Plante, M., Gregoire, J., Renaud, M.-C., … Bachvarov, D. (2013). The RUNX1 transcription factor is expressed in serous epithelial ovarian carcinoma and contributes to cell proliferation, migration and invasion. Cell Cycle, 12(6), 972–986. https://doi.org/10.4161/cc.23963

Kim, K.B., Kabra, A., Kim, D.W., Xue, Y., Huang, Y., Hou, P.C., Zhou, Y., Miranda, L.J., Park, J.I., Shi, X. and Bender, T.P., 2022. KIX domain determines a selective tumor-promoting role for EP300 and its vulnerability in small cell lung cancer. Science Advances, 8(7), p.eabl4618.

Kim, K.J., Cho, C.S. and Kim, W.U., 2012. Role of placenta growth factor in cancer and inflammation. Experimental & molecular medicine, 44(1), pp.10-19.

Kinsey, C.G., Camolotto, S.A., Boespflug, A.M., Guillen, K.P., Foth, M., Truong, A., Schuman, S.S., Shea, J.E., Seipp, M.T., Yap, J.T. and Burrell, L.D., 2019. Protective autophagy elicited by RAF→ MEK→ ERK inhibition suggests a treatment strategy for RAS-driven cancers. *Nature medicine*, *25*(4), pp.620-627.

Kleinsmith, L.J. and Martin, J. (2019) Principles of Cancer Biology. Jones & Bartlett Publishers.

Knight, J. and Ratnasingham, K., 2015. Metastasising pleomorphic adenoma: Systematic review. *International Journal of Surgery*, *19*, pp.137-145.

Koga, H., Sakisaka, S., Harada, M., Takagi, T., Hanada, S., Taniguchi, E., Kawaguchi, T., Sasatomi, K., Kimura, R., Hashimoto, O. and Ueno, T., 2001. Involvement of p21WAF1/Cip1, p27Kip1, and p18INK4c in troglitazone-induced cell-cycle arrest in human hepatoma cell lines. *Hepatology*, *33*(5), pp.1087-1097.

Kuehl, W.M., Bender, T.P., Stafford, J., McClinton, D., Segal, S. and Dmitrovsky, E., 1988. Expression and function of the c-myb oncogene during hematopoietic differentiation. In *Mechanisms in B-Cell Neoplasia 1988: Workshop at the National Cancer Institute, National Institutes of Health, Bethesda, MD, USA, March 23–25, 1988* (pp. 318-323). Springer Berlin Heidelberg.

Kumar, M., Molkentine, D. and Molkentine, J., 2020. CREBBP/EP300 mutation is associated with poor outcome in HNSCC and targetable with synthetic cytotoxicity. *bioRxiv*.

Landrette SF, Kuo Y-H, Hensen K, Barjesteh van Waalwijk van Doorn-Khosrovani S, Perrat PN, Van de Ven WJM, Delwel R & Castilla LH 2005 Plag1 and Plagl2 are oncogenes that induce acute myeloid leukemia in cooperation with Cbfb-MYH11. Blood 105 2900–2907.

Latysheva, N.S. and Babu, M.M., 2016. Discovering and understanding oncogenic gene fusions through data intensive computational approaches. Nucleic acids research, 44(10), pp.4487 4503.

Latysheva, N.S. and Babu, M.M., 2019. Molecular Signatures of Fusion Proteins in Cancer. ACS pharmacology & translational science, 2(2), pp.122-133.

Lavudi, K., Nuguri, S.M., Olverson, Z., Dhanabalan, A.K., Patnaik, S. and Kokkanti, R.R., 2023. Targeting the retinoic acid signaling pathway as a modern precision therapy against cancers. *Frontiers in Cell and Developmental Biology*, *11*, p.1254612.

Lee, J.H., Kang, H.J., Yoo, C.W., Park, W.S., Ryu, J.S., Jung, Y.S., Choi, S.W., Park, J.Y. and Han, N., 2019. PLAG1, SOX10, and Myb expression in benign and malignant salivary gland neoplasms. Journal of Pathology and Translational Medicine, 53(1), pp.23-30.

Lee, S.Y., Han, H.S., Lee, K.Y., Hwang, T.S., Kim, J.H., Sung, I.K., Park, H.S., Jin, C.J. and Choi, K.W., 2007. Sonic hedgehog expression in gastric cancer and gastric adenoma. *Oncology Reports*, *17*(5), pp.1051-1055.

Leegaard, T. and Lindeman, H., 1970. Salivary—gland tumours clinical picture and treatment. Acta Oto-Laryngologica, 69(sup263), pp.155-159.

Leone, P., Malerba, E., Susca, N., Favoino, E., Perosa, F., Brunori, G., Prete, M. and Racanelli, V., 2024. Endothelial cells in tumor microenvironment: insights and perspectives. *Frontiers in Immunology*, *15*, p.1367875.

Lescroart F, Zaffran S. Hox and Tale transcription factors in heart development and disease. Int J Dev Biol. 2018;62(11–12):837–46.

Lewis, J.E., Olsen, K.D. and Sebo, T.J., 2001. Carcinoma ex pleomorphic adenoma: pathologic analysis of 73 cases. Human pathology, 32(6), pp.596-604.

Li, J., Fan, W., Zhao, Y., Xue, M., Chen, T., Xie, W. and Huang, M., 2021. A Pan-cancer Analysis of CREBBP as a potential predictor for immune checkpoint therapy. Cancer Research, *81*(13_Supplement), pp.2236-2236

Li, Y., Jin, K., van Pelt, G.W., van Dam, H., Yu, X., Mesker, W.E., Ten Dijke, P., Zhou, F. and Zhang, L., 2016. c-Myb enhances breast cancer invasion and metastasis through the Wnt/β-catenin/Axin2 pathway. Cancer research, 76(11), pp.3364-3375.

Lin, L.C., Elkashty, O., Ramamoorthi, M., Trinh, N., Liu, Y., Sunavala-Dossabhoy, G., Pranzatelli, T., Michael, D.G., Chivasso, C., Perret, J. and Chiorini, J.A., 2018. Cross-contamination of the human salivary gland HSG cell line with HeLa cells: A STR analysis study. *Oral diseases*, *24*(8), pp.1477-1483.

Lin, Q.Q., Sun, J.L., Wang, F., Zhang, H.Z., Zhou, G. and Xi, Q., 2023. Current understanding of adenoid cystic carcinoma in the gene expression and targeted therapy. Holistic integrative Oncology, 2(1), p.7.

Liu, J., Cao, L., Chen, J., Song, S., Lee, I.H., Quijano, C., Liu, H., Keyvanfar, K., Chen, H., Cao, L.Y. and Ahn, B.H., 2009. Bmi1 regulates mitochondrial function and the DNA damage response pathway. Nature, 459(7245), pp.387-392.

Liu, X., Zhang, X., Wang, Y., Li, Y., & Liu, Y., 2018. MicroRNA expression profiles in pleomorphic adenoma and carcinoma ex pleomorphic adenoma of the salivary glands. Oncology Letters, 15(4), 4251-4258.

Liu, Y., Wang, J., Luo, S., Zhan, Y. and Lu, Q., 2020. The roles of PPARγ and its agonists in autoimmune diseases: A comprehensive review. *Journal of Autoimmunity*, *113*, p.102510.

Livak, K.J. and Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2− ΔΔCT method. *methods*, *25*(4), pp.402-408.

Mahmud, Z., Gomes, A.R., Lee, H.J., Aimjongjun, S., Jiramongkol, Y., Yao, S., Zona, S., Alasiri, G., Gong, G., Yagüe, E. and Lam, E.W.F., 2019. EP300 and SIRT1/6 co-regulate lapatinib sensitivity via modulating FOXO3-acetylation and activity in breast cancer. Cancers, 11(8), p.1067.

Manickavinayaham, S., Vélez-Cruz, R., Biswas, A.K., Bedford, E., Klein, B.J., Kutateladze, T.G., Liu, B., Bedford, M.T. and Johnson, D.G., 2019. E2F1 acetylation directs p300/CBP-mediated histone acetylation at DNA double-strand breaks to facilitate repair. Nature communications, 10(1), p.4951.

Mannarini, L., Kratochvil, V., Calabrese, L., Silva, L.G., Morbini, P., Betka, J. and Benazzo, M., 2009. Human Papilloma Virus (HPV) in head and neck region: review of literature. *Acta Otorhinolaryngologica Italica*, *29*(3), p.119.

Manucha, V. and Ioffe, O.B., 2008. Metastasizing pleomorphic adenoma of the salivary gland. *Archives of pathology & laboratory medicine*, *132*(9), pp.1445-1447.

Marini, K.D., Payne, B.J., Watkins, D.N. and Martelotto, L.G., 2011. Mechanisms of Hedgehog signalling in cancer. *Growth factors*, *29*(6), pp.221-234.

Martin, D., Abba, M.C., Molinolo, A.A., Vitale-Cross, L., Wang, Z., Zaida, M., Delic, N.C., Samuels, Y., Lyons, J.G. and Gutkind, J.S., 2014. The head and neck cancer cell oncogenome: a platform for the development of precision molecular therapies. *Oncotarget*, *5*(19), p.8906.

Martin, D., Abba, M.C., Molinolo, A.A., Vitale-Cross, L., Wang, Z., Zaida, M., Delic, N.C., Samuels, Y., Lyons, J.G. and Gutkind, J.S., 2014. The head and neck cancer cell oncogenome: a platform for the development of precision molecular therapies. Oncotarget, 5(19), p.8906.

Martins, C., Fonseca, I., Roque, L., Pereira, T., Ribeiro, C., Bullerdiek, J. and Soares, J., 2005. PLAG1 gene alterations in salivary gland pleomorphic adenoma and carcinoma ex-pleomorphic adenoma: a combined study using chromosome banding, in situ hybridization and immunocytochemistry. Modern Pathology, 18(8), pp.1048-1055.

Martins, C., Fonseca, I., Roque, L., Pereira, T., Ribeiro, C., Bullerdiek, J. and Soares, J., 2005. PLAG1 gene alterations in salivary gland pleomorphic adenoma and carcinoma ex pleomorphic adenoma: a combined study using chromosome banding, in situ hybridization and immunocytochemistry. Modern pathology, 18(8), pp.1048-1055.

Massagué, J., 2008. TGFβ in cancer. *Cell*, *134*(2), pp.215-230.

Mat Lazim, N., Yousaf, A., Abusalah, M.A.H., Sulong, S., Mohd Ismail, Z.I., Mohamud, R., Abu-Harirah, H.A., AlRamadneh, T.N., Hassan, R. and Abdullah, B., 2023. The Epigenesis of Salivary Glands Carcinoma: From Field Cancerization to Carcinogenesis. *Cancers*, *15*(7), p.2111.

Matsumiya-Matsumoto, Y., Morita, Y. and Uzawa, N., 2022. Pleomorphic adenoma of the salivary glands and epithelial-mesenchymal transition. Journal of Clinical Medicine, 11(14), p.4210.

Matsumiya-Matsumoto, Y., Morita, Y. and Uzawa, N., 2022. Pleomorphic adenoma of the salivary glands and epithelial–mesenchymal transition. Journal of clinical medicine, 11(14), p.4210.

Matsuyama, A., Hisaoka, M., Nagao, Y. and Hashimoto, H., 2011. Aberrant PLAG1 expression in pleomorphic adenomas of the salivary gland: a molecular genetic and immunohistochemical study. *Virchows Archiv*, *458*, pp.583-592.

Matta, A. and Ralhan, R., 2009. Overview of current and future biologically based targeted therapies in head and neck squamous cell carcinoma. *Head & neck oncology*, *1*, pp.1-8.

Medici, D.D., Hay, E.D., and Olsen, B.R., 2019. LEF1_UP.V1_UP gene set. GSEA-MSigDB, [online] Available at: https://www.gsea-msigdb.org/gsea/msigdb/cards/LEF1_UP.V1_UP [Accessed 13 August 2024].

Min, T.H., Kriebel, M., Hou, S. and Pera, E.M., 2011. The dual regulator Sufu integrates Hedgehog and Wnt signals in the early Xenopus embryo. *Developmental biology*, *358*(1), pp.262-276. Shi, I., Sadraei, N.H., Duan, Z.H. and Shi, T., 2011. Aberrant signaling pathways in squamous cell lung carcinoma. *Cancer informatics*, *10*, pp.CIN-S8283.

Mirza, A.Z., Althagafi, I.I. and Shamshad, H., 2019. Role of PPAR receptor in different diseases and their ligands: Physiological importance and clinical implications. *European journal of medicinal chemistry*, *166*, pp.502-513.

Mishra, S., Kelly, K.K., Rumian, N.L. and Siegenthaler, J.A., 2018. Retinoic acid is required for neural stem and progenitor cell proliferation in the adult hippocampus. Stem Cell Reports, 10(6), pp.1705-1720.

Mitani, Y., Li, J., Rao, P.H., Zhao, Y.J., Bell, D., Lippman, S.M., Weber, R.S., Caulin, C. and El-Naggar, A.K., 2010. Comprehensive analysis of the MYB-NFIB gene fusion in salivary adenoid cystic carcinoma: Incidence, variability, and clinicopathologic significance. Clinical cancer research, 16(19), pp.4722-4731.

Mitani, Y., Rao, P.H., Futreal, P.A., Roberts, D.B., Stephens, P.J., Zhao, Y.J., Zhang, L., Mitani, M., Weber, R.S., Lippman, S.M. and Caulin, C., 2011. Novel chromosomal rearrangements and break points at the t (6; 9) in salivary adenoid cystic carcinoma: association with MYB–NFIB chimeric fusion, MYB expression, and clinical outcome. *Clinical cancer research*, *17*(22), pp.7003-7014.

Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation. Nat Rev Cancer 2007;7:233– 45.

Miyazono, K., Ehata, S. and Koinuma, D., 2012. Tumor-promoting functions of transforming growth factor-β in progression of cancer. Upsala journal of medical sciences, 117(2), pp.143-152.

Moens, C.B. and Selleri, L., 2006. Hox cofactors in vertebrate development. *Developmental biology*, *291*(2), pp.193-206.

Molecular Signatures Database (MSigDB) (2024) 'KRAS.600_UP.V1_DN gene set'. Available at: https://www.gsea-msigdb.org/gsea/msigdb/human/geneset/KRAS.600_UP.V1_DN.html (Accessed: 17 February 2025).

Molofsky, A.V., Pardal, R., Iwashita, T., Park, I.K., Clarke, M.F. and Morrison, S.J., 2003. Bmi-1 dependence distinguishes neural stem cell self-renewal from progenitor proliferation. Nature, 425(6961), pp.962-967.

Muzaffar, J., Bari, S., Kirtane, K. and Chung, C.H., 2021. Recent advances and future directions in clinical management of head and neck squamous cell carcinoma. *Cancers*, *13*(2), p.338.

Nanci, A., 2013. Salivary glands. Ten Cate's oral histology. 8th ed. Nanci A, Development, structure, and function. St. Louis: Elsevier, pp.253-77.

Noguchi, S., Aihara, T., Yoshino, K., Motomura, K., Inaji, H., Imaoka, S. and Koyama, H., 1996. Demonstration of monoclonal origin of human parotid gland pleomorphic adenoma. *Cancer: Interdisciplinary International Journal of the American Cancer Society*, *77*(3), pp.431-435.

Nouraei SA, Hope KL, Kelly CG, et al. Carcinoma ex benign pleomorphic adenoma of the parotid gland. Plast Reconstr Surg. 2005;116:1206–1213.

Nowell, P.C. and Hungerford, D.A., 1960. Chromosome studies on normal and leukemic human leukocytes. Journal of the National Cancer Institute, 25(1), pp.85-109.

Nusse, R. and Clevers, H., 2017. Wnt/β-catenin signaling, disease, and emerging therapeutic modalities. Cell, 169(6), pp.985-999.

Ogawa Y. 2003. Immunocytochemistry of myoepithelial cells in the salivary glands. Prog Histochem Cytochem 38:343–426.

Olsen, K.D. and Lewis, J.E., 2001. Carcinoma ex pleomorphic adenoma: a clinicopathologic review. Head & Neck: Journal for the Sciences and Specialties of the Head and Neck, 23(9), pp.705-712.

Oniscu, A., James, R.M., Morris, R.G., Bader, S., Malcomson, R.D. and Harrison, D.J., 2004. Expression of Sonic hedgehog pathway genes is altered in colonic neoplasia. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland*, *203*(4), pp.909-917.

Ono Minagi, H., Sarper, S. E., Kurosaka, H., Kuremoto, K. I., Taniuchi, I., Sakai, T., & Yamashiro, T. (2017). Runx1 mediates the develop ment of the granular convoluted tubules in the submandibular glands. PLoS One, 12(9), e0184395. https://doi.org/10.1371/journ al.pone.0184395

Osanai, M., Takasawa, A., Takasawa, K., Kyuno, D., Ono, Y. and Magara, K., 2023. Retinoic acid metabolism in cancer: potential feasibility of retinoic acid metabolism blocking therapy. *Medical Molecular Morphology*, *56*(1), pp.1-10.

Padam, K.S.R., Chakrabarty, S., Hunter, K.D. and Radhakrishnan, R., 2024. Exploring the regulatory interactions between mutated genes and homeobox genes in the head and neck cancer progression. *Archives of oral biology*, *159*, p.105872.

Paiva, K.B.S., das Graças Silva-Valenzuela, M., Massironi, S.M.G., Ko, G.M., Siqueira, F.M. and Nunes, F.D., 2010. Differential Shh, Bmp and Wnt gene expressions during craniofacial development in mice. *Acta histochemica*, *112*(5), pp.508-517.

Panigrahy, D., Kaipainen, A., Huang, S., Butterfield, C.E., Barnés, C.M., Fannon, M., Laforme, A.M., Chaponis, D.M., Folkman, J. and Kieran, M.W., 2008. PPARα agonist fenofibrate suppresses tumor growth through direct and indirect angiogenesis inhibition. *Proceedings of the National Academy of Sciences*, *105*(3), pp.985-990.

Park, I.K., Morrison, S.J. and Clarke, M.F., 2004. Bmi1, stem cells, and senescence regulation. The Journal of clinical investigation, 113(2), pp.175-179.

Pascovici, D., Handler, D.C., Wu, J.X. and Haynes, P.A., 2016. Multiple testing corrections in quantitative proteomics: A useful but blunt tool. *Proteomics*, *16*(18), pp.2448-2453.

Peinado, H., Quintanilla, M. and Cano, A., 2003. Transforming growth factor β-1 induces snail transcription factor in epithelial cell lines: mechanisms for epithelial mesenchymal transitions. Journal of Biological Chemistry, 278(23), pp.21113-21123.

Persson, M., Andrén, Y., Mark, J., Horlings, H.M., Persson, F. and Stenman, G., 2009. Recurrent fusion of MYB and NFIB transcription factor genes in carcinomas of the breast and head and neck. Proceedings of the National Academy of Sciences, 106(44), pp.18740-18744.

Petrova, R. and Joyner, A.L., 2014. Roles for Hedgehog signaling in adult organ homeostasis and repair. *Development*, *141*(18), pp.3445-3457.

Pickup, M.W., Mouw, J.K. and Weaver, V.M., 2014. The extracellular matrix modulates the hallmarks of cancer. *EMBO reports*, *15*(12), pp.1243-1253.

Polakis, P., 2012. Wnt signaling in cancer. Cold Spring Harbor perspectives in biology, 4(5), p.a008052.

Pozzi, A., Ibanez, M.R., Gatica, A.E., Yang, S., Wei, S., Mei, S., Falck, J.R. and Capdevila, J.H., 2007. Peroxisomal proliferator-activated receptor-α-dependent inhibition of endothelial cell proliferation and tumorigenesis. *Journal of Biological Chemistry*, *282*(24), pp.17685-17695.

Pylayeva-Gupta, Y., Grabocka, E. and Bar-Sagi, D., 2011. RAS oncogenes: weaving a tumorigenic web. *Nature Reviews Cancer*, *11*(11), pp.761-774.

Quinonez, S.C. and Innis, J.W., 2014. Human HOX gene disorders. *Molecular genetics and metabolism*, *111*(1), pp.4-15.

Quintana, A.M., Liu, F., O'Rourke, J.P. and Ness, S.A., 2011. Identification and regulation of c-Myb target genes in MCF-7 cells. BMC cancer, 11, pp.1-12.

Rak, J., Mitsuhashi, Y., Sheehan, C., Tamir, A., Viloria-Petit, A., Filmus, J., Mansour, S.J., Ahn, N.G. and Kerbel, R.S., 2000. Oncogenes and tumor angiogenesis: differential modes of vascular endothelial growth factor up-regulation in ras-transformed epithelial cells and fibroblasts. Cancer research, 60(2), pp.490-498.

Ramsay, R.G. and Gonda, T.J., 2008. MYB function in normal and cancer cells. *Nature Reviews Cancer*, *8*(7), pp.523-534.

Reya, T. and Clevers, H., 2005. Wnt signalling in stem cells and cancer. Nature, 434(7035), pp.843-850.

Riad, M.A., Abdel-Rahman, H., Ezzat, W.F., Adly, A., Dessouky, O. and Shehata, M., 2011. Variables related to recurrence of pleomorphic adenomas: outcome of parotid surgery in 182 cases. The Laryngoscope, 121(7), pp.1467-1472.

Ring, A., Kaur, P. and Lang, J.E., 2020. EP300 knockdown reduces cancer stem cell phenotype, tumor growth and metastasis in triple negative breast cancer. BMC cancer, 20, pp.1-14.

Roden, A.C., Greipp, P.T., Knutson, D.L., Kloft-Nelson, S.M., Jenkins, S.M., Marks, R.S., Aubry, M.C. and García, J.J., 2015. Histopathologic and cytogenetic features of pulmonary adenoid cystic carcinoma. *Journal of Thoracic Oncology*, *10*(11), pp.1570-1575.

Röijer, E., Nordkvist, A., Ström, A.K., Ryd, W., Behrendt, M., Bullerdiek, J., Mark, J. and Stenman, G., 2002. Translocation, deletion/amplification, and expression of HMGIC and MDM2 in a carcinoma ex pleomorphic adenoma. The American journal of pathology, 160(2), pp.433-440.

Romer, A.M.A., Thorseth, M.L. and Madsen, D.H., 2021. Immune modulatory properties of collagen in cancer. *Frontiers in Immunology*, *12*.

Rowley, J.D., 2001. Chromosome translocations: dangerous liaisons revisited. Nature Reviews Cancer, 1(3), pp.245-250.

Sangeetha Priya, P., Anitha, N., Rajesh, E. and Masthan, K.M.K., 2020. Embryology and development of salivary gland. European Journal of Molecular & Clinical Medicine, 7(10), pp.764-70.

Scarini, J.F., de Lima-Souza, R.A., Lavareze, L., Ribeiro de Assis, M.C.F., Damas, I.I., Altemani, A., Egal, E.S.A., Dos Santos, J.N., Bello, I.O. and Mariano, F.V., 2023. Heterogeneity and versatility of

the extracellular matrix during the transition from pleomorphic adenoma to carcinoma ex pleomorphic adenoma: cumulative findings from basic research and new insights. *Frontiers in Oral Health*, *4*, p.942604.

Schwabe, R.F., Bradham, C.A., Uehara, T., Hatano, E., Bennett, B.L., Schoonhoven, R. and Brenner, D.A., 2003. c-Jun-N-terminal kinase drives cyclin D1 expression and proliferation during liver regeneration. *Hepatology*, *37*(4), pp.824-832.

Seeler, J.S. and Dejean, A., 2017. SUMO and the robustness of cancer. *Nature Reviews Cancer*, *17*(3), pp.184-197.

Sene, K.H., Porter, C.J., Palidwor, G., Perez-Iratxeta, C., Muro, E.M., Campbell, P.A., Rudnicki, M.A. and Andrade-Navarro, M.A., 2007. Gene function in early mouse embryonic stem cell differentiation. *BMC genomics*, *8*, pp.1-21.

Shah, A.A.K., Mulla, A.F. and Mayank, M., 2016. Pathophysiology of myoepithelial cells in salivary glands. *Journal of Oral and Maxillofacial Pathology*, *20*(3), pp.480-490.

Shah, N. and Sukumar, S., 2010. The Hox genes and their roles in oncogenesis. *Nature Reviews Cancer*, *10*(5), pp.361-371.

Shaulian, E. and Karin, M., 2002. AP-1 as a regulator of cell life and death. *Nature cell biology*, *4*(5), pp.E131-E136.

Shenoy, U.S., Adiga, D., Kabekkodu, S.P., Hunter, K.D. and Radhakrishnan, R., 2022. Molecular implications of HOX genes targeting multiple signaling pathways in cancer. *Cell biology and toxicology*, pp.1-30.

Shiama, N., 1997. The p300/CBP family: integrating signals with transcription factors and chromatin. *Trends in cell biology*, *7*(6), pp.230-236.

Sjöström, J. and Bergh, J., 2001. How apoptosis is regulated, and what goes wrong in cancer. *Bmj*, *322*(7301), pp.1538-1539.

Skálová, A., Bradová, M., Michal Jr, M., Mosaieby, E., Klubíčková, N., Vaněček, T. and Leivo, I., 2024. Molecular pathology in diagnosis and prognostication of head and neck tumors. *Virchows Archiv*, *484*(2), pp.215-231.

Smith, J., Zyoud, A. and Allegrucci, C., 2019. A case of identity: HOX genes in normal and cancer stem cells. *Cancers*, *11*(4), p.512.

Soares, A.B., Demasi, A.P., Tincani, A.J., Martins, A.S., Altemani, A. and de Araújo, V.C., 2012. The increased PDGF-A, PDGF-B and FGF-2 expression in recurrence of salivary gland pleomorphic adenoma. Journal of clinical pathology, 65(3), pp.272-277.

Soares, C.D., de Cáceres, C.V.B.L., Rodrigues-Fernandes, C.I., de Lima Morais, T.M., de Almeida, O.P., de Carvalho, M.G.F. and Fonseca, F.P., 2021. Prognostic importance of RUNX1 expression for head and neck adenoid cystic carcinoma. *Oral Diseases*, *27*(2), pp.266-276.

Som, P.M. and Miletich, I., 2015. The embryology of the salivary glands: an update. Neurographics, 5(4), pp.167-177.

Song, M., Xiao, C., Wang, T., Pei, Q., Wang, S., Xu, L. and Chen, W., 2011. Study of the differentially expressed genes in pleomorphic adenoma using cDNA microarrays. *Pathology & Oncology Research*, *17*, pp.765-769.

Song, P., Gao, Z., Bao, Y., Chen, L., Huang, Y., Liu, Y., Dong, Q. and Wei, X., 2024. Wnt/β-catenin signaling pathway in carcinogenesis and cancer therapy. Journal of Hematology & Oncology, 17(1), p.46.

Speight, P.M. and Barrett, A.W., 2020. Salivary gland tumours: diagnostic challenges and an update on the latest WHO classification. *Diagnostic histopathology*, *26*(4), pp.147-158.

Srivastava, S.K., Bhardwaj, A., Arora, S., Singh, S., Azim, S., Tyagi, N., Carter, J.E., Wang, B. and Singh, A.P., 2015. MYB is a novel regulator of pancreatic tumour growth and metastasis. *British* journal of cancer, 113(12), pp.1694-1703.

Sroczynska, P., Lancrin, C., Kouskoff, V., & Lacaud, G. (2009). The differential activities of Runx1 promoters define milestones during embryonic hematopoiesis. Blood, 114(26), 5279–5289. https://doi. org/10.1182/blood-2009-05-222307

Stenman, G., 2013. Fusion oncogenes in salivary gland tumours: molecular and clinical consequences. Head and neck pathology, 7(1), pp.12-19.

Stenman, G., Fehr, A., Skálová, A., Vander Poorten, V., Hellquist, H., Mikkelsen, L.H., Saba, N.F., Guntinas-Lichius, O., Chiesa-Estomba, C.M., Andersson, M.K. and Ferlito, A., 2022. Chromosome translocations, gene fusions, and their molecular consequences in pleomorphic salivary gland adenomas. *Biomedicines*, *10*(8), p.1970.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. and Mesirov, J.P., 2005. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences*, *102*(43), pp.15545-15550.

Suzuki, T., Kano, S., Suzuki, M., Yasukawa, S., Mizumachi, T., Tsushima, N., Hatanaka, K.C., Hatanaka, Y., Matsuno, Y. and Homma, A., 2021. Enhanced angiogenesis in salivary duct carcinoma ex-pleomorphic adenoma. Frontiers in Oncology, 10, p.603717.

Suzuki, T., Kano, S., Suzuki, M., Yasukawa, S., Mizumachi, T., Tsushima, N., Hatanaka, K.C., Hatanaka, Y., Matsuno, Y. and Homma, A., 2021. Enhanced angiogenesis in salivary duct carcinoma ex-pleomorphic adenoma. *Frontiers in Oncology*, *10*, p.603717.

Swid, M.A., Li, L., Drahnak, E.M., Idom, H. and Quinones, W., 2023. Updated Salivary Gland Immunohistochemistry A Review. Archives of Pathology & Laboratory Medicine, 147(12), pp.1383-1389.

Szabo, E., Riffe, M.E., Steinberg, S.M., Birrer, M.J. and Linnoila, R.I., 1996. Altered cJUN expression: an early event in human lung carcinogenesis. Cancer Research, 56(2), pp.305-315.

Tamayo, P., 2024. ESC_V6.5_UP_EARLY.V1_DN. Available at: https://www.gsea-msigdb.org/gsea/msigdb/cards/ESC_V6.5_UP_EARLY.V1_DN [Accessed 13 August 2024].

Tan, Y., Wang, Z., Xu, M., Li, B., Huang, Z., Qin, S., Nice, E.C., Tang, J. and Huang, C., 2023. Oral squamous cell carcinomas: state of the field and emerging directions. *International journal of oral science*, *15*(1), p.44.

Thompson, C.B., Challoner, P.B., Neiman, P.E. and Groudine, M., 1986. Expression of the c-myb proto-oncogene during cellular proliferation. Nature, 319(6052), pp.374-380.

Tian, J., Hu, L., Li, X., Geng, J., Dai, M. and Bai, X., 2016. MicroRNA-130b promotes lung cancer progression via PPARγ/VEGF-A/BCL-2-mediated suppression of apoptosis. *Journal of Experimental & Clinical Cancer Research*, *35*, pp.1-15.

Tondi-Resta, I., Hobday, S.B., Gubbiotti, M.A., Jalaly, J.B., Rassekh, C.H., Montone, K.T. and Baloch, Z.W., 2023. Carcinoma ex pleomorphic adenomas: an institutional experience and literature review. *American journal of clinical pathology*, *159*(5), pp.502-515.

Toretsky, J.A. and Helman, L.J., 1996. Involvement of IGF-II in human cancer. *Journal of endocrinology*, *149*(3), pp.367-372.

Toth, C., Funke, S., Nitsche, V., Liverts, A., Zlachevska, V., Gasis, M., Wiek, C., Hanenberg, H., Mahotka, C., Schirmacher, P. and Heikaus, S., 2017. The role of apoptosis repressor with a CARD domain (ARC) in the therapeutic resistance of renal cell carcinoma (RCC): the crucial role of ARC in the inhibition of extrinsic and intrinsic apoptotic signalling. *Cell Communication and Signaling*, *15*, pp.1-14.

Tuo, Z., Zhang, Y., Wang, X., Dai, S., Liu, K., Xia, D., Wang, J. and Bi, L., 2022. RUNX1 is a promising prognostic biomarker and related to immune infiltrates of cancer-associated fibroblasts in human cancers. *BMC cancer*, *22*(1), p.523.

Vidal, M.T.A., Lourenço, S.V., Soares, F.A., Gurgel, C.A., Studart, E.J., Valverde, L.D.F., Araújo, I.B.D.O., Ramos, E.A.G., Xavier, F.C.D.A. and Dos Santos, J.N., 2016. The sonic hedgehog signaling

pathway contributes to the development of salivary gland neoplasms regardless of perineural infiltration. *Tumor Biology*, *37*, pp.9587-9601.

Vleugel, M.M., Greijer, A.E., Bos, R., van der Wall, E. and van Diest, P.J., 2006. c-Jun activation is associated with proliferation and angiogenesis in invasive breast cancer. Human pathology, 37(6), pp.668-674.

Voronov, D., Gromova, A., Liu, D., Zoukhri, D., Medvinsky, A., Meech, R. and Makarenkova, H.P., 2013. Transcription factors Runx1 to 3 are expressed in the lacrimal gland epithelium and are involved in regulation of gland morphogenesis and regeneration. *Investigative ophthalmology & visual science*, *54*(5), pp.3115-3125.

Voz ML, Agten NS, Van de Ven WJM & Kas K 2000 PLAG1, the main translocation target in pleomorphic adenoma of the salivary glands, is a positive regulator of IGF-II. Cancer Research 60 106–113.

Voz ML, Åström A-K, Kas K, Mark J, Stenman G & Van de Ven WJM 1998 The recurrent translocation t(5;8)(p13;q12) in pleomorphic adenomas results in upregulation of PLAG1 gene expression under control of the LIFR promoter. Oncogene 16 1409–1416.

Voz, M.L., Agten, N.S., Van de Ven, W.J. and Kas, K., 2000. PLAG1, the main translocation target in pleomorphic adenoma of the salivary glands, is a positive regulator of IGF-II. *Cancer research*, *60*(1), pp.106-113.

Voz, M.L., Mathys, J., Hensen, K., Pendeville, H., Van Valckenborgh, I., Van Huffel, C., Chavez, M., Van Damme, B., De Moor, B., Moreau, Y. and Van de Ven, W.J., 2004. Microarray screening for target genes of the proto-oncogene PLAG1. *Oncogene*, *23*(1), pp.179-191.

Wang, H., Birkenbach, M. and Hart, J., 2000. Expression of Jun family members in human colorectal adenocarcinoma. Carcinogenesis, 21(7), pp.1313-1317.

Wang, J., Huang, Y., Xu, J., Yue, B., Wen, Y., Wang, X., Lei, C. and Chen, H., 2022. Pleomorphic adenoma gene 1 (PLAG1) promotes proliferation and inhibits apoptosis of bovine primary myoblasts through the PI3K-Akt signaling pathway. *Journal of Animal Science*, *100*(4), p.skac098.

Wang, W., Wu, S., Shi, Y., Miao, Y., Luo, X., Ji, M., Yao, K. and He, J., 2015. c-MYB regulates cell growth and DNA damage repair through modulating MiR-143. *FEBS letters*, *589*(5), pp.555-564.

Wang, W.H., Yuan, T., Qian, M.J., Yan, F.J., Yang, L., He, Q.J., Yang, B., Lu, J.J. and Zhu, H., 2021. Post-translational modification of KRAS: potential targets for cancer therapy. *Acta Pharmacologica Sinica*, *42*(8), pp.1201-1211.

Wang, Y., Kaiser, C.E., Frett, B. and Li, H.Y., 2013. Targeting mutant KRAS for anticancer therapeutics: a review of novel small molecule modulators. Journal of medicinal chemistry, 56(13), pp.5219-5230.

Wenig, B.M., Jo, V.Y., Nakaguro, M. and Skálová, A., 2023. Histologic Considerations and Salivary Gland Tumor Classification in Surgical Pathology. In *The Milan System for Reporting Salivary Gland Cytopathology* (pp. 229-235). Cham: Springer International Publishing.

Witt, R.L., 2002. The significance of the margin in parotid surgery for pleomorphic adenoma. The Laryngoscope, 112(12), pp.2141-2154.

Xu, X., Ding, H., Rao, G., Arora, S., Saclarides, C.P., Esparaz, J., Gattuso, P., Solorzano, C.C. and Prinz, R.A., 2012. Activation of the Sonic Hedgehog pathway in thyroid neoplasms and its potential role in tumor cell proliferation. *Endocrine-Related Cancer*, *19*(2), pp.167-179.

Xu, Y., Li, H., Yang, F., Yang, D. and Zhou, B.B.S., 2020. Cell plasticity and genomic instability in cancer evolution. Genome Instability & Disease, 1, pp.301-309.

Yang, B., Liu, Y., Zhao, J., Hei, K., Zhuang, H., Li, Q., Wei, W., Chen, R., Zhang, N. and Li, Y., 2017. Ectopic overexpression of filamin C scaffolds MEK1/2 and ERK1/2 to promote the progression of human hepatocellular carcinoma. *Cancer Letters*, *388*, pp.167-176.

Yang, R. M., Nanayakkara, D., Kalimutho, M., Mitra, P., Khanna, K. K., Dray, E., et al. (2019). MYB regulates the DNA damage response and components of the homology-directed repair pathway in human estrogen receptor-positive breast cancer cells. Oncogene

Ye, L., Bokobza, S.M. and Jiang, W.G., 2009. Bone morphogenetic proteins in development and progression of breast cancer and therapeutic potential. *International journal of molecular medicine*, *24*(5), pp.591-597.

Young, A. and Okuyemi, O.T., 2023. Malignant salivary gland tumors. In StatPearls [Internet]. StatPearls Publishing.

Yurdakok-Dikmen, B., Filazi, A. and Ince, S., 2017. Retinoids. In *Reproductive and Developmental Toxicology* (pp. 481-492). Academic Press.

Zagozewski, J.L., Zhang, Q., Pinto, V.I., Wigle, J.T. and Eisenstat, D.D., 2014. The role of homeobox genes in retinal development and disease. *Developmental biology*, *393*(2), pp.195-208.

Zatkova A, Rouillard J-M, Hartmann W, Lamb BJ, Kuick R, Eckart M, von Schweinitz D, Koch A, Fonatsch C & Pietsch T et al. 2004 Amplification and overexpression of the IGF2 regulator PLAG1 in hepatoblastoma. Genes Chromosomes and Cancer 39 126–137.

Zbären, P., Zbären, S., Caversaccio, M.D. and Stauffer, E., 2008. Carcinoma ex pleomorphic adenoma: diagnostic difficulty and outcome. Otolaryngology—Head and Neck Surgery, 138(5), pp.601-605.

Zhan, K.Y., Khaja, S.F., Flack, A.B. and Day, T.A., 2016. Benign parotid tumors. Otolaryngol Clin North Am, 49(2), pp.327-342.

Zhan, T., Rindtorff, N. and Boutros, M., 2017. Wnt signaling in cancer. Oncogene, 36(11), pp.1461-1473.

Zhang, J., Vlasevska, S., Wells, V.A., Nataraj, S., Holmes, A.B., Duval, R., Meyer, S.N., Mo, T., Basso, K., Brindle, P.K. and Hussein, S., 2017. The CREBBP acetyltransferase is a haploinsufficient tumor suppressor in B-cell lymphoma. *Cancer discovery*, *7*(3), pp.322-337.

Zhang, J., Zheng, X., Wang, P., Wang, J. and Ding, W., 2021. Role of apoptosis repressor with caspase recruitment domain (ARC) in cell death and cardiovascular disease. *Apoptosis*, *26*, pp.24-37.

Zhang, Y., Pu, X., Shi, M., Chen, L., Qian, L., Song, Y., Yuan, G., Zhang, H., Yu, M., Hu, M. and Shen, B., 2007. c-Jun, a crucial molecule in metastasis of breast cancer and potential target for biotherapy. *Oncology reports*, *18*(5), pp.1207-1212.

Zhang, Y., Yang, Y., Zhang, J., Chen, H., Xu, Z., Zhang, Y., … & Li, J. (2020). Genomic landscape of salivary gland tumors: A comprehensive review. Cancer Letters, 471, 129-141.

Zhang, Y., Yang, Y., Zhang, J., et al. (2015). DNA methylation profiling of pleomorphic adenoma and carcinoma ex pleomorphic adenoma of the salivary glands. PLoS One, 10(1), e0114559.

Zhao, H., Ming, T., Tang, S., Ren, S., Yang, H., Liu, M., Tao, Q. and Xu, H., 2022. Wnt signaling in colorectal cancer: pathogenic role and therapeutic target. *Molecular cancer*, *21*(1), p.144.

Zhao, X., Ren, W., Yang, W., Wang, Y., Kong, H., Wang, L., Yan, L., Xu, G., Fei, J., Fu, J. and Zhang, C., 2006. Wnt pathway is involved in pleomorphic adenomas induced by overexpression of PLAG1 in transgenic mice. *International journal of cancer*, *118*(3), pp.643-648.

Zheng, G. and Yang, Y.C., 2005. Sumoylation and acetylation play opposite roles in the transactivation of PLAG1 and PLAGL2. *Journal of Biological Chemistry*, *280*(49), pp.40773-40781.

Zhong, Z. and Virshup, D.M., 2020. Wnt signaling and drug resistance in cancer. Molecular pharmacology, 97(2), pp.72-89.

Zhu, Y., Wang, Z., Li, Y., Peng, H., Liu, J., Zhang, J. and Xiao, X., 2023. The role of CREBBP/EP300 and its therapeutic implications in hematological malignancies. *Cancers*, *15*(4), p.1219.

8 Appendix

zinc finger protein PLAG1 [Homo sapiens]

Sequence ID: AAC50995.1 **Length:** 500 **Number of Matches:** 1

Range 1: 1 to 386 GenPept Graphics ▼ Next Match ▲ Previous Match

Score	Expect	Method	Identities	Positives	Gaps
805 bits(2079)	0.0	Compositional matrix adjust.	386/386(100%)	386/386(100%)	0/386(0%)

```
Query  1    MATVIPGDLSEVRDTQKVPSGKRKRGETKPRKNFPCQLCDKAFNSVEKLKVHSYSHTGER  60
            MATVIPGDLSEVRDTQKVPSGKRKRGETKPRKNFPCQLCDKAFNSVEKLKVHSYSHTGER
Sbjct  1    MATVIPGDLSEVRDTQKVPSGKRKRGETKPRKNFPCQLCDKAFNSVEKLKVHSYSHTGER  60

Query  61   PYKCIQQDCTKAFVSKYKLQRHMATHSPEKTHKCNYCEKMFHRKDHLKNHLHTHDPNKET  120
            PYKCIQQDCTKAFVSKYKLQRHMATHSPEKTHKCNYCEKMFHRKDHLKNHLHTHDPNKET
Sbjct  61   PYKCIQQDCTKAFVSKYKLQRHMATHSPEKTHKCNYCEKMFHRKDHLKNHLHTHDPNKET  120

Query  121  FKCEECGKNYNTKLGFKRHLALHAATSGDLTCKVCLQTFESTGVLLEHLKSHAGKSSGGV  180
            FKCEECGKNYNTKLGFKRHLALHAATSGDLTCKVCLQTFESTGVLLEHLKSHAGKSSGGV
Sbjct  121  FKCEECGKNYNTKLGFKRHLALHAATSGDLTCKVCLQTFESTGVLLEHLKSHAGKSSGGV  180

Query  181  KEKKHQCEHCDRRFYTRKDVRRHMVVHTGRKDFLCQYCAQRFGRKDHLTRHMKKSHNQEL  240
            KEKKHQCEHCDRRFYTRKDVRRHMVVHTGRKDFLCQYCAQRFGRKDHLTRHMKKSHNQEL
Sbjct  181  KEKKHQCEHCDRRFYTRKDVRRHMVVHTGRKDFLCQYCAQRFGRKDHLTRHMKKSHNQEL  240

Query  241  LKVKTEPVDFLDPFTCNVSVPIKDELLPVMSLPSSELLSKPFTNTLQLNLYNTPFQSMQS  300
            LKVKTEPVDFLDPFTCNVSVPIKDELLPVMSLPSSELLSKPFTNTLQLNLYNTPFQSMQS
Sbjct  241  LKVKTEPVDFLDPFTCNVSVPIKDELLPVMSLPSSELLSKPFTNTLQLNLYNTPFQSMQS  300

Query  301  SGSAHQMITTLPLGMTCPIDMDTVHPSHHLSFKYPFSSTSYAISIPEKEQPLKGEIESYL  360
            SGSAHQMITTLPLGMTCPIDMDTVHPSHHLSFKYPFSSTSYAISIPEKEQPLKGEIESYL
Sbjct  301  SGSAHQMITTLPLGMTCPIDMDTVHPSHHLSFKYPFSSTSYAISIPEKEQPLKGEIESYL  360

Query  361  MELQGGVPSSSQDSQASSSSKLGLDP  386
            MELQGGVPSSSQDSQASSSSKLGLDP
Sbjct  361  MELQGGVPSSSQDSQASSSSKLGLDP  386
```

Biomatik
Tel: (519) 489-7195, (800) 836-8089
Fax: (519) 231-0140, (877) 221-3515
Email: info@biomatik.com
http://www.biomatik.com

Gene Synthesis QA Report

Gene Name	PLAG1Flag	Order ID	SG210269-4
Lot#	VF62034	Cloning Vector	pCDNA3.1(+)
Length (bp)	1564	Cloning Sites	NheI(GCTAGC)-XhoI(CTCGAG)

Quality Control

Test Items	Specifications	Results
Sequencing Alignment	Sequencing data consistent with target	[X] Pass
Vector Sequence	Flanking sequence of cloning sites are correct	[X] Pass
Restriction Digest	Insert size is correct and no contaminated bands	[X] Pass
ORF Across Junction	Correct and consistent with target	N/A
PCR Amplification	Correct and no contaminated bands	[X] Pass
Endotoxin Level	Verified, <0.1 EU/μg (Endo-Free Preps Only)	N/A
Appearance	Clear, no foreign particles	[X] Pass
DNA Purity	Purity (A 260/A280 = 1.8 - 2.0)	[X] Pass
DNA Quantity	Actual yield (by A 260)	5μg/5μg

Comments

Delivery form: lyophilized plasmid (TE lyophilized) containing the gene insert. It is stable at room temperature for extended period of time during shipping. The lyophilized plasmid can be dissolved in sterile TE buffer or nuclease-free water (neutral pH) depending on the established laboratory practice. After reconstitution, store the stock solution at -20ºC or -80ºC for long term storage. The lyophilized plasmid dissolved in TE buffer is stable for at least 6 months at 4°C while the lyophilized DNA dissolved in water is NOT STABLE at 4°C.

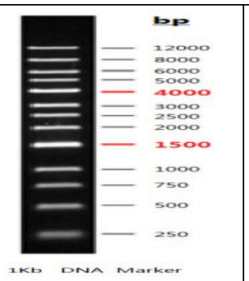
a) Before opening the tube containing the plasmid, please briefly centrifuge the tube. Lyophilized plasmid could attach to the wall of the tube. Opening without centrifugation could cause DNA loss.
b) Stock Solution: Reconstitute lyophilized plasmid (4ug or 10ug) in 40ul or 100ul of TE buffer or nuclease-free water (final concentration - 100ng/ul). To accurately determine the quantity of DNA present, please measure OD value of original stock at OD_{260}nm after reconstitution.
c) Working Solution: make a 1:10 dilution of stock solution using TE buffer or nuclease-free water (final concentration - 10ng/ul)

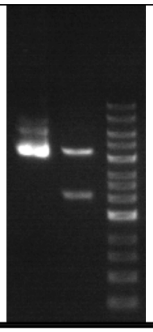
Transformation and Replating:

Transform 2ul of stock solution into appropriate *E. coli* competent cells according to standard laboratory protocol. Plate the mixture on LB agar (with desired antibiotic selection) and incubate at 37°C for overnight. Select a well separated, SINGLE colony and inoculate in LB medium with desired antibiotic selection for overnight culture. Purify Plasmid DNA from overnight cultures, verify sequences and continue with project of interest. *It is important to select only a SINGLE colony for overnight culture.

Restriction Digestion

Biomatik
Tel: (519) 489-7195, (800) 836-8089
Fax: (519) 231-0140, (877) 221-3515
Email: info@biomatik.com
http://www.biomatik.com



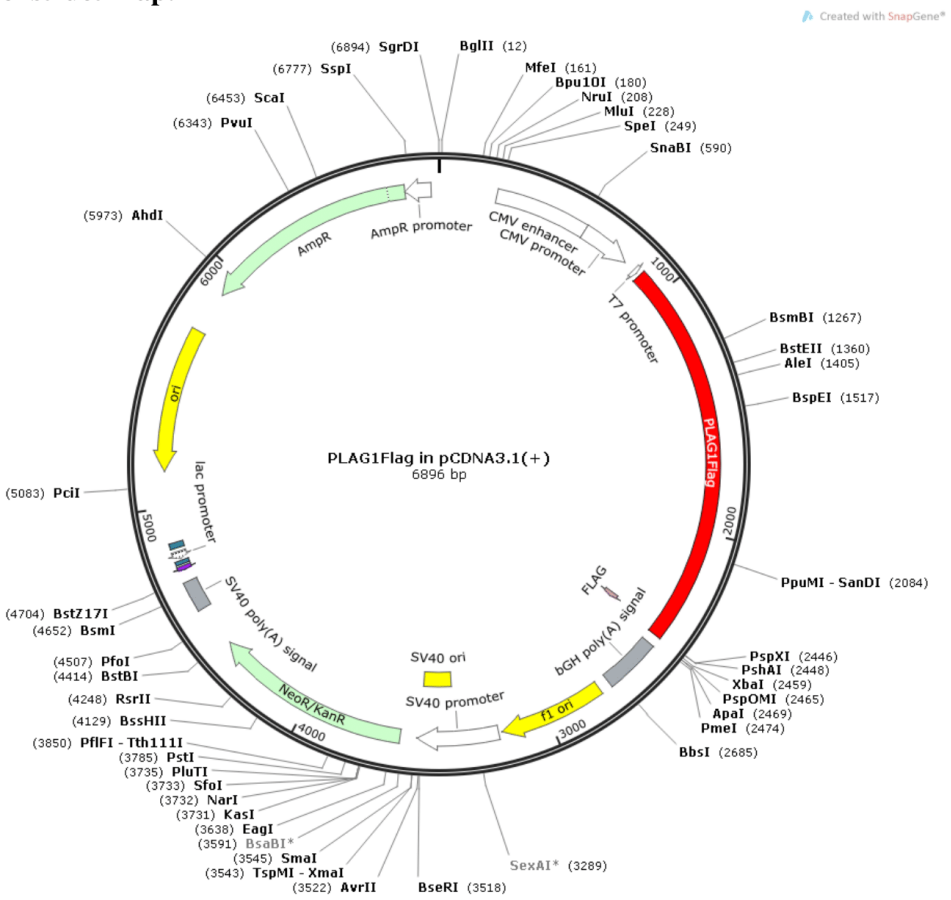


Gene name: PLAG1Flag

Clone ID#: VF62034

RES: MluI-XhoI

Construct Map:



Biomatik
Tel: (519) 489-7195, (800) 836-8089
Fax: (519) 231-0140, (877) 221-3515
Email: info@biomatik.com
http://www.biomatik.com

Detailed Sequence of the Whole Construct:

```
   1 GACGGATCGG GAGATCTCCC GATCCCCTAT GGTGCACTCT CAGTACAATC TGCTCTGATG
  61 CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG
 121 CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC
 181 TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT
 241 GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
 301 TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
 361 CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC
 421 ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT
 481 ATCATATGCC AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT
 541 ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
 601 TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG
 661 ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC
 721 AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG
 781 GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA
 841 CTGCTTACTG GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTGGCTAGC
 901 ACGCTAGCCG CCGCCACCAT GGCCACTGTC ATTCCTGGTG ATTTGTCAGA AGTAAGAGAT
 961 ACCCAGAAAG TCCCTTCAGG GAAACGTAAG CGTGGTGAAA CCAAACCAAG AAAAAACTTT
1021 CCTTGCCAAC TGTGTGACAA GGCCTTTAAC AGTGTTGAGA AATTAAAGGT TCACTCCTAC
1081 TCTCACACAG GAGAGAGGCC CTACAAGTGC ATACAACAAG ACTGCACCAA GGCCTTTGTT
1141 TCTAAGTACA AATTACAAAG GCACATGGCT ACTCATTCTC CTGAGAAAAC CCACAAGTGT
1201 AATTATTGTG AGAAAATGTT TCACCGGAAA GATCATCTGA AGAATCACCT CCATACACAC
1261 GACCCTAACA AAGAGACGTT TAAGTGCGAA GAATGTGGCA AGAACTACAA TACCAAGCTT
1321 GGATTTAAAC GTCACTTGGC CTTGCATGCC GCAACAAGTG GTGACCTCAC CTGTAAGGTA
1381 TGTTTGCAAA CTTTTGAAAG CACGGGAGTG CTTCTGGAGC ACCTTAAATC TCATGCAGGC
1441 AAGTCGTCTG GTGGGGTTAA AGAAAAAAAG CACCAGTGCG AACATTGTGA TCGCCGGTTC
1501 TACACCCGAA AGGATGTCCG GAGACACATG GTGGTGCACA CTGGAAGAAA GGACTTCCTC
1561 TGTCAGTATT GTGCACAGAG ATTTGGGCGA AAGGATCACC TGACTCGACA TATGAAGAAG
1621 AGTCACAATC AAGAGCTTCT GAAGGTCAAA ACAGAACCAG TGGATTTCCT TGACCCATTT
1681 ACCTGCAATG TGTCTGTGCC TATAAAAGAC GAGCTCCTTC CGGTGATGTC CTTACCTTCC
1741 AGTGAACTGT TATCAAAGCC ATTCACAAAC ACTTTGCAGT TAAACCTCTA CAACACTCCA
1801 TTTCAGTCCA TGCAGAGCTC GGGATCTGCC CACCAAATGA TCACAACTTT ACCTTTGGGA
1861 ATGACATGCC CAATAGATAT GGACACTGTT CATCCCTCTC ACCACCTTTC TTTCAAATAT
1921 CCGTTCAGTT CTACCTCATA TGCAATTTCT ATTCCTGAAA AAGAACAGCC ATTAAAGGGG
1981 GAAATTGAGA GTTACCTGAT GGAGTTACAA GGTGGCGTGC CCTCTTCATC CCAAGATTCT
2041 CAAGCATCGT CATCATCTAA GCTAGGGTTG GATCCTCAGA TTGGGTCCCT AGATGATGGT
2101 GCAGGAGACC TCTCCCTATC CAAAAGCTCT ATCTCCATCA GTGACCCCCT AAACACACCA
2161 GCATTGGATT TTTCTCAGTT GTTTAATTTC ATACCTTTAA ATGGTCCTCC CTATAATCCT
2221 CTATCAGTGG GGAGCCTTGG AATGAGCTAT TCCCAGGAAG AAGCACATTC TTCTGTTTCC
2281 CAGCTCCCCC CACAAACACA GGATCTTCAG GATCCTGCAA ACACTATAGG GCTTGGGTCT
2341 CTGCACTCAC TGTCAGCAGC TTTCACCAGC AGTTTAAGCA CAAGTACCAC CCTCCCACGT
2401 TTCCATCAAG CTTTTCAGGA CTACAAGGAC GACGATGACA AGTGACTCGA GTCTCGAGTC
2461 TAGAGGGCCC GTTTAAACCC GCTGATCAGC CTCGACTGTG CCTTCTAGTT GCCAGCCATC
2521 TGTTGTTTGC CCCTCCCCCG TGCCTTCCTT GACCCTGGAA GGTGCCACTC CCACTGTCCT
2581 TTCCTAATAA AATGAGGAAA TTGCATCGCA TTGTCTGAGT AGGTGTCATT CTATTCTGGG
2641 GGGTGGGGTG GGGCAGGACA GCAAGGGGGA GGATTGGGAA GACAATAGCA GGCATGCTGG
2701 GGATGCGGTG GGCTCTATGG CTTCTGAGGC GGAAAGAACC AGCTGGGGCT CTAGGGGGTA
2761 TCCCCACGCG CCCTGTAGCG GCGCATTAAG CGCGGCGGGT GTGGTGGTTA CGCGCAGCGT
2821 GACCGCTACA CTTGCCAGCG CCCTAGCGCC CGCTCCTTTC GCTTTCTTCC CTTCCTTTCT
2881 CGCCACGTTC GCCGGCTTTC CCCGTCAAGC TCTAAATCGG GGGCTCCCTT TAGGGTTCCG
2941 ATTTAGTGCT TTACGGCACC TCGACCCCAA AAAACTTGAT TAGGGTGATG GTTCACGTAG
3001 TGGGCCATCG CCCTGATAGA CGGTTTTTCG CCCTTTGACG TTGGAGTCCA CGTTCTTTAA
3061 TAGTGGACTC TTGTTCCAAA CTGGAACAAC ACTCAACCCT ATCTCGGTCT ATTCTTTTGA
3121 TTTATAAGGG ATTTTGCCGA TTTCGGCCTA TTGGTTAAAA AATGAGCTGA TTTAACAAAA
3181 ATTTAACGCG AATTAATTCT GTGGAATGTG TGTCAGTTAG GGTGTGGAAA GTCCCCAGGC
3241 TCCCCAGCAG GCAGAAGTAT GCAAAGCATG CATCTCAATT AGTCAGCAAC CAGGTGTGGA
```

Biomatik
Tel: (519) 489-7195, (800) 836-8089
Fax: (519) 231-0140, (877) 221-3515
Email: info@biomatik.com
http://www.biomatik.com

```
3301 AAGTCCCCAG GCTCCCCAGC AGGCAGAAGT ATGCAAAGCA TGCATCTCAA TTAGTCAGCA
3361 ACCATAGTCC CGCCCCTAAC TCCGCCCATC CCGCCCCTAA CTCCGCCCAG TTCCGCCCAT
3421 TCTCCGCCCC ATGGCTGACT AATTTTTTTT ATTTATGCAG AGGCCGAGGC CGCCTCTGCC
3481 TCTGAGCTAT TCCAGAAGTA GTGAGGAGGC TTTTTTGGAG GCCTAGGCTT TTGCAAAAAG
3541 CTCCCGGGAG CTTGTATATC CATTTTCGGA TCTGATCAAG AGACAGGATG AGGATCGTTT
3601 CGCATGATTG AACAAGATGG ATTGCACGCA GGTTCTCCGG CCGCTTGGGT GGAGAGGCTA
3661 TTCGGCTATG ACTGGGCACA ACAGACAATC GGCTGCTCTG ATGCCGCCGT GTTCCGGCTG
3721 TCAGCGCAGG GGCGCCCGGT TCTTTTTGTC AAGACCGACC TGTCCGGTGC CCTGAATGAA
3781 CTGCAGGACG AGGCAGCGCG GCTATCGTGG CTGGCCACGA CGGGCGTTCC TTGCGCAGCT
3841 GTGCTCGACG TTGTCACTGA AGCGGGAAGG GACTGGCTGC TATTGGGCGA AGTGCCGGGG
3901 CAGGATCTCC TGTCATCTCA CCTTGCTCCT GCCGAGAAAG TATCCATCAT GGCTGATGCA
3961 ATGCGGCGGC TGCATACGCT TGATCCGGCT ACCTGCCCAT TCGACCACCA AGCGAAACAT
4021 CGCATCGAGC GAGCACGTAC TCGGATGGAA GCCGGTCTTG TCGATCAGGA TGATCTGGAC
4081 GAAGAGCATC AGGGGCTCGC GCCAGCCGAA CTGTTCGCCA GGCTCAAGGC GCGCATGCCC
4141 GACGGCGAGG ATCTCGTCGT GACCCATGGC GATGCCTGCT TGCCGAATAT CATGGTGGAA
4201 AATGGCCGCT TTTCTGGATT CATCGACTGT GGCCGGCTGG GTGTGGCGGA CCGCTATCAG
4261 GACATAGCGT TGGCTACCCG TGATATTGCT GAAGAGCTTG GCGGCGAATG GGCTGACCGC
4321 TTCCTCGTGC TTTACGGTAT CGCCGCTCCC GATTCGCAGC GCATCGCCTT CTATCGCCTT
4381 CTTGACGAGT TCTTCTGAGC GGGACTCTGG GGTTCGAAAT GACCGACCAA GCGACGCCCA
4441 ACCTGCCATC ACGAGATTTC GATTCCACCG CCGCCTTCTA TGAAAGGTTG GGCTTCGGAA
4501 TCGTTTTCCG GGACGCCGGC TGGATGATCC TCCAGCGCGG GGATCTCATG CTGGAGTTCT
4561 TCGCCCACCC CAACTTGTTT ATTGCAGCTT ATAATGGTTA CAAATAAAGC AATAGCATCA
4621 CAAATTTCAC AAATAAAGCA TTTTTTTCAC TGCATTCTAG TTGTGGTTTG TCCAAACTCA
4681 TCAATGTATC TTATCATGTC TGTATACCGT CGACCTCTAG CTAGAGCTTG GCGTAATCAT
4741 GGTCATAGCT GTTTCCTGTG TGAAATTGTT ATCCGCTCAC AATTCCACAC AACATACGAG
4801 CCGGAAGCAT AAAGTGTAAA GCCTGGGGTG CCTAATGAGT GAGCTAACTC ACATTAATTG
4861 CGTTGCGCTC ACTGCCCGCT TTCCAGTCGG GAAACCTGTC GTGCCAGCTG CATTAATGAA
4921 TCGGCCAACG CGCGGGGAGA GGCGGTTTGC GTATTGGGCG CTCTTCCGCT TCCTCGCTCA
4981 CTGACTCGCT GCGCTCGGTC GTTCGGCTGC GGCGAGCGGT ATCAGCTCAC TCAAAGGCGG
5041 TAATACGGTT ATCCACAGAA TCAGGGGATA ACGCAGGAAA GAACATGTGA GCAAAAGGCC
5101 AGCAAAAGGC CAGGAACCGT AAAAAGGCCG CGTTGCTGGC GTTTTTCCAT AGGCTCCGCC
5161 CCCCTGACGA GCATCACAAA AATCGACGCT CAAGTCAGAG GTGGCGAAAC CCGACAGGAC
5221 TATAAAGATA CCAGGCGTTT CCCCCTGGAA GCTCCCTCGT GCGCTCTCCT GTTCCGACCC
5281 TGCCGCTTAC CGGATACCTG TCCGCCTTTC TCCCTTCGGG AAGCGTGGCG CTTTCTCATA
5341 GCTCACGCTG TAGGTATCTC AGTTCGGTGT AGGTCGTTCG CTCCAAGCTG GGCTGTGTGC
5401 ACGAACCCCC CGTTCAGCCC GACCGCTGCG CCTTATCCGG TAACTATCGT CTTGAGTCCA
5461 ACCCGGTAAG ACACGACTTA TCGCCACTGG CAGCAGCCAC TGGTAACAGG ATTAGCAGAG
5521 CGAGGTATGT AGGCGGTGCT ACAGAGTTCT TGAAGTGGTG GCCTAACTAC GGCTACACTA
5581 GAAGAACAGT ATTTGGTATC TGCGCTCTGC TGAAGCCAGT TACCTTCGGA AAAAGAGTTG
5641 GTAGCTCTTG ATCCGGCAAA CAAACCACCG CTGGTAGCGG TTTTTTTGTT TGCAAGCAGC
5701 AGATTACGCG CAGAAAAAAA GGATCTCAAG AAGATCCTTT GATCTTTTCT ACGGGGTCTG
5761 ACGCTCAGTG GAACGAAAAC TCACGTTAAG GGATTTTGGT CATGAGATTA TCAAAAAGGA
5821 TCTTCACCTA GATCCTTTTA AATTAAAAAT GAAGTTTTAA ATCAATCTAA AGTATATATG
5881 AGTAAACTTG GTCTGACAGT TACCAATGCT TAATCAGTGA GGCACCTATC TCAGCGATCT
5941 GTCTATTTCG TTCATCCATA GTTGCCTGAC TCCCCGTCGT GTAGATAACT ACGATACGGG
6001 AGGGCTTACC ATCTGGCCCC AGTGCTGCAA TGATACCGCG AGACCCACGC TCACCGGCTC
6061 CAGATTTATC AGCAATAAAC CAGCCAGCCG GAAGGGCCGA GCGCAGAAGT GGTCCTGCAA
6121 CTTTATCCGC CTCCATCCAG TCTATTAATT GTTGCCGGGA AGCTAGAGTA AGTAGTTCGC
6181 CAGTTAATAG TTTGCGCAAC GTTGTTGCCA TTGCTACAGG CATCGTGGTG TCACGCTCGT
6241 CGTTTGGTAT GGCTTCATTC AGCTCCGGTT CCCAACGATC AAGGCGAGTT ACATGATCCC
6301 CCATGTTGTG CAAAAAAGCG GTTAGCTCCT TCGGTCCTCC GATCGTTGTC AGAAGTAAGT
6361 TGGCCGCAGT GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT ACTGTCATGC
6421 CATCCGTAAG ATGCTTTTCT GTGACTGGTG AGTACTCAAC CAAGTCATTC TGAGAATAGT
6481 GTATGCGGCG ACCGAGTTGC TCTTGCCCGG CGTCAATACG GGATAATACC GCGCCACATA
6541 GCAGAACTTT AAAAGTGCTC ATCATTGGAA AACGTTCTTC GGGGCGAAAA CTCTCAAGGA
6601 TCTTACCGCT GTTGAGATCC AGTTCGATGT AACCCACTCG TGCACCCAAC TGATCTTCAG
6661 CATCTTTTAC TTTCACCAGC GTTTCTGGGT GAGCAAAAAC AGGAAGGCAA AATGCCGCAA
```

Biomatik
Tel: (519) 489-7195, (800) 836-8089
Fax: (519) 231-0140, (877) 221-3515
Email: info@biomatik.com
http://www.biomatik.com

```
6721 AAAAGGGAAT AAGGGCGACA CGGAAATGTT GAATACTCAT ACTCTTCCTT TTTCAATATT
6781 ATTGAAGCAT TTATCAGGGT TATTGTCTCA TGAGCGGATA CATATTTGAA TGTATTTAGA
6841 AAAATAAACA AATAGGGGTT CCGCGCACAT TTCCCCGAAA AGTGCCACCT GACGTC
```

Certified by:

Date:04/02/2021

Biomatik
Tel: (519) 489-7195, (800) 836-8089
Fax: (519) 231-0140, (877) 221-3515
Email: info@biomatik.com
http://www.biomatik.com

Gene Synthesis QA Report

Gene Name	FLagTurboID PLAG1	**Order ID**	SG230310-1
Lot#	TF92273	**Cloning Vector**	pcDNA3.1(+)
Length (bp)	2508	**Cloning Sites**	NheI(GCTAGC)-XhoI(CTCGAG)

Quality Control

Test Items	Specifications	Results
Sequencing Alignment	Sequencing data consistent with target	[X] Pass
Vector Sequence	Flanking sequence of cloning sites are correct	[X] Pass
Restriction Digest	Insert size is correct and no contaminated bands	[X] Pass
ORF Across Junction	Correct and consistent with target	N/A
PCR Amplification	Correct and no contaminated bands	[X] Pass
Endotoxin Level	Verified, <0.1 EU/μg (Endo-Free Preps Only)	N/A
Appearance	Clear, no foreign particles	[X] Pass
DNA Purity	Purity (A 260/A280 = 1.8 - 2.0)	[X] Pass
DNA Quantity	Actual yield (by A 260)	5μg/5μg

Comments

Delivery form: lyophilized plasmid (TE lyophilized) containing the gene insert. It is stable at room temperature for extended period of time during shipping. The lyophilized plasmid can be dissolved in sterile TE buffer or nuclease-free water (neutral pH) depending on the established laboratory practice. After reconstitution, store the stock solution at -20ºC or -80ºC for long term storage. The lyophilized plasmid dissolved in TE buffer is stable for at least 6 months at 4°C while the lyophilized DNA dissolved in water is NOT STABLE at 4°C.

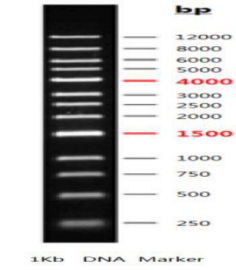
a) Before opening the tube containing the plasmid, please briefly centrifuge the tube. Lyophilized plasmid could attach to the wall of the tube. Opening without centrifugation could cause DNA loss.
b) Stock Solution: Reconstitute lyophilized plasmid (4ug or 10ug) in 40ul or 100ul of TE buffer or nuclease-free water (final concentration - 100ng/ul). To accurately determine the quantity of DNA present, please measure OD value of original stock at OD_{260}nm after reconstitution.
c) Working Solution: make a 1:10 dilution of stock solution using TE buffer or nuclease-free water (final concentration - 10ng/ul)

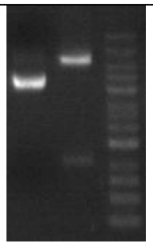
Transformation and Replating:

Transform 2ul of stock solution into appropriate *E. coli* competent cells according to standard laboratory protocol. Plate the mixture on LB agar (with desired antibiotic selection) and incubate at 37°C for overnight. Select a well separated, SINGLE colony and inoculate in LB medium with desired antibiotic selection for overnight culture. Purify Plasmid DNA from overnight cultures, verify sequences and continue with project of interest. *It is important to select only a SINGLE colony for overnight culture.

Restriction Digestion

Biomatik
Tel: (519) 489-7195, (800) 836-8089
Fax: (519) 231-0140, (877) 221-3515
Email: info@biomatik.com
http://www.biomatik.com



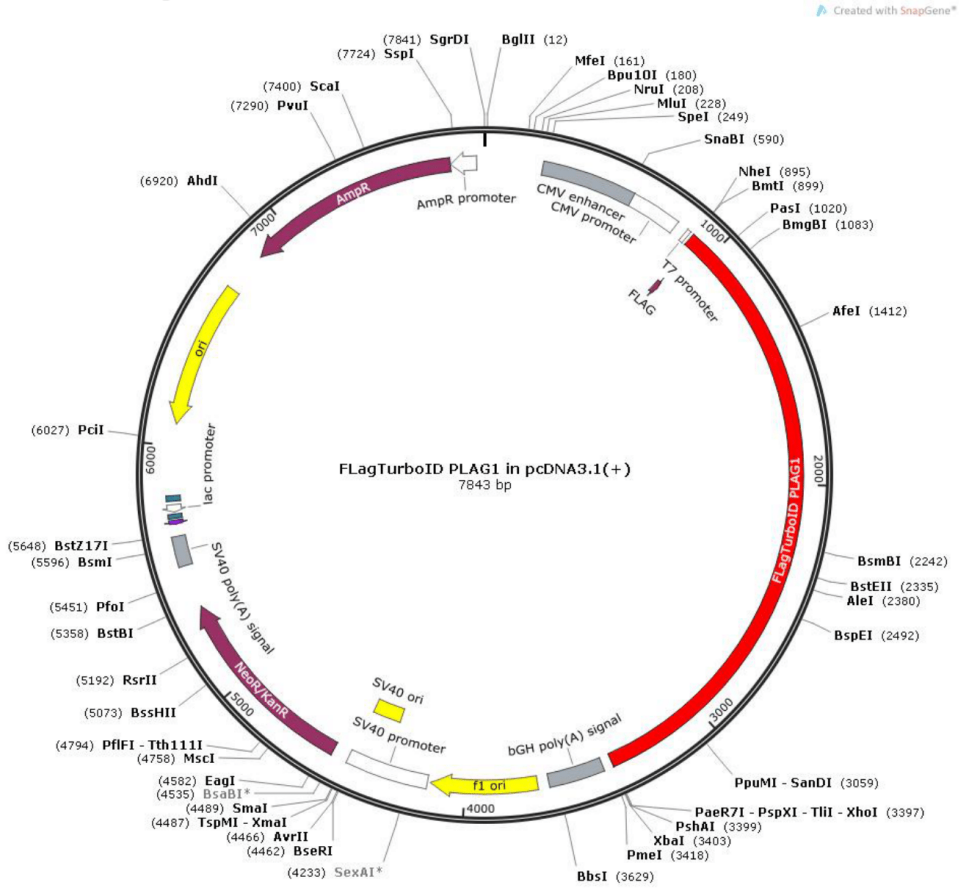


Gene name: FLagTurboID PLAG1

Clone ID#: TF92273

RES: HindIII

Construct Map:



Biomatik
Tel: (519) 489-7195, (800) 836-8089
Fax: (519) 231-0140, (877) 221-3515
Email: info@biomatik.com
http://www.biomatik.com

Detailed Sequence of the Whole Construct:

```
   1 GACGGATCGG GAGATCTCCC GATCCCCTAT GGTGCACTCT CAGTACAATC TGCTCTGATG
  61 CCGCATAGTT AAGCCAGTAT CTGCTCCCTG CTTGTGTGTT GGAGGTCGCT GAGTAGTGCG
 121 CGAGCAAAAT TTAAGCTACA ACAAGGCAAG GCTTGACCGA CAATTGCATG AAGAATCTGC
 181 TTAGGGTTAG GCGTTTTGCG CTGCTTCGCG ATGTACGGGC CAGATATACG CGTTGACATT
 241 GATTATTGAC TAGTTATTAA TAGTAATCAA TTACGGGGTC ATTAGTTCAT AGCCCATATA
 301 TGGAGTTCCG CGTTACATAA CTTACGGTAA ATGGCCCGCC TGGCTGACCG CCCAACGACC
 361 CCCGCCCATT GACGTCAATA ATGACGTATG TTCCCATAGT AACGCCAATA GGGACTTTCC
 421 ATTGACGTCA ATGGGTGGAG TATTTACGGT AAACTGCCCA CTTGGCAGTA CATCAAGTGT
 481 ATCATATGCC AAGTACGCCC CCTATTGACG TCAATGACGG TAAATGGCCC GCCTGGCATT
 541 ATGCCCAGTA CATGACCTTA TGGGACTTTC CTACTTGGCA GTACATCTAC GTATTAGTCA
 601 TCGCTATTAC CATGGTGATG CGGTTTTGGC AGTACATCAA TGGGCGTGGA TAGCGGTTTG
 661 ACTCACGGGG ATTTCCAAGT CTCCACCCCA TTGACGTCAA TGGGAGTTTG TTTTGGCACC
 721 AAAATCAACG GGACTTTCCA AAATGTCGTA ACAACTCCGC CCCATTGACG CAAATGGGCG
 781 GTAGGCGTGT ACGGTGGGAG GTCTATATAA GCAGAGCTCT CTGGCTAACT AGAGAACCCA
 841 CTGCTTACTG GCTTATCGAA ATTAATACGA CTCACTATAG GGAGACCCAA GCTGGCTAGC
 901 TCCACCATGG ATTACAAGGA TGACGACGAT AAGATCCGCA AAGACAATAC TGTGCCTCTG
 961 AAGCTGATCG CTCTCCTGGC TAATGGCGAG TTCCATAGTG GCGAACAGCT GGGAGAAACC
1021 CTGGGCATGT CCAGGGCCGC TATCAACAAG CACATTCAGA CTCTGCGCGA CTGGGGCGTG
1081 GACGTGTTCA CCGTGCCCGG AAAGGGCTAC TCTCTGCCCG AGCCTATCCC GCTGCTGAAC
1141 GCTAAACAGA TTCTGGGACA GCTGGACGGC GGGAGCGTGG CAGTCCTGCC TGTGGTCGAC
1201 TCCACCAATC AGTACCTGCT GGATCGAATC GGCGAGCTGA AGAGTGGGGA TGCTTGCATT
1261 GCAGAATATC AGCAGGCAGG GAGAGGAAGC AGAGGGAGGA AATGGTTCTC TCCTTTTGGA
1321 GCTAACCTGT ACCTGAGTAT GTTTTGGCGC CTGAAGCGGG GACCAGCAGC AATCGGCCTG
1381 GGCCCGGTCA TCGGAATTGT CATGGCAGAA GCGCTGCGAA AGCTGGGAGC AGACAAGGTG
1441 CGAGTCAAAT GGCCCAATGA CCTGTATCTG CAGGATAGAA AGCTGGCAGG CATCCTGGTG
1501 GAGCTGGCCG GAATAACAGG CGATGCTGCA CAGATCGTCA TTGGCGCCGG GATTAACGTG
1561 GCTATGAGGC GCGTGGAGGA AAGCGTGGTC AATCAGGGCT GGATCACACT GCAGGAAGCA
1621 GGGATTAACC TGGACAGGAA TACTCTGGCC GCTACGCTGA TCCGAGAGCT GCGGGCAGCC
1681 CTGGAACTGT TCGAGCAGGA AGGCCTGGCT CCATATCTGC CACGGTGGGA GAAGCTGGAT
1741 AACTTCATCA ATAGACCCGT GAAGCTGATC ATTGGGGACA AAGAGATTTT CGGGATTAGC
1801 CGGGGGATTG ATAAACAGGG AGCCCTGCTG CTGGAACAGG ACGGAGTTAT CAAACCCTGG
1861 ATGGGCGGAG AAATCAGTCT GCGGTCTGCC GAAAAGGCCA CTGTCATTCC TGGTGATTTG
1921 TCAGAAGTAA GAGATACCCA GAAAGTCCCT TCAGGGAAAC GTAAGCGTGG TGAAACCAAA
1981 CCAAGAAAAA ACTTTCCTTG CCAACTGTGT GACAAGGCCT TTAACAGTGT TGAGAAATTA
2041 AAGGTTCACT CCTACTCTCA CACAGGAGAG AGGCCCTACA AGTGCATACA ACAAGACTGC
2101 ACCAAGGCCT TTGTTTCTAA GTACAAATTA CAAAGGCACA TGGCTACTCA TTCTCCTGAG
2161 AAAACCCACA AGTGTAATTA TTGTGAGAAA ATGTTTCACC GGAAAGATCA TCTGAAGAAT
2221 CACCTCCATA CACACGACCC TAACAAAGAG ACGTTTAAGT GCGAAGAATG TGGCAAGAAC
2281 TACAATACCA AGCTTGGATT TAAACGTCAC TTGGCCTTGC ATGCCGCAAC AAGTGGTGAC
2341 CTCACCTGTA AGGTATGTTT GCAAACTTTT GAAAGCACGG GAGTGCTTCT GGAGCACCTT
2401 AAATCTCATG CAGGCAAGTC GTCTGGTGGG GTTAAAGAAA AAAAGCACCA GTGCGAACAT
2461 TGTGATCGCC GGTTCTACAC CCGAAAGGAT GTCCGGAGAC ACATGGTGGT GCACACTGGA
2521 AGAAAGGACT TCCTCTGTCA GTATTGTGCA CAGAGATTTG GGCGAAAGGA TCACCTGACT
2581 CGACATATGA AGAAGAGTCA CAATCAAGAG CTTCTGAAGG TCAAAACAGA ACCAGTGGAT
2641 TTCCTTGACC CATTTACCTG CAATGTGTCT GTGCCTATAA AAGACGAGCT CCTTCCGGTG
2701 ATGTCCTTAC CTTCCAGTGA ACTGTTATCA AAGCCATTCA CAAACACTTT GCAGTTAAAC
2761 CTCTACAACA CTCCATTTCA GTCCATGCAG AGCTCGGGAT CTGCCCACCA AATGATCACA
2821 ACTTTACCTT TGGGAATGAC ATGCCCAATA GATATGGACA CTGTTCATCC CTCTCACCAC
2881 CTTTCTTTCA AATATCCGTT CAGTTCTACC TCATATGCAA TTTCTATTCC TGAAAAAGAA
2941 CAGCCATTAA AGGGGGAAAT TGAGAGTTAC CTGATGGAGT TACAAGGTGG CGTGCCCTCT
3001 TCATCCCAAG ATTCTCAAGC ATCGTCATCA TCTAAGCTAG GGTTGGATCC TCAGATTGGG
3061 TCCCTAGATG ATGGTGCAGG AGACCTCTCC CTATCCAAAA GCTCTATCTC CATCAGTGAC
3121 CCCCTAAACA CACCAGCATT GGATTTTTCT CAGTTGTTTA ATTTCATACC TTTAAATGGT
```

Biomatik
Tel: (519) 489-7195, (800) 836-8089
Fax: (519) 231-0140, (877) 221-3515
Email: info@biomatik.com
http://www.biomatik.com

```
3181 CCTCCCTATA ATCCTCTATC AGTGGGGAGC CTTGGAATGA GCTATTCCCA GGAAGAAGCA
3241 CATTCTTCTG TTTCCCAGCT CCCCCCACAA ACACAGGATC TTCAGGATCC TGCAAACACT
3301 ATAGGGCTTG GGTCTCTGCA CTCACTGTCA GCAGCTTTCA CCAGCAGTTT AAGCACAAGT
3361 ACCACCCTCC CACGTTTCCA TCAAGCTTTT CAGTGACTCG AGTCTAGAGG GCCCGTTTAA
3421 ACCCGCTGAT CAGCCTCGAC TGTGCCTTCT AGTTGCCAGC CATCTGTTGT TTGCCCCTCC
3481 CCCGTGCCTT CCTTGACCCT GGAAGGTGCC ACTCCCACTG TCCTTTCCTA ATAAAATGAG
3541 GAAATTGCAT CGCATTGTCT GAGTAGGTGT CATTCTATTC TGGGGGGTGG GGTGGGGCAG
3601 GACAGCAAGG GGGAGGATTG GGAAGACAAT AGCAGGCATG CTGGGGATGC GGTGGGCTCT
3661 ATGGCTTCTG AGGCGGAAAG AACCAGCTGG GGCTCTAGGG GGTATCCCCA CGCGCCCTGT
3721 AGCGGCGCAT TAAGCGCGGC GGGTGTGGTG GTTACGCGCA GCGTGACCGC TACACTTGCC
3781 AGCGCCCTAG CGCCCGCTCC TTTCGCTTTC TTCCCTTCCT TTCTCGCCAC GTTCGCCGGC
3841 TTTCCCCGTC AAGCTCTAAA TCGGGGGCTC CCTTTAGGGT TCCGATTTAG TGCTTTACGG
3901 CACCTCGACC CCAAAAAACT TGATTAGGGT GATGGTTCAC GTAGTGGGCC ATCGCCCTGA
3961 TAGACGGTTT TTCGCCCTTT GACGTTGGAG TCCACGTTCT TTAATAGTGG ACTCTTGTTC
4021 CAAACTGGAA CAACACTCAA CCCTATCTCG GTCTATTCTT TTGATTTATA AGGGATTTTG
4081 CCGATTTCGG CCTATTGGTT AAAAAATGAG CTGATTTAAC AAAAATTTAA CGCGAATTAA
4141 TTCTGTGGAA TGTGTGTCAG TTAGGGTGTG GAAAGTCCCC AGGCTCCCCA GCAGGCAGAA
4201 GTATGCAAAG CATGCATCTC AATTAGTCAG CAACCAGGTG TGGAAAGTCC CCAGGCTCCC
4261 CAGCAGGCAG AAGTATGCAA AGCATGCATC TCAATTAGTC AGCAACCATA GTCCCGCCCC
4321 TAACTCCGCC CATCCCGCCC CTAACTCCGC CCAGTTCCGC CCATTCTCCG CCCCATGGCT
4381 GACTAATTTT TTTTATTTAT CAGAGGCCCG AGGCCGCCTC TGCCTCTGAG CTATTCCAGA
4441 AGTAGTGAGG AGGCTTTTTT GGAGGCCTAG GCTTTTGCAA AAAGCTCCCG GGAGCTTGTA
4501 TATCCATTTT CGGATCTGAT CAAGAGACAG GATGAGGATC GTTTCGCATG ATTGAACAAG
4561 ATGGATTGCA CGCAGGTTCT CCGGCCGCTT GGGTGGAGAG GCTATTCGGC TATGACTGGG
4621 CACAACAGAC AATCGGCTGC TCTGATGCCG CCGTGTTCCG GCTGTCAGCG CAGGGGCGCC
4681 CGGTTCTTTT TGTCAAGACC GACCTGTCCG GTGCCCTGAA TGAACTGCAG GACGAGGCAG
4741 CGCGGCTATC GTGGCTGGCC ACGACGGGCG TTCCTTGCGC AGCTGTGCTC GACGTTGTCA
4801 CTGAAGCGGG AAGGGACTGG CTGCTATTGG GCGAAGTGCC GGGGCAGGAT CTCCTGTCAT
4861 CTCACCTTGC TCCTGCCGAG AAAGTATCCA TCATGGCTGA TGCAATGCGG CGGCTGCATA
4921 CGCTTGATCC GGCTACCTGC CCATTCGACC ACCAAGCGAA ACATCGCATC GAGCGAGCAC
4981 GTACTCGGAT GGAAGCCGGT CTTGTCGATC AGGATGATCT GGACGAAGAG CATCAGGGGC
5041 TCGCGCCAGC CGAACTGTTC GCCAGGCTCA AGGCGCGCAT GCCCGACGGC GAGGATCTCG
5101 TCGTGACCCA TGGCGATGCC TGCTTGCCGA ATATCATGGT GGAAAATGGC CGCTTTTCTG
5161 GATTCATCGA CTGTGGCCGG CTGGGTGTGG CGGACCGCTA TCAGGACATA GCGTTGGCTA
5221 CCCGTGATAT TGCTGAAGAG CTTGGCGGCG AATGGGCTGA CCGCTTCCTC GTGCTTTACG
5281 GTATCGCCGC TCCCGATTCG CAGCGCATCG CCTTCTATCG CCTTCTTGAC GAGTTCTTCT
5341 GAGCGGGACT CTGGGGTTCG AAATGACCGA CCAAGCGACG CCCAACCTGC CATCACGAGA
5401 TTTCGATTCC ACCGCCGCCT TCTATGAAAG GTTGGGCTTC GGAATCGTTT TCCGGGACGC
5461 CGGCTGGATG ATCCTCCAGC GCGGGGATCT CATGCTGGAG TTCTTCGCCC ACCCCAACTT
5521 GTTTATTGCA GCTTATAATG GTTACAAATA AAGCAATAGC ATCACAAATT TCACAAATAA
5581 AGCATTTTTT TCACTGCATT CTAGTTGTGG TTTGTCCAAA CTCATCAATG TATCTTATCA
5641 TGTCTGTATA CCGTCGACCT CTAGCTAGAG CTTGGCGTAA TCATGGTCAT AGCTGTTTCC
5701 TGTGTGAAAT TGTTATCCGC TCACAATTCC ACACAACATA CGAGCCGGAA GCATAAAGTG
5761 TAAAGCCTGG GGTGCCTAAT GAGTGAGCTA ACTCACATTA ATTGCGTTGC GCTCACTGCC
5821 CGCTTTCCAG TCGGGAAACC TGTCGTGCCA GCTGCATTAA TGAATCGGCC AACGCGCGGG
5881 GAGAGGCGGT TTGCGTATTG GGCGCTCTTC CGCTTCCTCG CTCACTGACT CGCTGCGCTC
5941 GGTCGTTCGG CTGCGGCGAG CGGTATCAGC TCACTCAAAG GCGGTAATAC GGTTATCCAC
6001 AGAATCAGGG GATAACGCAG GAAAGAACAT GTGAGCAAAA GGCCAGCAAA AGGCCAGGAA
6061 CCGTAAAAAG GCCGCGTTGC TGGCGTTTTT CCATAGGCTC CGCCCCCCTG ACGAGCATCA
6121 CAAAAATCGA CGCTCAAGTC AGAGGTGGCG AAACCCGACA GGACTATAAA GATACCAGGC
6181 GTTTCCCCCT GGAAGCTCCC TCGTGCGCTC TCCTGTTCCG ACCCTGCCGC TTACCGGATA
6241 CCTGTCCGCC TTTCTCCCTT CGGGAAGCGT GGCGCTTTCT CATAGCTCAC GCTGTAGGTA
6301 TCTCAGTTCG GTGTAGGTCG TTCGCTCCAA GCTGGGCTGT GTGCACGAAC CCCCCGTTCA
6361 GCCCGACCGC TGCGCCTTAT CCGGTAACTA TCGTCTTGAG TCCAACCCGG TAAGACACGA
6421 CTTATCGCCA CTGGCAGCAG CCACTGGTAA CAGGATTAGC AGAGCGAGGT ATGTAGGCGG
6481 TGCTACAGAG TTCTTGAAGT GGTGGCCTAA CTACGGCTAC ACTAGAAGAA CAGTATTTGG
```

Biomatik
Tel: (519) 489-7195, (800) 836-8089
Fax: (519) 231-0140, (877) 221-3515
Email: info@biomatik.com
http://www.biomatik.com

```
6541 TATCTGCGCT CTGCTGAAGC CAGTTACCTT CGGAAAAAGA GTTGGTAGCT CTTGATCCGG
6601 CAAACAAACC ACCGCTGGTA GCGGTGGTTT TTTTGTTTGC AAGCAGCAGA TTACGCGCAG
6661 AAAAAAAGGA TCTCAAGAAG ATCCTTTGAT CTTTTCTACG GGGTCTGACG CTCAGTGGAA
6721 CGAAAACTCA CGTTAAGGGA TTTTGGTCAT GAGATTATCA AAAAGGATCT TCACCTAGAT
6781 CCTTTTAAAT TAAAAATGAA GTTTTAAATC AATCTAAAGT ATATATGAGT AAACTTGGTC
6841 TGACAGTTAC CAATGCTTAA TCAGTGAGGC ACCTATCTCA GCGATCTGTC TATTTCGTTC
6901 ATCCATAGTT GCCTGACTCC CCGTCGTGTA GATAACTACG ATACGGGAGG GCTTACCATC
6961 TGGCCCCAGT GCTGCAATGA TACCGCGAGA CCCACGCTCA CCGGCTCCAG ATTTATCAGC
7021 AATAAACCAG CCAGCCGGAA GGGCCGAGCG CAGAAGTGGT CCTGCAACTT TATCCGCCTC
7081 CATCCAGTCT ATTAATTGTT GCCGGGAAGC TAGAGTAAGT AGTTCGCCAG TTAATAGTTT
7141 GCGCAACGTT GTTGCCATTG CTACAGGCAT CGTGGTGTCA CGCTCGTCGT TTGGTATGGC
7201 TTCATTCAGC TCCGGTTCCC AACGATCAAG GCGAGTTACA TGATCCCCCA TGTTGTGCAA
7261 AAAAGCGGTT AGCTCCTTCG GTCCTCCGAT CGTTGTCAGA AGTAAGTTGG CCGCAGTGTT
7321 ATCACTCATG GTTATGGCAG CACTGCATAA TTCTCTTACT GTCATGCCAT CCGTAAGATG
7381 CTTTTCTGTG ACTGGTGAGT ACTCAACCAA GTCATTCTGA GAATAGTGTA TGCGGCGACC
7441 GAGTTGCTCT TGCCCGGCGT CAATACGGGA TAATACCGCG CCACATAGCA GAACTTTAAA
7501 AGTGCTCATC ATTGGAAAAC GTTCTTCGGG GCGAAAACTC TCAAGGATCT TACCGCTGTT
7561 GAGATCCAGT TCGATGTAAC CCACTCGTGC ACCCAACTGA TCTTCAGCAT CTTTTACTTT
7621 CACCAGCGTT TCTGGGTGAG CAAAAACAGG AAGGCAAAAT GCCGCAAAAA AGGGAATAAG
7681 GGCGACACGG AAATGTTGAA TACTCATACT CTTCCTTTTT CAATATTATT GAAGCATTTA
7741 TCAGGGTTAT TGTCTCATGA GCGGATACAT ATTTGAATGT ATTTAGAAAA ATAAACAAAT
7801 AGGGGTTCCG CGCACATTTC CCCGAAAAGT GCCACCTGAC GTC
```

Certified by:

Date:05/15/2023

Table 8.1. Analysis ready molecules.
The table comprises proteins identified through biotin-based proximity labelling, processed for downstream analysis in Ingenuity Pathway Analysis (IPA).

ID	Symbol	Expr Log Ratio	Expr p-value	Expr False Discovery Rate (q-value)
P01040	CSTA	10.00	0.00E+00	0.00E+00
Q7LFL8	CXXC5	10.00	0.00E+00	0.00E+00
Q09472	EP300	10.00	0.00E+00	0.00E+00
P05412	JUN	10.00	0.00E+00	0.00E+00
Q13118-2	KLF10	10.00	0.00E+00	0.00E+00
P61626	LYZ	10.00	0.00E+00	0.00E+00
P10242-4	MYB	10.00	0.00E+00	0.00E+00
P49116	NR2C2	10.00	0.00E+00	0.00E+00
Q6DJT9	PLAG1	10.00	0.00E+00	0.00E+00
O43435-3	TBX1	10.00	0.00E+00	0.00E+00
Q8NFU7	TET1	10.00	0.00E+00	0.00E+00
Q8WUH6	TMEM263	10.00	0.00E+00	0.00E+00
O75132	ZBED4	10.00	0.00E+00	0.00E+00
Q96JP5-2	ZFP91	1.24	6.45E-04	4.70E-02
P62854	RPS26	0.43	9.27E-04	4.70E-02
P11498	PC	2.07	1.06E-03	4.70E-02
Q9NRZ9-2	HELLS	1.95	1.13E-03	4.70E-02
Q5T3J3	LRIF1	1.46	1.63E-03	4.70E-02
Q9BXF3	CECR2	1.67	1.66E-03	4.70E-02
Q9Y2X9	ZNF281	2.44	1.67E-03	4.70E-02
Q6W2J9	BCOR	2.68	1.76E-03	4.70E-02
Q96DT7	ZBTB10	2.29	2.50E-03	4.85E-02
Q15723-2	ELF2	1.44	2.84E-03	5.02E-02
Q8NEZ4	KMT2C	4.31	3.42E-03	5.17E-02
Q14687	GSE1	3.09	3.43E-03	5.17E-02
Q2KHR3	QSER1	2.77	3.90E-03	5.41E-02
Q86T24	ZBTB33	2.29	4.40E-03	5.67E-02
Q9ULL5-2	PRR12	1.62	5.31E-03	6.10E-02
O00763-3	ACACB	2.23	5.40E-03	6.11E-02
O75925	PIAS1	1.87	5.46E-03	6.11E-02
Q9UQR1	ZNF148	2.15	5.50E-03	6.11E-02
O00268	TAF4	2.63	5.64E-03	6.20E-02
O60315-2	ZEB2	2.83	5.73E-03	6.20E-02
Q8IWI9	MGA	1.62	6.23E-03	6.28E-02
P50219	MNX1	1.37	6.58E-03	6.48E-02
O14497	ARID1A	3.01	7.33E-03	6.77E-02
Q8IVW6	ARID3B	1.89	7.57E-03	6.80E-02
Q02413	DSG1	1.49	7.69E-03	6.83E-02
Q92766	RREB1	1.53	7.94E-03	6.89E-02

Q04726-2	TLE3	2.71	8.32E-03	6.93E-02
O15014	ZNF609	2.06	8.69E-03	7.02E-02
P14923	JUP	1.04	9.69E-03	7.29E-02
Q9H0E3-3	SAP130	1.22	9.78E-03	7.29E-02
Q9ULK4-4	MED23	2.43	1.13E-02	7.78E-02
Q9UPN9	TRIM33	1.89	1.13E-02	7.80E-02
P62979	RPS27A	0.60	1.16E-02	7.86E-02
Q5H9F3	BCORL1	3.11	1.18E-02	7.86E-02
Q13085	ACACA	2.25	1.18E-02	7.86E-02
Q6P4R8-3	NFRKB	1.38	1.23E-02	8.09E-02
Q8NFD5	ARID1B	3.00	1.34E-02	8.17E-02
O14686	KMT2D	4.78	1.37E-02	8.17E-02
P05165	PCCA	2.33	1.38E-02	8.17E-02
O43847	NRDC	0.83	1.38E-02	8.17E-02
P17535	JUND	1.34	1.40E-02	8.17E-02
O75376	NCOR1	1.87	1.41E-02	8.18E-02
Q9NPI1	BRD7	0.70	1.44E-02	8.21E-02
Q8WXX7	AUTS2	1.44	1.47E-02	8.27E-02
P38646	HSPA9	1.07	1.71E-02	8.90E-02
P11021	HSPA5	0.72	1.73E-02	8.95E-02
P62937	PPIA	1.72	1.76E-02	9.00E-02
Q9UKD1	GMEB2	1.73	1.78E-02	9.05E-02
Q9ULM3	YEATS2	1.60	1.80E-02	9.09E-02
P15924	DSP	1.00	1.81E-02	9.09E-02
Q8NHM5-3	KDM2B	1.79	1.87E-02	9.26E-02
Q9Y618-5	NCOR2	3.12	2.00E-02	9.72E-02
Q68CP9-3	ARID2	1.52	2.02E-02	9.76E-02
P05089-3	ARG1	1.90	2.13E-02	9.77E-02
O15164-2	TRIM24	1.52	2.15E-02	9.77E-02
Q6ZW49	PAXIP1	1.13	2.23E-02	9.98E-02
Q5T749	KPRP	1.13	2.36E-02	1.04E-01
P14859-5	POU2F1	1.33	2.39E-02	1.04E-01
Q96RQ3	MCCC1	2.42	2.46E-02	1.06E-01
Q8N2W9	PIAS4	1.74	2.52E-02	1.07E-01
Q7Z3K3-5	POGZ	0.75	2.67E-02	1.12E-01
Q2TAL8	QRICH1	1.63	2.69E-02	1.12E-01
P19793-2	RXRA	1.64	2.75E-02	1.12E-01
Q9UPW6	SATB2	3.40	2.75E-02	1.12E-01
O15550	KDM6A	3.16	2.76E-02	1.12E-01
Q15652	JMJD1C	1.58	2.96E-02	1.16E-01
Q86V15	CASZ1	2.12	2.96E-02	1.16E-01
Q6ZRI6	C15orf39	2.23	3.11E-02	1.19E-01
O60281	ZNF292	1.91	3.11E-02	1.19E-01
Q14686	NCOA6	2.67	3.17E-02	1.20E-01
P13056	NR2C1	1.24	3.26E-02	1.21E-01

Q5T5X7	BEND3	3.27	3.38E-02	1.24E-01
P06702	S100A9	1.83	3.46E-02	1.26E-01
Q96RN5-3	MED15	2.61	3.58E-02	1.29E-01
Q92793	CREBBP	2.42	3.59E-02	1.29E-01
Q99697-3	PITX2	1.86	3.68E-02	1.31E-01
Q15084-3	PDIA6	0.37	3.77E-02	1.32E-01
P39880-2	CUX1	1.39	3.81E-02	1.33E-01
Q9HCK8	CHD8	1.24	4.04E-02	1.38E-01
P62263	RPS14	0.42	4.07E-02	1.39E-01
Q92610	ZNF592	0.37	4.20E-02	1.42E-01
Q7Z589	EMSY	1.56	4.22E-02	1.42E-01
P38398-4	BRCA1	0.69	4.23E-02	1.42E-01
Q99081	TCF12	1.71	4.24E-02	1.42E-01
Q9H2F5-3	EPC1	1.04	4.24E-02	1.42E-01
P46937-6	YAP1	1.33	4.31E-02	1.44E-01
P62826	RAN	0.12	4.53E-02	1.49E-01
P62888	RPL30	0.19	4.60E-02	1.51E-01
Q8NB78	KDM1B	0.80	4.61E-02	1.51E-01
Q9H334-6	FOXP1	2.33	4.67E-02	1.52E-01
P85037	FOXK1	0.46	4.80E-02	1.55E-01
Q8NEM7-2	SUPT20H	1.12	4.83E-02	1.55E-01
P11142	HSPA8	0.57	4.85E-02	1.55E-01
Q6MZP7	LIN54	1.02	4.87E-02	1.55E-01
P05166	PCCB	1.68	4.87E-02	1.55E-01

Table 8.2. Activation of Anterior HOX Genes.
The table lists proteins identified in the dataset that are associated with the activation of anterior HOX genes. The proteins include transcription factors and chromatin regulators that influence HOX gene expression.

Symbol	Entrez Gene Name	Expr Log Ratio	Expr p-value	Expr False Discover (q-value)	Expected
CREBBP	CREB binding protein	2.416	0.0359	0.129	Up
EP300	E1A binding protein p300	10	0	0	Up
JUN	Jun proto-oncogene, AP-1 transcription factor subunit	10	0	0	Up
KDM6A	lysine demethylase 6A	3.156	0.0276	0.112	Up
KMT2C	lysine methyltransferase 2C	4.311	0.00342	0.0517	Up
KMT2D	lysine methyltransferase 2D	4.78	0.0137	0.0817	Up
NCOA6	nuclear receptor coactivator 6	2.667	0.0317	0.12	Up
NCOR1	nuclear receptor corepressor 1	1.866	0.0141	0.0818	Up
PAXIP1	PAX interacting protein 1	1.125	0.0223	0.0998	Up
RXRA	retinoid X receptor alpha	1.64	0.0275	0.112	Up

Table 8.3. Transcriptional Regulation by RUNX1.
The table displays proteins from the dataset that are involved in RUNX1-mediated transcriptional regulation. This includes co-factors, chromatin remodelers, and transcriptional regulators that may interact with RUNX1.

Symbol	Entrez Gene Name	Expr Log Ratio	Expr p-value	Expr FDR (q-value)	Expected
ARID2	AT-rich interaction domain 2	1.516	0.0202	0.0976	Up
ARID1A	AT-rich interaction domain 1A	3.009	0.00733	0.0677	Up
ARID1B	AT-rich interaction domain 1B	3.001	0.0134	0.0817	Up
AUTS2	activator of transcription and developmental regulator AUTS2	1.439	0.0147	0.0827	Up
CREBBP	CREB binding protein	2.416	0.0359	0.129	Up
ELF2	E74 like ETS transcription factor 2	1.439	0.00284	0.0502	Up
EP300	E1A binding protein p300	10	0	0	Up
MYB	MYB proto-oncogene, transcription factor	10	0	0	Up
RPS27A	ribosomal protein S27a	0.596	0.0116	0.0786	Up
TCF12	transcription factor 12	1.715	0.0424	0.142	Up
YAP1	Yes1 associated transcriptional regulator	1.328	0.0431	0.144	Up

Table 8.4. Down-Regulation of the PPARα/RXRα Pathway.
This table highlights proteins that are linked to the PPARα/RXRα signalling pathway. The data suggest downregulation of this pathway, as indicated by changes in protein expression levels.

Symbol	Entrez Gene Name	Expr Log Ratio	Expr p-value	Expr FDR (q-value)	Expected
CREBBP	CREB binding protein	2.416	0.0359	0.129	Down
EP300	E1A binding protein p300	10	0	0	Down
JUN	Jun proto-oncogene, AP-1 transcription factor subunit	10	0	0	Down
MED23	mediator complex subunit 23	2.429	0.0113	0.0778	
NCOA6	nuclear receptor coactivator 6	2.667	0.0317	0.12	
NCOR1	nuclear receptor corepressor 1	1.866	0.0141	0.0818	Down
NCOR2	nuclear receptor corepressor 2	3.121	0.02	0.0972	Down
NR2C2	nuclear receptor subfamily 2 group C member 2	10	0	0	Down
RXRA	retinoid X receptor alpha	1.64	0.0275	0.112	Up

Table 8.5. Inhibition of the Sumoylation Pathway.
This table presents proteins associated with the SUMOylation pathway, a post-translational modification process that regulates protein function and stability. The dataset suggests inhibition of this pathway based on protein expression changes.

Symbol	Entrez Gene Name	Expr Log Ratio	Expr p-value	Expr FDR (q-value)	Expected
CREBBP	CREB binding protein	2.416	0.0359	0.129	Down
EP300	E1A binding protein p300	10	0	0	Down
JUN	Jun proto-oncogene, AP-1 transcription factor subunit	10	0	0	Up
MYB	MYB proto-oncogene, transcription factor	10	0	0	Down
PIAS1	protein inhibitor of activated STAT 1	1.866	0.00546	0.0611	Down
PIAS4	protein inhibitor of activated STAT 4	1.743	0.0252	0.107	Down
RAN	RAN, member RAS oncogene family	0.123	0.0453	0.149	

Table 8.6. MYB Mechanistic Network.
This table illustrates a mechanistic network involving MYB, a key transcription factor. It maps MYB's interactions with other molecules, showing activation, inhibition, or other functional relationships.

Target	Expr Log Ratio	Molecule Type	MYB	NFKBIA	JNK (family)	TP53	FOS	JUN
ACACA	2.253	Enzyme					Inhibited	
ACACB	2.231	Enzyme				Inhibited		
ARG1	1.902	Enzyme				Activated		Inhibited
BRCA1	0.686	transcription regulator	Activated		Activated	Inhibited	Affected	Affected
CASZ1	2.118	Enzyme					Affected	
CREBBP	2.416	transcription regulator					Affected	
CSTA	10	Other					Activated	Activated
HSPA5	0.721	Enzyme	Affected	Inhibited	Activated	Inhibited	Activated	
HSPA8	0.572	Enzyme	Affected	Inhibited		Inhibited		
JMJD1C	1.577	Enzyme				Inhibited		
JUN	10	transcription regulator	Affected	Activated	Activated	Inhibited	Affected	Activated
JUND	1.345	transcription regulator			Activated	Affected		
KMT2D	4.78	transcription regulator				Activated		
LYZ	10	enzyme		Affected		Activated		
MYB	10	transcription regulator	Activated					Activated
NCOR2	3.121	transcription regulator				Activated		
NR2C1	1.237	transcription regulator				Affected		
PCCA	2.326	enzyme				Activated		
PDIA6	0.368	enzyme				Activated		
PIAS1	1.866	transcription regulator					Affected	
RAN	0.123	enzyme				Affected		
RPS26	0.429	Other				Activated		
RXRA	1.64	ligand-dependent nuclear receptor			Inhibited		Activated	Activated
S100A9	1.827	Other		Affected		Activated	Inhibited	
TCF12	1.715	transcription regulator		Activated				

TET1	10	enzyme		Activated	Affected	
YAP1	1.328	transcription regulator	Inhibited			
ZEB2	2.829	transcription regulator		Inhibited		Affected

Table 8.7. JUN Mechanistic Network.
This table describes the mechanistic interactions of JUN, a component of the AP-1 transcription factor complex. It outlines how JUN interacts with other regulatory proteins, transcription factors, and signalling moleculesThe table includes information on activation and inhibition relationships.

Target	Expr Log Ratio	Molecule Type	JUN	JNK (family)	TGFB1	SPI1	SNAI2	FOS	TP53
JUN	10	transcription regulator	Activated	Activated	Activated	Inhibited		Affected	Inhibited
AUTS2	1.439	Other			Inhibited				
NCOR2	3.121	transcription regulator							Activated
RPS26	0.429	Other							Activated
PITX2	1.861	transcription regulator			Activated				
DSP	1	Other			Affected		Inhibited		
PDIA6	0.368	enzyme							Activated
ZEB2	2.829	transcription regulator	Affected		Activated				Inhibited
YAP1	1.328	transcription regulator			Activated		Inhibited		
KLF10	10	transcription regulator			Activated				
PIAS1	1.866	transcription regulator						Affected	
JUND	1.345	transcription regulator		Activated	Inhibited				Affected
ZNF281	2.435	transcription regulator			Activated				
HSPA8	0.572	enzyme				Affected			Inhibited
RXRA	1.64	ligand-dependent nuclear receptor	Activated	Inhibited	Activated			Activated	
CSTA	10	Other	Activated					Activated	
ARG1	1.902	enzyme	Inhibited		Activated	Activated			Activated
CREBBP	2.416	transcription regulator						Affected	
PCCA	2.326	enzyme							Activated

RAN	0.123	enzyme							Affected
CASZ1	2.118	enzyme						Affected	
JMJD1C	1.577	enzyme							Inhibited
BRCA1	0.686	transcription regulator	Affected	Activated	Inhibited		Inhibited	Affected	Inhibited
TCF12	1.715	transcription regulator			Affected				
TET1	10	enzyme						Affected	Activated
S100A9	1.827	Other						Inhibited	Activated
JUP	1.038	Other			Inhibited		Inhibited		
MYB	10	transcription regulator	Activated		Inhibited	Inhibited			
TRIM33	1.887	transcription regulator			Inhibited				
GSE1	3.086	Other			Affected				
LYZ	10	enzyme				Activated			Activated
HSPA5	0.721	enzyme		Activated	Activated			Activated	Inhibited
KMT2D	4.78	transcription regulator			Inhibited				Activated
ACACB	2.231	enzyme							Inhibited
NR2C1	1.237	transcription regulator							Affected
PCCB	1.677	enzyme			Affected				
ACACA	2.253	enzyme			Activated			Inhibited	
ARID1A	3.009	transcription regulator				Affected			