A GENOMIC, PROTEOMIC AND TRANSCRIPTOMIC INVESTIGATION INTO GROWTH HORMONE ACTIONS

LEKHA JAIN

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Biomedical Science, the University of Auckland, 2020

Abstract

Interactions between genomic loci have been implicated in playing a critical role in the regulation of gene transcription. Growth hormone (GH) is a peptide hormone predominantly produced in the pituitary that is crucial for normal growth and metabolism. GH actions are activated through binding to its cell-surface receptor, the GH receptor (GHR), with consequent activation of downstream signalling cascades. Genetic variation in the GH locus genes and aberrant GH signalling have been implicated in diseases such as cancer. This led to the hypothesis that genetic variation that occurs in the GH locus may potentially result in alteration of gene regulatory mechanisms. Common SNPs in the regions across the GH locus that physically interacted with genes, both in cis and trans (intra and inter-chromosomal), were identified. These SNPs were also found to be associated with altered expression of these genes. This suggests that regions encompassing the GH locus function as regulatory hubs for multiple genes, some of which are involved in cellular and cancer pathways, related to GH/GHR signalling.

Nuclear localisation of GHR and its association with increased tumorigenesis has previously been reported, but the specific consequences of this phenomenon remain unclear. To determine whether there is a functional role for the GHR in the nucleus, combined immunoprecipitation-mass spectrometry was used to assess whether the GHR interacts with proteins in the nucleus and microarray analysis was used to determine the consequent effects of this nuclear import on gene expression. Multiple proteins were found to localise with the GHR, including two transcription factors, HMGN1 and SUMO1. Moreover, targets genes of HMGN1 and SUMO1 were found to be differentially expressed following GH treatment. This suggests that GHR nuclear translocation potentially serves as an auxiliary mechanism for regulation of GH actions.

The competitive endogenous RNA (ceRNA) theory suggests that miRNA can be sequestered by lncRNA through competition for shared binding sites, causing changes in miRNA targeted gene expression. GH/GHR signalling has been linked to altered non-coding RNA transcription profiles. As such, a GH-induced ceRNA network was constructed to elucidate the molecular regulatory mechanisms involved in mediating GH functions in a mammary epithelial cell line.

This thesis provides novel insights into the mechanisms underlying GH actions, including a functional role for the GHR in the nucleus and the coordinated regulation of endogenous RNA networks.

Dedicated to my parents, Deepak and Ranjana Jain

Acknowledgements

I am immensely grateful for every bit of this PhD experience—the highs and the lows—and am very thankful to everyone who has been a part of it. I am grateful and indebted to:

Dr. Jo Perry, for being the kindest, most generous, and helpful supervisor one could hope to learn from. A big thanks to Associate Prof. Justin O'Sullivan, for always pushing me to think more, and Prof. Mark Vickers, for always having time and offering guidance

The Liggins support staff, particularly Jean Leonard, Elizabeth McIntosh and Eric Thorstensen, for your efficiency and generosity in all things organisational

Man Lu, Yue, Karla, and Julia for all the lab-chats and coffee-breaks, and to *JOS* team members, particularly Tayaza and William, for all the insightful discussions around work and life

Vishakha and Dharani, for being my dearest friends. Thank you for your endless love, support, and generosity in getting me through these years

Elsie, Pania, and Shikha, for always being there to share sadness and happiness with cake and drink

Aratrika and Ellyn, for helping me settle into a new country, always making time for me, and providing friendship outside of my work bubble

The entire Jain clan, for their love and generosity in all forms. A very special thanks to Chote Dada for always being my guiding light and for always looking out for me

The Poeschls, for their unwavering love, warmth, and kindness. Thank you for being my *Familie* here. A special thanks to JP for being the most beautiful companion

My brother, Anuj, for being my lifelong supporter and voice of reason. Thank you for always going out of your way to make my life easier

My partner, Lorenz, for being my pillar of love and strength through everything. Words cannot describe how grateful I am for you

My parents, Deepak and Ranjana. Thank you for always believing in me. Without your unconditional love, support, and sacrifices for me, I wouldn't be where I am in life. For that, I dedicate this thesis to you.

Table of Contents

List of Figures

List of Tables

Glossary

ATCC	American tissue culture collection
CoDeS3D	Contextualizing Developmental SNPs using 3D information
CTCF	CCCTC-binding factor
EGF/R	Epidermal growth factor/receptor
eQTL	Expression quantitative trait locus
ERK	Extracellular signal-regulated kinase
FC	Fold change
FDR	False discovery rate
GH1	Growth hormone
GHR	Growth hormone receptor
GTEx	Genotype-Tissue Expression
GWAS	Genome-wide association study
Hi-C	High-throughput chromosome conformation capture
HMGN1	High Mobility Group Nucleosome Binding Domain 1
IGF1	Insulin-like growth factor-1
JNK	c-Jun N-terminal kinase
Kb	Kilo base
kDa	Kilo Dalton
KEGG	Kyoto encyclopedia of genes and genomes
LCR	Locus control region
MAPK	Mitogen-activated protein kinase
Mb	Mega base
mRNA	Messenger RNA
PI3K	Phosphoinositide 3-kinase
SNP	Single nucleotide polymorphism
SUMO1	Small ubiquitin-like modifier 1
TAD	Topologically associated domain

CHAPTER 1

General Introduction

1.1 Overview

The role of genetics in increased susceptibility to many diseases has been accepted for decades now. Essential processes, such as cell differentiation, metabolism, and development, are associated with coordinated changes in gene expression profiles. Genetic changes (both coding and intragenic) that alter gene expression can contribute to disease risk (Kilpinen et al., 2013). It is now widely accepted that a majority of the non-coding genome is involved in regulating cellular functions, either directly or indirectly. While the mechanism of action of non-coding transcripts ,such as microRNA (miRNA) and long non-coding RNA (lncRNA), is still not completely understood, there is no doubt regarding their significance (Esteller, 2011).

Changes in chromatin organisation, structure, and interactions are found to play a crucial role in the regulation of gene expression (Marti-Renom, & Mirny, 2013; Gibcus & Dekker, 2013). Regulation of gene promoters can be mediated through both proximal and distal enhancer regions, and the latter may be located on different chromosomes (*trans* interactions). Thus, the impact of intragenic genetic variation needs to be considered in the context of the 3D structure of the genome. Non-random and non-linear arrangement of chromosomes into discrete and precise units within the nucleus of a cell serves to bring together distal genomic loci in close proximity of each other. This facilitates physical interactions between these regions, the consequences of which are not understood very well. These long-range contacts are often involved in regulation of gene expression, and may contribute to genome stability (Ciabrelli and Cavalli, 2015; Rao et al., 2014). Therefore, genetic variation underlying regulatory regions can potentially disrupt these contacts (Kilpinen et al., 2013; Krijger and De Laat, 2016). Investigating variations in the context of long-range gene regulation might provide some novel insights into disease mechanisms.

The growth hormone/insulin-like growth factor-1 (GH/IGF1) axis performs the crucial function of mediating normal growth and metabolism, along with several additional functions critical to normal physiology (Roelfsema and Clark, 2001). Compromised GH/IGF1 signalling and polymorphisms in genes related to this axis are associated with multiple disease states. GH exerts its impact on cell growth and differentiation by activation of downstream signalling cascades, following binding with the GH receptor (GHR) (Waters, 2016). In addition, it is now apparent that the GHR translocates into the nucleus following

stimulation with GH, although the significance of this is unclear (Lobie et al., 1994a). Signal transduction pathways activated by GH mediate effects by alteration of gene transcription profiles, by direct stimulation of transcription, and by regulation of histone modifications (Álvarez-Nava and Lanes, 2017; Dehkhoda et al., 2018; Rotwein and Chia, 2010). The *GH* gene locus consists of five evolutionarily related genes clustered together on chromosome 17 (*GH1*, *GH2*, *CSH1*, *CSHL1*, and *CSH2*). Tissue-specific expression pattern of this gene cluster is controlled by an upstream locus-control region, through the process of chromatin looping (Ho et al., 2004; Tsai et al., 2016). This gives rise to the possibility that some GH functions may be mediated by spatial interaction of the *GH* gene locus with distal loci. Therefore, the study of polymorphisms across the *GH* locus could potentially elucidate novel regulatory networks involving genes associated with GH/IGF function, and identify novel functions of this essential hormone axis.

This thesis was an attempt to unravel the complexity of GH-induced functions by dissecting spatial interactions, the coding and the non-coding genome, and the consequence of GHR nuclear translocation. In this Chapter, concepts of spatial organisation of the genome and the consequences of genetic variation in non-coding regions are introduced.

1.2 Chromatin organisation and gene expression

Development and evolution of multicellular organisms is subject to the cells' capability to acquire new fates. A specific cell fate can be characterised by an amalgamation of cellular regulatory and expression pathways that result in a specific phenotypic outcome (Davidson, 2010). Environmental stimuli, signal transduction pathways, transcriptional regulation, and cell–cell interactions are some of the factors that drive cell-fate decisions. External stimuli (for example, hormones) activate transduction of signalling pathways, which results in binding of transcription factors to specific DNA motifs, causing either activation (via enhancers) or silencing (via repressors) of gene transcription (Lambert et al., 2018; Vaquerizas et al., 2009).

The DNA in the nucleus is wrapped around histone proteins to form a nucleosome, which gives rise to the characteristic "beads on a string" appearance. These nucleosomes form the repeating units of chromatin (Laskey et al, 1978). Chromatin can be broadly termed as heterochromatic or euchromatic based on how densely or sparsely the DNA is packaged into

a nucleosome. Heterochromatin refers to the highly condensed and inaccessible domain in which the genes are poorly expressed, and euchromatin is the highly accessible, transcriptionally active domain. It was previously thought that heterochromatin was completely and irreversibly inactive, but more recent studies have shown that this region is characterised by low transcriptional activity and can sometimes revert to a euchromatin state, if repressive conditions are eliminated (Feng and Michaels, 2015; Politz et al., 2013; van Steensel and Belmont, 2017). Based on permanency of this condensation process, it is categorised as either facultative heterochromatin or constitutive heterochromatin. Facultative heterochromatin consists of chromatin regions that are poorly expressed due to certain cellular or epigenetic modifications, whereas constitutive heterochromatin corresponds to permanently condensed chromatin regions. These regions are characterised by the presence of modified histone residues, which are often the markers for gene silencing, tandem repeats, and satellite DNA (Haddad et al., 2017; Politz et al., 2013).

Pairs of four different types of histone proteins, namely H2A, H2B, H3, and H4, form the building block of each nucleosome molecule along with H1 histone and DNA (Luger et al., 1997) (Figure 1.1). Each histone protein has a unique function in formation and maintenance of nucleosomes (Bannister and Kouzarides, 2011). These histones often undergo post-translational modifications (e.g. acetylation, methylation and phosphorylation) at multiple sites. Specific modifications in these histone proteins are responsible for epigenetic changes and modulation of the genetic material by formation of functionally distinct clusters (Hahn, Dambacher, & Schotta, 2010, Cairns, 2009). These modifications are dynamic, and are highly dependent on a variety of factors, such as cell type, timing, stimulus, signalling conditions, and the availability of modifying enzymes. There is a vast array of enzymes involved in modification of histones, which vary depending on the type of modification. These processes function in accordance with the type of DNA machinery involved. This characteristic feature of histone modifications dictates its role in regulation of essential nuclear functions, such as gene transcription, regulation of gene expression, cell cycle, and DNA repair (Zhang and Reinberg, 2001).

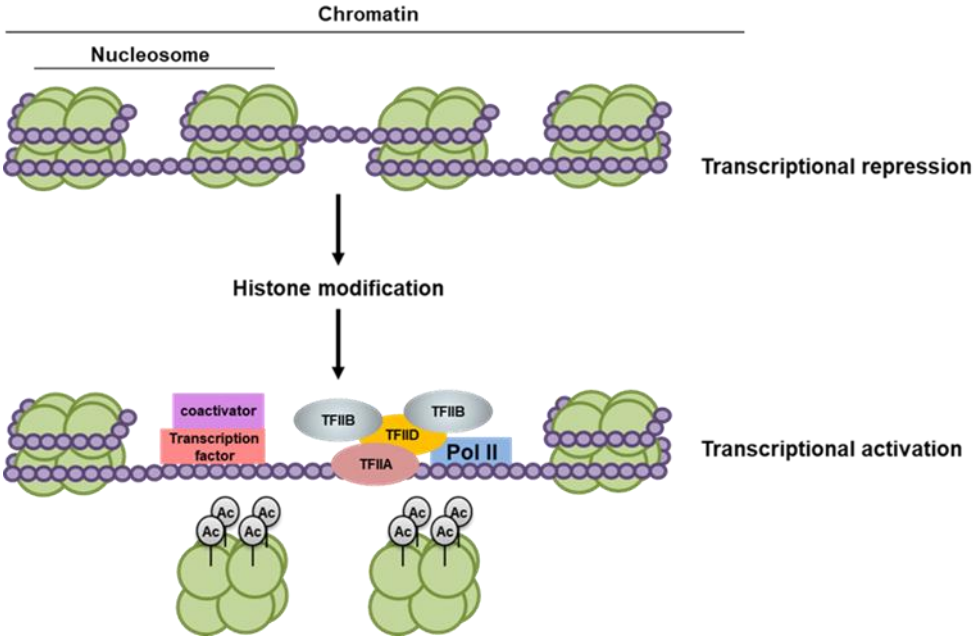

Figure 1.1. Histone modifications as regulators of transcription.
Unmodified histones (top) tightly interact with DNA, making regulatory regions, such as, promoters and enhancers inaccessible, resulting in transcriptional repression. Histone modifications, such as, acetylation (bottom) can disrupt this histone-DNA interaction, making regulatory regions accessible to the transcriptional machinery, involving RNA polymerase (Pol II), transcription factors (TFIIs), and coactivators, resulting in activation of transcription. Adapted from (McGee and Hargreaves, 2011).

1.2.1 Gene expression and regulation

Gene transcription in eukaryotes is characterised by interactions between several units in the cell nucleus forming a transcription complex in a coordinated and dynamic manner. This complex includes gene promoters, which function as the drivers of gene expression and are often found at the 5'-end of the gene. Other factors which become proximal to the promoter site include proteins, such as multiple transcription factors and cofactors, and DNA elements, such as enhancers and repressors (Lemon & Tjian, 2000; Nightingale et al, 2006). Enhancers serve as activators of gene expression and function by direct physical contact with the promoters. Because of this characteristic, it was previously assumed that enhancers were found in close proximity to the gene they regulate. More recently, though, multiple studies have shown that enhancers are often found far away from the gene, sometimes even on a different chromosome (Sanyal et al., 2012). Long-distance enhancers and promoters are brought together for regulation of gene expression by complicated three-dimensional

mechanisms and are dependent on a multitude of factors (Furlong and Levine, 2018; Sanyal et al., 2012). Enhancers often contain several transcription factor binding sites, which facilitate their function as stimulators of gene expression by recruitment of relevant transcription factors and RNA polymerase II. An important point to remember is that these interactions are often highly tissue specific, which would explain differences in gene expression levels across multiple tissues and also during different developmental stages (He et al., 2012).

Enhancer sequences are often characterised by an open chromatin configuration, which makes these sequences more accessible for interactions in the cell nucleus. Other epigenetic markers for these gene regulatory regions involve the presence of dense clusters of DNase I hypersensitive sites (Gross and Garrard, 1988; Thurman et al., 2012). DNase I hypersensitive sites (DHS) are markers of open chromatin, which are susceptible to digestion by DNase I endonuclease. These regions serve as distinct markers of cis-regulatory elements, such as enhancers, promoters, locus control regions (LCRs) and insulators (Chen and Chen, 2019; Ma et al., 2014). As mentioned previously, the mode of histone modification can also be used to characterise enhancers in a cell type (Barski et al., 2007). Addition of one methyl group and acetylation of histone H3 (H3K4me1, H3K27ac) has been found to be hallmarks of active enhancers (Cremer et al, 2000; Creyghton et al., 2010). Besides these, there are a multitude of histone modifications that have been found in different enhancers in different tissues and can be used to distinguish between active and repressed enhancers. However, some modifications are not specific to enhancers and can be attributed to other regulatory regions in the genome, such as promoters, open chromatin, and insulators (Cremer et al., 2000; Koch et al., 2007).

Several regulatory sequences that are involved in regulation of genes crucial for development and survival are often found to be conserved across species, but this parameter alone cannot be a reliable factor for enhancer prediction (Berthelot et al., 2018; Meireles-Filho and Stark, 2009). A combination of these characteristics can, however, be helpful in a more confident prediction of presence of an enhancer region, which can then be corroborated by *in vitro* analysis, for example through using a luciferase promoter assay.

Regulation of gene expression, in particular assembly of the transcription apparatus, is yet to be understood properly. Though enhancers and promoters may be located large distances from each other, sometimes even on different chromosomes, physical contact between these

regulatory elements is crucial for regulation of gene transcription. This is often found to be facilitated by formation of chromatin loops (Marsman and Horsfield, 2012), as the next section of this chapter discusses.

1.2.2 Chromosomal organisation in the cell nucleus

Previously believed to be linear, it is now well-established that chromosomal arrangement in the nucleus of a cell is three-dimensional (3D) and highly precise (Dekker and Mirny, 2016). This 3D conformation has crucial functional consequences in every facet of the nuclear process (Bonev and Cavalli, 2016). Each chromosome exists within discrete and spatially confined territories known as chromosomal territories (Figure 1.2). The boundaries of each territory restricts overlapping of different chromosomes (Branco and Pombo, 2006). The nuclear positioning of these territories varies across different cell types, and is dependent on the genetic and epigenetic features of the genome involved. Chromosomes that contain a higher density of genes and transcriptionally active loci (euchromatin) tend to cluster towards the centre of the nucleus, whereas the ones with low gene density and an abundance of silenced genes (heterochromatin) are preferentially positioned near the periphery of the nucleus (Croft et al., 1999; Yu and Ren, 2017). The significance of this preferential placement in genome architecture and genome interactions is not yet identified.

Genomic regions in the cell nucleus are also found to be partitioned into two distinct compartments: A and B compartments. Gene loci that are found in compartment A are characterised by high gene density and are transcriptionally active. These regions are rich in epigenetic markers, such as H3K36 methylation and DNase I hypersensitive sites. Loci found in compartment B are characterised by the presence of transcriptionally repressed regions (Haddad et al., 2017; Politz et al., 2013). Epigenetic modifications, cell type, and cellular processes can lead to switching of genomic regions between the two compartments (Dixon et al., 2015; van Steensel and Belmont, 2017). Studies have demonstrated that this A–B compartment switching can be induced by transcription factors and chromatin modifiers (Therizols et al., 2014; Wijchers et al., 2016). An example of this phenomenon could be reprogramming of B-cells into induced pluripotent stem cells, brought about by A–B compartment switching. Such switching occurs because of the changes in expression patterns

of CCAAT/enhancer-binding protein alpha (C/EBPα) and four transcription factors, OCT4, SOX2, KLF4, and MYC (Stadhouders et al., 2018).

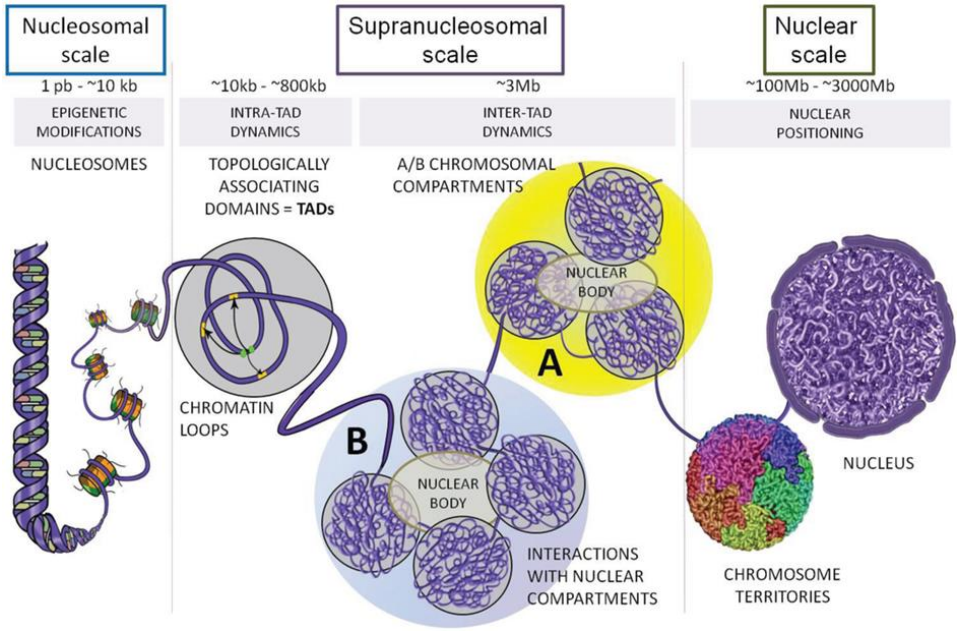

Figure 1.2. Representation of genome organization at different scales in mammals
The nucleosomal scale is composed of nucleosomes (repeating units of histone and DNA). The supranucleosomal scale, which contains chromosomal compartments and TADs, is key in understanding 3D interactions of the genome, and leads to formation of chromosomal territories (nuclear scale). Taken from (Ea et al., 2015)

Chromosomal organisation into territories and compartments undergo further arrangement into modular units within the nucleus, which has been determined using Hi-C. Hi-C is a type of chromosome capture technique that identifies regions of the genome which are in close physical proximity of each other at the given point. It can be considered a snapshot of the genome in 3D at a specific point. Chromatin is folded in such a way that it forms discrete spatial units of interacting chromosomes that are referred to as "topologically associating domains" (TADs) (Lonfat and Duboule, 2015; Symmons et al., 2014) (Figure 1.2 and 1.3). These domains are characterised by internal territorial interactions, rather than interactions within different domains, and can vary from a few kilo-bases to a few mega-bases in size (Dekker and Heard, 2015; Lonfat and Duboule, 2015). Unlike chromosomal territories and

compartments, TAD positioning is very stable and is conserved evolutionarily (Ciabrelli and Cavalli, 2015).

The boundaries of each TAD are often defined by insulator and housekeeping genes. Other factors that mark TAD boundaries are histone methylation, transcription start sites, and transfer RNA genes. These borders might undergo significant changes (particularly when exposed to different conditions), which could be responsible for altered activity profiles across different cell types (Dixon et al., 2012). Well characterised insulators that mark TAD boundaries include CCCTC binding factor, CTCF, and the Cohesin complex. CTCF is a DNA binding regulatory element that functions as an insulator motif and facilitates formation of chromatin loops to mediate long-range spatial interactions (Rao et al., 2014). A well-studied example for this phenomenon is the physical interaction between the promoter region of a gene and its long-distance enhancer region, which can drive regulation of gene expression (Ji et al., 2016; Ma et al., 2014).

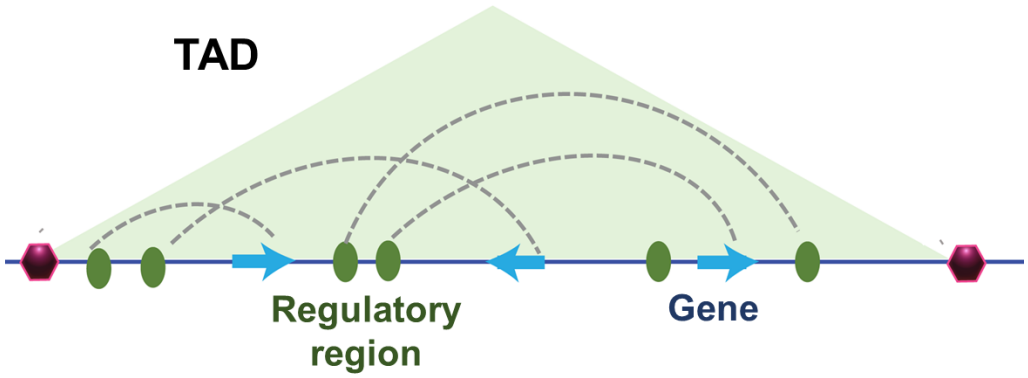

Figure 1.3. Topologically associating domain (TAD)
A simple drawing of a TAD (represented by the triangular region), showing internal interactions between a gene and its regulatory region. Boundaries of TADs are defined by housekeeping genes (purple blocks) and insulators, such as the CTCF motif.

The features of chromatin material form the basis of clustering of different genes into a specific TAD (Tang et al., 2015). Highly accessible euchromatin forms active TADs, which contain transcriptionally active genes, associated factors, and regulatory elements. This clustering is highly variable across different cell types, owing to differences in their gene expression profiles. This reflects highly on the specificity of the 3D genome organisation (Dixon et al., 2016). External interactions within active regions of different TADs can occur, which gives rise to the now widely accepted phenomenon of chromatin looping (Sanborn et

al., 2015). Interactions between distal loci, such as between promoters and enhancers, take place by physical looping of chromatin material, bringing these regions into close proximity, thus enabling physical interaction (Jin et al., 2013). This looping is mediated by dimer formation of the insulator protein, CTCF (Dixon et al., 2012). This phenomenon also substantiates the theory of co-regulation of different genes by the same set of regulatory elements, which may be located several mega-bases apart, or even on different chromosomes. These interactions between distal loci are termed "trans" interactions (Lieberman-Aiden et al., 2009; Ma et al., 2014) (Figure 1.4).

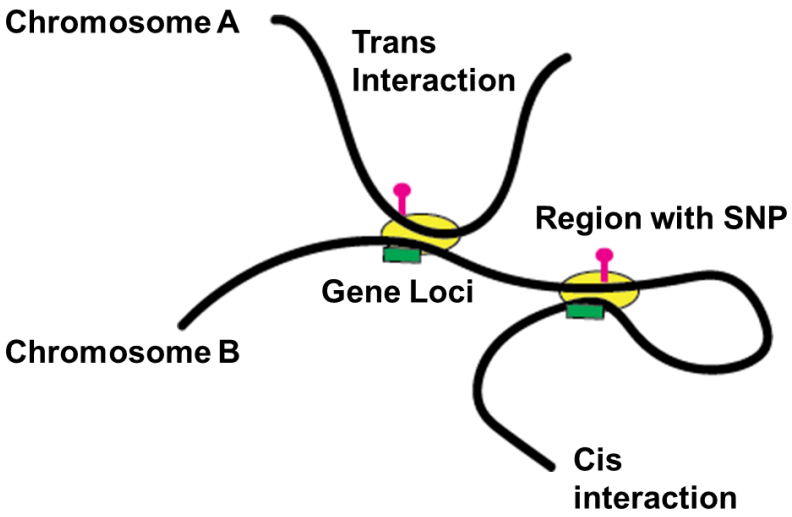

Figure 1.4. Representation of chromatin looping and interactions
Regulatory regions of chromosomes, such as gene promoters (green) and enhancers (yellow), are brought into close proximity of each other and may regulate gene expression. These interactions can be in cis (loci located nearby, <1MB) or in trans (distal loci, >1MB on the same chromosome, or located on a different chromosome). Presence of a SNP (pink) may disrupt these interactions, resulting in altered gene expression profile.

1.2.3 Genome-wide association studies and eQTLs

In recent years, there has been an increasing emphasis on study of genetic variation with respect to increased/decreased susceptibility to disease among a population or in an individual. The magnitude of this approach made it necessary to aggregate data on a population level. Genome-wide association studies (GWAS) have proven to be an important tool for identification of subsets of common single nucleotide polymorphisms (SNPs), which

are associated with different traits related to a specific disease (Hirschhorn and Gajdos, 2011; Manolio, 2010).

Most of the disease-associated variants have been found in non-coding genomic regions instead of coding regions, as was predicted earlier. This trend would imply that these variants are more likely to be involved in regulating gene expression. Further, most of these variants were not found in close proximity to the locus affected (Manolio, 2010). In certain cases, these variants were found to be located thousands of base-pairs away. For example, SNPs within the Fat-mass and obesity associated protein (FTO) gene were found to be associated with *IRX3* expression, which is located around 50,000 bases downstream. Abundant experimental data links obesity-associated SNPs within the *FTO* gene to changes in expression levels of the *IRX3* gene, which indicates that the *FTO* locus is involved in long-distance regulation of *IRX3* (Smemo et al., 2014).

Another important feature of this approach was that these associations were found to be specific to cell types and tissues. When a change in gene expression level is linked to genetic variation, it is known as an expression quantitative trait locus (eQTL). This is a simple approach that integrates common genetic variation with a measurable phenotype, and can be applied to a large number of individuals and populations. This approach has the potential to identify novel loci associated with disease states (Dixon et al., 2007; Emilsson et al., 2008; Storey et al., 2007). *Cis*-eQTLs are characterised by the distance from the affected gene being under 1 mega-base, whereas *trans*-eQTLs refer to long-distance interactions above 1 mega-base on the same chromosome, or those occurring with other chromosomes. These eQTLS can be observed across single or multiple tissues, which is to be expected, considering the complex nature of gene regulatory networks (Grundberg et al., 2012; Stranger et al., 2012).

Since GWAS does not include sequencing of the entire genomes, the identified variants associated with disease-risk may not be causal/driver variants but could just be inherited (linkage disequilibrium) with the actual causal variant/variants (Kellis et al., 2014; Schierding et al., 2014; Tak and Farnham, 2015). Some of these identified variants can instead be involved in inactivation of enhancers by disruption of transcription factor-binding sites, resulting in transcriptional suppression of the target genes, which then could lead to a disease state (Benko et al., 2009; Fukami et al., 2012; Lettice et al., 2003). These regulatory actions can be explained by the 3D conformation of the genome. For example, the *β-globin* genes are under the regulatory control of five regions that function together to increase *β-globin* gene

expression by ~25–100 fold. Deleting any of these individual sites reduces this expression by only two-fold (Tolhuis et al., 2002). The explanation for this regulation lies in the 3D structure as each of these regions physically contacts the active *β-globin* genes and potentially forms a regulatory hub to spatially control the expression of the target genes (Tolhuis et al., 2002; Van De Werken et al., 2012).

Therefore, study of variants in the context of the 3D genome is crucial to progress from GWAS to characterisation of the molecular mechanisms underlying the disease. Recent studies carried out by the O'Sullivan group reiterate this point by integrating the analysis of disease-associated variants with physical interactions in the 3D genome, in order to gain novel insights into the genetic architecture of diseases and potential therapeutic measures (Fadason et al., 2017; Gokuladhas et al., 2020; Nyaga et al., 2018; Schierding et al., 2020).

1.3 The non-coding genome

It is important to understand the elegant regulatory circuit that exists between the coding and the non-coding genome before delving into genetic association studies, considering the obscure nature of GWAS data. Coding genes account for roughly 2% of the total genome, but multiple studies suggest that at least 70% of the genome is transcribed. This suggests that a major chunk of the human transcriptome is composed of non-coding RNA (Dunham et al., 2012). Previously believed to be "junk" DNA, the significance of the non-coding genome has been recognised only in the past few decades. Evolutionary studies conducted on a genome-wide scale have further corroborated the importance of non-coding RNAs by illustrating that, more often than not, alterations in non-coding RNA expression, rather than in protein-coding genes, is what separates different species from each other (Frith et al., 2005; Shabalina and Spiridonov, 2004). Aberrations in the non-coding genome transcription and regulation have been implicated in metabolic dysfunction and diseases, such as cancer. However, functional characterisation of the majority of non-coding RNA has not been carried out.

1.3.1 Classification of the non-coding RNAs

Non-coding RNAs vary in length, transcription mechanisms, and functions. They are classified on the basis of their size: small non-coding RNAs include micro RNA (miRNA), small interfering RNA (siRNA), small nuclear RNA (snRNA), small nucleolar RNA (snoRNA), and transfer RNA (tRNA). The functions of these small non-coding RNA extend

from ribosomal processing, to messenger RNA (mRNA) processing, post-transcriptional RNA silencing, and splicing (Salmena et al., 2011; Tay et al., 2014). Longer non-coding RNAs include long non-coding RNA (lncRNA) and pseudogenes, which are implicated in the regulation of gene activity, both in *cis* and trans (Nagano and Fraser, 2011; Sabin et al., 2013). Later in this chapter, miRNA and lncRNA will be discussed in further.

1.3.2 MicroRNA

miRNAs are small (approximately 18-24 nucleotides long) and are the most well characterised class of non-coding RNAs. These are implicated in a variety of cellular processes and malignancies, and are a potential target for therapies. Multiple studies have reported that miRNAs are linked to different cancer types, such as breast cancer, gastric cancer, lung cancer, prostate cancer, and pancreatic cancer (Catto et al., 2011; Lin et al., 2010; Mulrane et al., 2013; Rachagani et al., 2010; Shrestha et al., 2014). The first miRNA, *lin-4* , was described in *Caenorhabditis elegans*, which was responsible for post-transcriptional silencing of *lin-14* mRNA (B Wightman et al., 1993; Lee et al., 1993). This was followed by the discovery of *let-7*, first in *C. elegans* and then in *Drosophila melanogaster* and humans, making it one of the most widely studied miRNAs in human development and disease today (Lee et al., 2016).

miRNAs are either encoded as single or multiple miRNAs in a group, or within intergenic regions of a coding gene. The classical pathway of miRNA biogenesis begins with RNA polymerase II transcribing these regions into primary miRNA (pri-miRNA), which then undergoes pre-processing to form precursor miRNAs (pre-miRNAs). Pre-miRNAs are characterised by a single hairpin structure. This pre-processing involves loading of pri-miRNA to an enzyme complex that includes Drosha, the double-stranded RNA (dsRNA)-binding protein (dsRBP), and DiGeorge critical region 8 (DGCR8) as the major proteins. This is followed by translocation of pre-miRNAs to the cytoplasm with the help of a transporter protein called exportin 5 (EXP5), and cytoplasmic processing of pre-miRNA to a dsRNA by a protein known as Dicer. This dsRNA is then loaded onto a protein complex comprising the Argonaute (AGO) protein family, which leads to degradation of one strand of dsRNA, leaving the mature miRNA. This process of miRNA maturation is called RNA-induced silencing complex (RISC) loading (Bartel, 2004; Chaulk, 2011; Lund et al., 2004; Maas, 2012; RW Carthew, 2009; T Conrad, 2014; Treiber et al., 2019).

Target sites of miRNAs are generally located in the 3′ UTR of mRNAs (Bartel, 2009). miRNAs usually cause gene silencing by repressing protein translation in some cases and mRNA decay in others (Jonas and Izaurralde, 2015). The process of mRNA decay is irreversible and accounts for ~ 80% of the gene silencing effect of miRNAs (Guo et al., 2010). This is often promoted by the process of deadenylation, which results in increased susceptibility of mRNA to degradation by exoribonuclease enzymes (Braun et al., 2012). The processes involved in inhibition of translation initiation and translational repression are mediated by multiple proteins, such as eukaryotic initiation factor 4 A (eIF4A-I and II) (Fukao et al., 2014). Interestingly, one miRNA can target multiple genes, and potentially several genes present in a cellular pathway. A single gene can also be targeted by multiple miRNAs (Mestdagh et al., 2010; Selbach et al., 2008; Uhlmann et al., 2012) .

Although deregulation of miRNA is implicated in multiple cancer types and developmental disorders, it is difficult to assign a specific biological role to individual miRNAs. In the majority of cases, knockdown of a single miRNA does not yield a particularly dramatic effect, even in its direct targets. This has been observed for even the most abundantly expressed miRNA. Exceptions to this observation occur when the direct target of a miRNA is extremely critical to cellular functioning, so that even a small alteration in its expression can lead to severe consequences. For example, hsa-miR-128 targets and de-represses multiple members involved in MAPK signalling pathway. Knockout of this miRNA in mice caused a fatal epileptic phenotype, as the resulting de-repression of crucial genes led to a massive increase in ERK-2 phosphorylation (Tan et al., 2013).

The reasons behind the subtle effects of miRNA deactivation could be four-fold. Firstly, most genes contain recognition sites for multiple miRNAs, which may not all belong to the same family, and which may work together to silence the gene. Therefore, knockdown of a single miRNA may not produce a significant change in the gene expression level. Secondly, subtle changes in gene transcription and its regulation are often endured well by an organism, given the presence of compensating mechanisms and regulatory pathways. Thirdly, most functional miRNA studies involve overexpression and knockdown in laboratory cell-based models, meaning that the complexities of a multicellular environment are not always represented. Overexpression studies in animals can often lead to expression of miRNA in tissues in which it is generally not expressed, which can alter the miRNA-target relations (Chi et al., 2009). Lastly, most miRNAs share their sequence, often even the pre-miRNA sequence, with other miRNA which may not always function together (Ambros et al., 2003). This suggests that

inactivation of a single miRNA would likely only cause a modest change in the expression of its direct targets.

These observations indicate that most miRNA potentially function as micromanagers of gene translation and serve as rheostats that adjust the gene transcription machinery to balance phenotypes that are associated with different pathways and mechanisms (Bartel and Chen, 2004; Ebert and Sharp, 2012; Hornstein and Shomron, 2006). Cancer often being a consequence of dysregulated cellular machinery could explain why miRNAs are associated with so many diverse types of cancer (Reddy, 2015).

1.3.3 Long non-coding RNA

As studies on miRNAs and their significance gained momentum, the existence and importance of lncRNAs remained obscure for decades. The first lncRNA gene, called *H19*, was discovered in the 1980s, but was initially classified as a coding RNA (Pachnis et al., 1984). This gene was linked to the newly discovered phenomenon of genomic imprinting, along with another gene present in the same cluster, *Igf2* gene. The puzzling aspect of this gene was the absence of translation despite the presence of an open reading frame (Barlow et al., 1991; Bartolomei et al., 1991; Brannan et al., 1990). This was followed by characterisation of another lncRNA encoded by the X-inactivation centre (Xic) locus known as *Xist* (X-inactive-specific transcript). *Xist* was found to be involved in dosage compensation by initiation of X- chromosome silencing. This was followed by a multitude of chromatin and epigenetic changes which results in systematic repression of inactivated X chromosome (Borsani et al., 1991; Brown et al., 1991; Gendrel and Heard, 2014; Loda and Heard, 2019). These non-coding genes brought the biological significance and functional relevance of lncRNA into the spotlight. Since then, multiple lncRNAs have been identified through next-generation sequencing and more advanced transcriptomic techniques.

Broadly, RNA transcripts which are more than 200 nucleotides in length, lack a conserved open reading frame, and are non-homologous to annotated protein sequences are called lncRNAs (Harrow et al., 2012). Similar to coding genes, lncRNA genes are often composed of multiple exons and can be transcribed into different transcript isoforms, but the length of the exons is longer when compared to coding genes (Derrien et al., 2012). Human lncRNA gene promoters tend to be enriched in AT nucleotides and deficient in CG which is in contrast to the coding gene promoter makeup (Alam et al., 2014). lncRNAs are

predominantly categorised based on their genomic location in relation to the protein-coding genes. lncRNAs can be classified as intergenic (lincRNAs), transcribed from a promoter sequence, either independent (plncRNAs) or shared with a protein-coding gene (pancRNAs) and transcribed from an enhancer (eRNAs) (St.Laurent et al., 2015).

A crucial thing that distinguishes lncRNA from mRNA is its biogenesis. The striking feature of lncRNA expression lies in its precision. LncRNA biogenesis is very specific to cell-type and cell-state (Akerman et al., 2017; Batista and Chang, 2013; Flynn and Chang, 2014; Morán et al., 2012). The transcription of the majority of human lncRNAs is mediated by RNA polymerase II. An exception to this is the human neuroblastoma associated *NDM29* gene which requires RNA polymerase III for its synthesis (Massone et al., 2012). The majority of lncRNAs are reported to undergo post-transcriptional modifications, such as 5′ end capping and 3′-end polyadenylation, although this largely depends on the source of lncRNA. Some lncRNA transcripts, like *MALAT1,* exist in both processed and unprocessed forms (Djebali et al., 2012). While the transcription machinery for biosynthesis of most lncRNA resembles that of coding genes, the promoters of both are quite distinct (Quinn and Chang, 2016).

LncRNAs often function as regulators of coding gene expression, both in *cis* as well as *trans*. Previous studies have reported the relevance of lncRNA in cellular processes, development and disease. LncRNAs have the potential to interplay with chromatin and epigenetic machinery to drive gene expression (scaffolding) (Hanly et al., 2018; Quinn and Chang, 2016). For example, lncRNA *HOTAIR* interacts with transcription complexes and redirects them to different locations, thus altering histone modifications and gene transcription (Tsai et al., 2010). A class of lncRNAs known as guide lncRNAs can form RNA-DNA triplexes to regulate transcription of specific genes, for example, *MEG3* lncRNA can guide transcription of genes related to the TGFβ pathway (Mondal et al., 2015). lncRNAs can also function as decoy molecules by blocking DNA and/or protein binding sites which can have multiple functional consequences. A classic example of this mechanism is lncRNA *GAS5* which blocks glucocorticoid receptor binding to its DNA response element, thus blocking expression of genes related to the glucocorticoid pathway (Kino et al., 2010).

LncRNAs may also function as decoys for miRNA. The competitive endogenous RNA (ceRNA) hypothesis states that lncRNAs can sequester miRNA by competing for shared binding sites, which results in increased expression of the miRNA target gene (Salmena et al.,

2011; Szcześniak and Makałowska, 2016; Tay et al., 2014). This hypothesis may provide an explanation for the complexity of gene transcription and its correlation with protein translation. This is a very popular theory that is being widely explored currently, but its claim that significant physiological alterations can occur due to the sponging effect is still controversial. Theories against this hypothesis state that sequestration of miRNA by a single lncRNA is likely not sufficient enough to result in a significant spike in target gene expression, also considering the subtle physiological effects of most miRNAs (Broderick and Zamore, 2014; Thomson and Dinger, 2016). Further exploration into the mechanisms of action of different RNAs, and into the effect of competition between them in a cellular context, is needed to confirm the general applicability of the ceRNA hypothesis.

1.4 Growth Hormone

Human GH is a peptide hormone that is synthesised by the somatotrophic cells of the anterior pituitary gland (Evans et al, 2016). Pulsatile secretion of GH from the anterior pituitary is modulated by GHRH and somatostatin produced in the hypothalamus, with the former exerting a positive effect on GH secretion and the latter inhibiting its secretion. Another hormone released from the stomach lining and the hypothalamus, ghrelin, has a positive effect on secretion of GH (Boguszewski, 2003). Following secretion from the pituitary, GH binds to its cell surface receptor, GHR, a type I cytokine receptor, which induces a cascade of downstream signal transduction. GH also mediates its functions by promoting the secretion of IGF1 in the liver, which inhibits GH secretion in a negative feedback loop (Ohlsson et al., 2009; Perrini et al., 2010) (Figure 1.5).

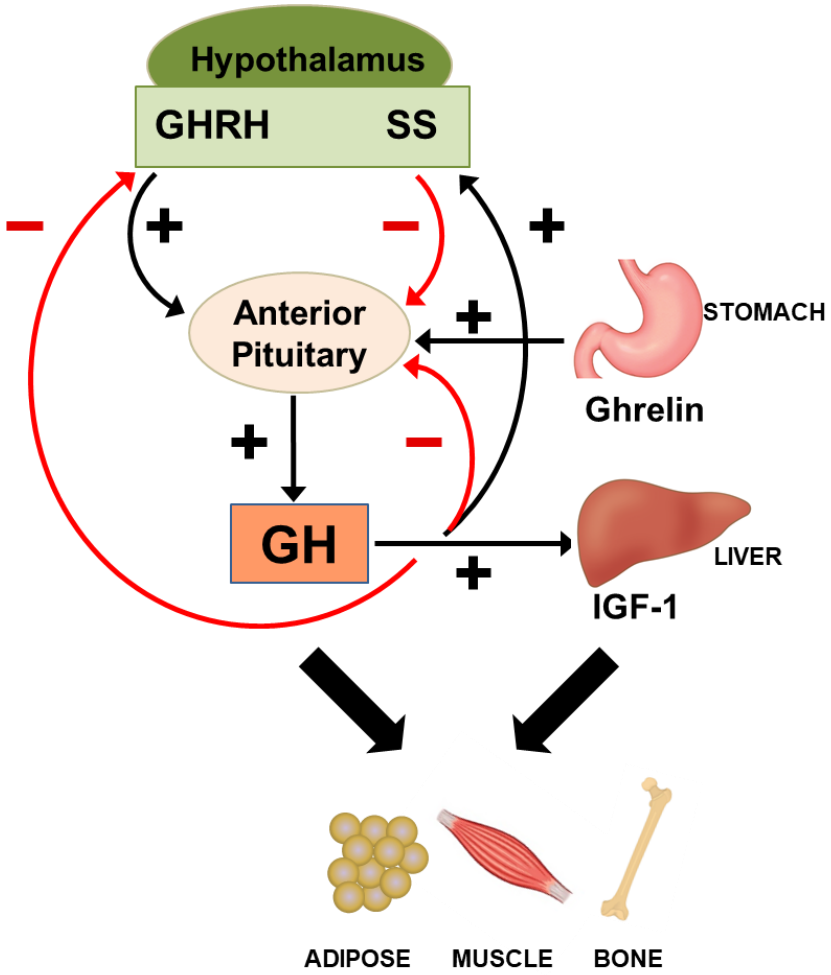

Figure 1.5. The feedback loop involved in GH secretion and function.
GH is released from the anterior pituitary, after which it stimulates the release of IGF1 in the liver. IGF1 mediates some of the functions of GH, such as lipolysis, bone and muscle growth. GHRH and Ghrelin have a positive effect on secretion of GH, whereas somatostatin inhibits its secretion. IGF1 further limits GH release through a negative feedback loop.

As the name suggests, GH has a crucial function in mediation of normal longitudinal growth in childhood and puberty. GH induces the secretion of IGF1 in liver which acts on the growth plate and promotes cellular growth and metabolism (Giustina et al, 2008; Le Roith et al, 2001; LeRoith & Yakar, 2007). GH plays an important role in metabolism. It is a glucogenic hormone and works to increase blood glucose levels (Kim et al., 2012). GH also is also a major lipolytic hormone, it converts complex lipids to fatty acids (Møller et al., 2009; Perrini et al., 2010). Other functions attributed to the GH/IGF1 axis involve increase in fat and muscle mass, regeneration, and repair of skeletal muscle cells by increasing protein synthesis. It also regulates bone growth by promoting chondrogenesis and increasing bone formation

(Bonert and Melmed, 2017; Giustina et al., 2008). Levels of GH progressively decrease with age; consequently it has been proposed that GH has anti-ageing effects (Hermann and Berger, 2001). However, recent studies have demonstrated that lower levels of GH are associated with healthy ageing (Gesing et al., 2017).

1.4.1 *GH* locus and gene structure

GH1 gene is located on the long arm of chromosome 17 (chromosome 17q23) and is part of a cluster of five genes located between galactokinase and thymidine kinase (Xu et al., 1988). This gene cluster is composed of an array of homologous, tissue-specific genes, namely *GH1* (also known as GH-N), *GH2* (earlier known as *GH-V*), and chorionic somatomammotropin genes (*CSH1, CSH2 and CSHL1)*. *GH1* is expressed predominantly in the pituitary, whereas the rest of the locus genes are expressed in the syncytiotrophoblast layer and invasive trophoblast cells of the placenta (Chen et al., 1989; Liao et al., 2018).

The *GH1* gene is approximately 1,640 bp long, and composed of five exons and four introns. The GH protein is 191 amino acid residues in length, with two disulphide bridges, and a molecular mass of 22 kDa (Niall et al, 1971). There are three major isoforms of *GH1*: the full-length 22kDa isoform (the main product); a 20 kDa variant with a deletion of amino acid residues 32-46 (accounts for 5-10% of expressed transcripts), which is generated as a result of alternative splicing on exon 3; and a third isoform with a deletion of amino acid residues 32-71, which is generally expressed under pathological conditions (Baumann, 2009).

From an evolutionary perspective, there are several noteworthy facts pertaining to the *GH* locus. Firstly, all five of the genes are structurally similar, i.e. they consist of five exons that are separated by four introns. Secondly, the introns all occur at the same sites in all the genes. Lastly, these five genes occur in the same transcriptional orientation. This suggests a very high structural homology (Ho et al, 2002).

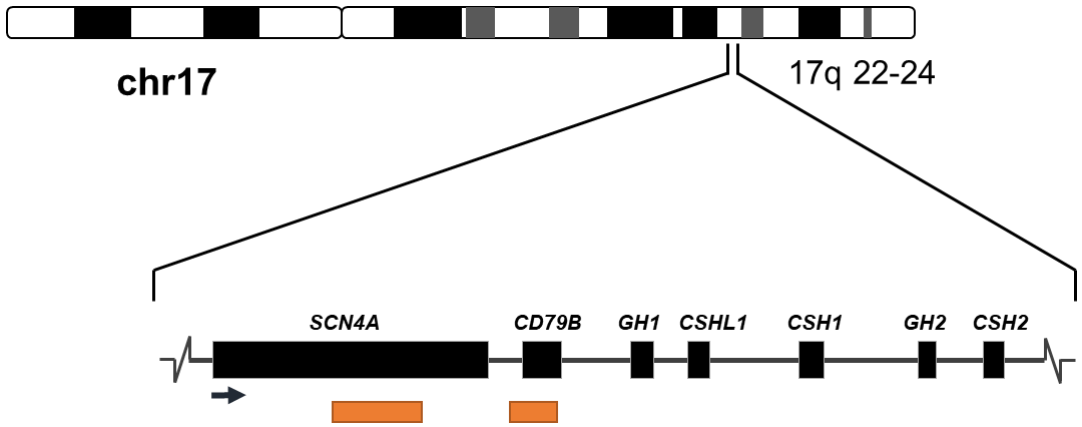

Figure 1.6. The GH gene cluster on chromosome 17
The GH gene cluster includes five evolutionarily related genes, *GH1*, *GH2*, *CSH1*, *CSH2*, and *CSHL1*, controlled by a locus control region situated upstream of this cluster (orange block). The locus control region is responsible for tissue-specific expression of these genes.

The expression of genes within the *GH* locus is regulated by a region that lies 15-32 kb upstream of the *GH1* gene, now referred to as the locus control region (LCR) (Ho et al., 2002) (Figure 1.6). Five DNase I hypersensitive sites located 5'- of the *GH1* promoter form major elements of this region. This LCR is a *cis*-acting regulatory site that is believed to drive/repress expression of these genes in pituitary and placenta in a coordinated manner (Tsai et al, 2016) (Figure 1.7).

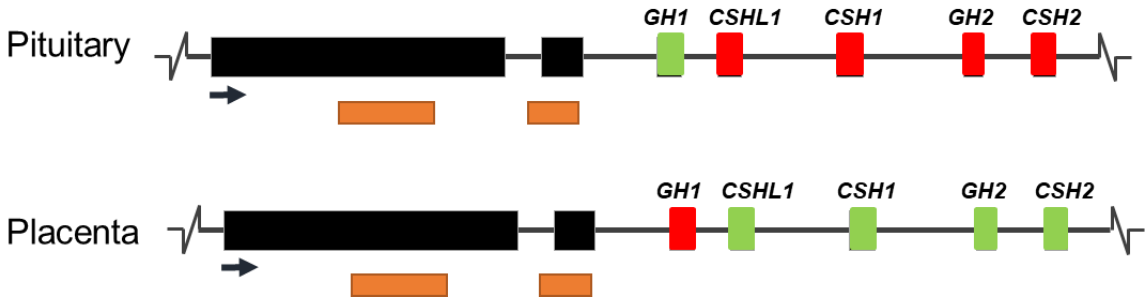

Figure 1.7. Tissue-specific expression of the *GH* gene cluster
GH1 is expressed predominantly in the pituitary, whereas the other four genes are expressed in the placenta. Expressed genes are in green, and switched off genes are in red. The locus control region responsible for this coordinated expression is represented by the orange block below.

In the pituitary, chromatin looping brings HSI and HSII into close proximity of the *GH1* promoter region (Figure 1.8). This interaction drives the expression of *GH1*, switching off expression of other genes in the process. In the placenta, HSI and HSII are repressed, which turns off transcription of *GH1*. HSIII, HSIV and HSV are activated in the placenta and interact with the promoters of *GH2*, *CSH1*, *CSH2*, and *CSHL1* through a long-range looping mechanism, which results in expression of these genes (Tsai et al., 2016).

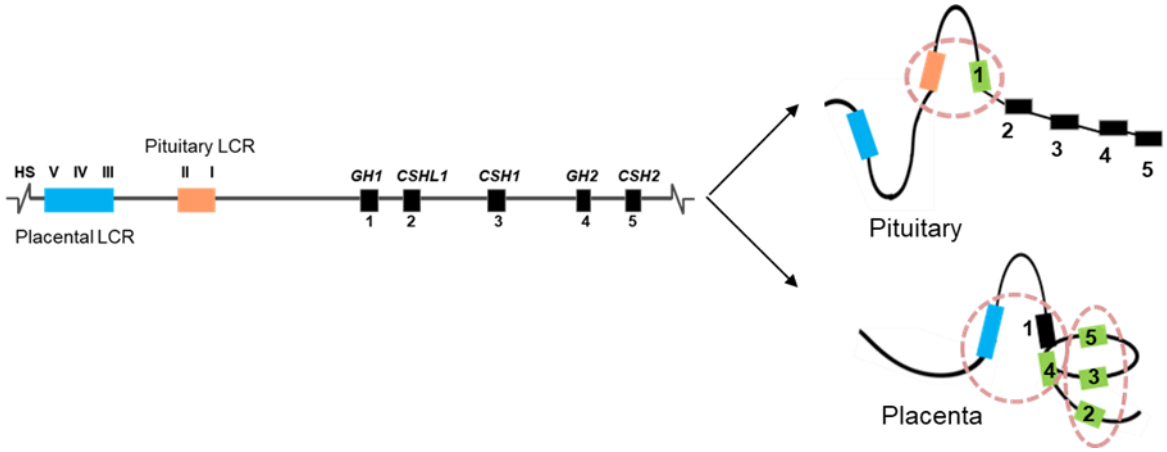

Figure 1.8. Differential expression of *GH* locus genes in the pituitary and placenta
The looping mechanism for interaction between the locus control region (LCR) and the *GH* gene cluster results in differential gene expression in the pituitary and placenta. The orange block represents the pituitary LCR, which is composed of two hypersensitive (HS I and II), and the blue block is the placental LCR with HS III, IV, and V. HS sites I and II are brought into close proximity of the *GH1* gene promoter as a result of chromatin looping. This causes the expression of *GH1* to be turned on in the pituitary, while the expression of the other four genes is turned off. Conversely, in the placenta, the expression of *GH2*, *CSH1*, *CSH2* and *CSHL1* is turned on, whereas *GH1* gene expression is turned off. Adapted from (Tsai et al., 2016).

1.4.2 Mediators of GH action and related proteins

1.4.2.1 GHR: structure and activation

The human *GHR* gene is located at chromosomal location 5p13.1-p12 (Barton et al, 1989). The structure of this gene comprises nine exons, with multiple additional exons in the 5'-untranslated region. (Godowski et al., 1989). The GHR protein is encoded by nine coding exons, which code for a 638 amino acid protein, including an 18 amino signal peptide. The mature protein is 620 amino acids and contains an extracellular domain (encoded by exons 3

to 7), a single transmembrane domain, and a cytoplasmic domain. As mentioned above, the main mechanism of GH action is mediated through the activation of the GHR and downstream signalling pathways. The receptor is part of the cytokine type I receptor family, which is characterised by being single pass transmembrane receptors containing a minimum of one classic cytokine receptor homology (CRH) domain. This domain, in turn, has two fibronectin III-like (FNIII) domains, each containing an immunoglobulin-like β sandwich domain with seven strands in two layers. These two FNIII domains form two asymmetrical ligand binding sites, which have different binding affinities. Upon ligand binding, the dimerisation interface between two receptor subunits gives rise to a site 3, which has been shown to be essential for signal transduction, and is particularly important for the GHR. The GHR is one of the simplest of these receptors, as it contains only a single CRH domain. The intracellular domain of this receptor family consists of a proline-rich Box-1 motif, which facilitates binding of JAK2 molecules to the inner membrane, along with a Box-2 motif. The GHR has no intrinsic kinase activity and relies on recruitment of non-receptor tyrosine kinases JAK2 and c-SRC for activation (Waters, 2016).

The extracellular domain of the GHR is known as GH-binding protein (GHBP). GHBP is found in both humans and animals, but one remarkable difference is that in humans it is generated by cleavage, whereas in animals it is formed by alternative splicing. It forms a complex with GH in human plasma, and most likely functions to prolong the half-life of GH, and, furthermore, modulates GH bioactivity by competing with GHR for GH. Three isoforms of GHR are generated in humans by alternative splicing: a full length form, an isoform with a deletion of exon 3 (d3GHR), and a smaller isoform (GHRtr) that codes for a truncated protein lacking 97.5% of the GHR intracellular domain (Amit et al., 1997). The expression and abundance of these three isoforms of GHR is found to be highly tissue specific and can vary from person to person. The different expression/abundance patterns may correspond to varied sensitivity to GH; however, it is still unclear whether deletion of exon 3 alters sensitivity to GH.

1.4.2.2 GH/GHR signalling pathways

As discussed earlier, the GHR exists as a pre-formed homodimer in the cell membrane. GH binds to the extracellular domain of GHR via two sites, 1 and 2, with higher affinity towards site 1. GH binding is followed by conformational changes in the intracellular domain of the

receptor which leads to activation of the classical signalling protein tyrosine kinase, JAK2, thus initiating an entire cascade of downstream signal transduction (Waters, 2016) (Figure 1.9). Key signalling pathways pertaining to GH-GHR system include JAK2 signalling via STATs (1, 3 and 5). This pathway involves phosphorylation of important tyrosine residues, as well as recruitment and phosphorylation of STATs, primarily STAT5 (5A and 5B) by binding to their SH2 domains. Following phosphorylation, STAT5 forms a dimer and translocates into the nucleus, to regulate gene expression by binding to STAT-responsive DNA elements. GH also signals through the MAPK pathway, as JAK2 leads to activation of ERK via adaptors SHC, GRB and, SOS, which in turn regulates transcription of target genes. ERK can also be activated by PLCγ, Ras, and c-SRC. Other important signalling cascades through which GH functions are the JNK pathway, in which JAK2 leads to upregulation of the non-receptor tyrosine kinase c-SRC. This, in turn, can activate JNK, resulting in gene transcription via the RAP and PI3K pathway. This pathway functions by phosphorylation of IRS by JAK2, resulting in the formation of docking sites for GRB2, SHP2 and PI3K, which leads to the activation of mTOR (Carter-Su et al., 2016; Lu et al., 2019).

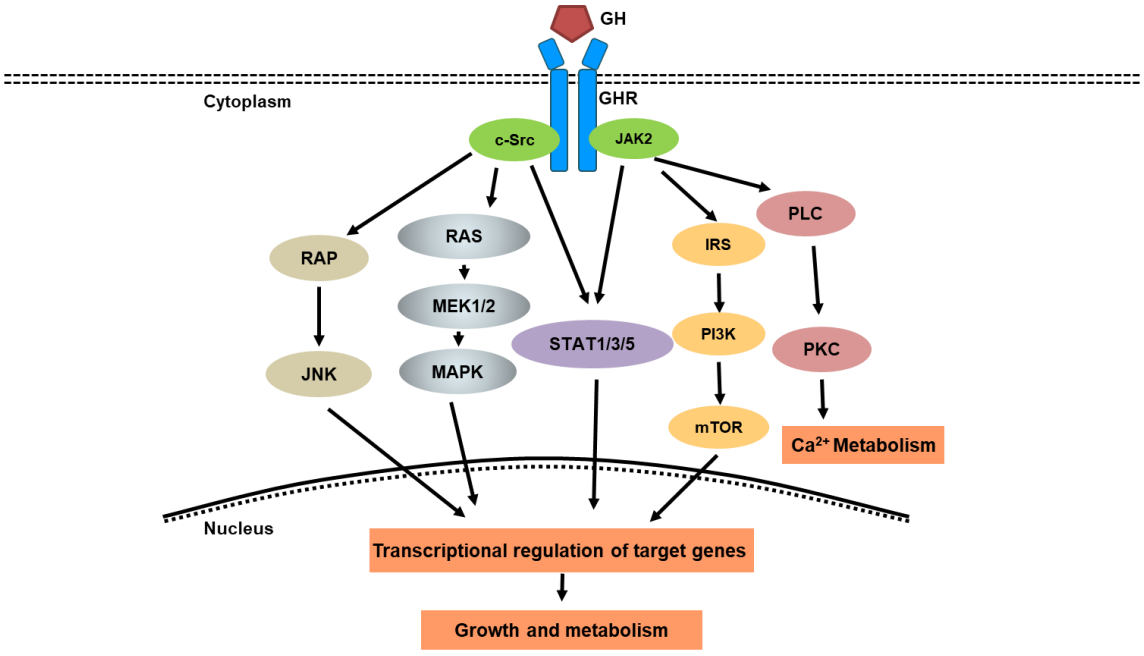

Figure 1.9. Signalling pathways activated by binding of GH to the GHR
Adapted from (Waters, 2016)

1.4.3 GHR internalisation and degradation

Following activation by the GH, the GHR is downregulated in the cell through several mechanisms. These includes internalisation of the receptor, ubiquitination, nuclear translocation and proteasomal degradation. GH binding to GHR leads to the internalisation of GH/GHR complex (Lobie et al., 1994b; Sachse et al., 2001). Internalisation of the GHR is mostly carried out by endocytosis. Endocytosis involves transport of molecules into the cell via engulfment of cell membrane and formation of vesicles (Du Toit, 2015). This can occur through clathrin-dependent, caveolae-dependent, and clathrin/caveolae-independent pathways (Aguilar and Wendland, 2005). Common targets of these internalised vesicles include lysosomes via late endosomes for degradation, endosomes for recycling to the cell membrane and signal transduction, as well as the nucleus for direct DNA binding and/or regulation of gene transcription (Aguilar and Wendland, 2005; McMahon and Boucrot, 2011; Du Toit, 2015). One of the major destinations for the internalised GH/GHR complex is the lysosomes, where the GHR is degraded (Van Kerkhof et al., 2000). This is crucial for maintenance of GHR levels in the cell in a normal physiological state (Strous and Van Kerkhof, 2002). It was previously suggested that the GH/GHR complex was recycled to the cell-surface, but this is likely to be minimal (Sachse et al., 2001). Lobie *et al*. showed that one of the cellular destinations of this internalised complex is the nucleus, and that the internalisation of GH is subject to GHR binding, as it does not occur in cells transfected with only the extracellular part of GHR (GHBP) (Lobie et al., 1994b).

Several studies have suggested that internalisation of GH/GHR is possibly mediated by clathrin or caveolae-related pathways (Lobie et al., 1999; Sachse et al., 2001; Yang et al., 2004). The salient finding implicates suppressor of cytokine signalling 2 (SOCS2) protein in mediation of ubiquitination for proteasome-dependent degradation of GHR. SOCS2 is a component of a ubiquitin ligase that is induced by GH stimulation and is a crucial down-regulator of GHR expression and signal transduction (Greenhalgh et al., 2002). GHR activation leads to phosphorylation of Y487 residue of GHR. This site (pY487) functions as a binding site for both STAT5 and SOCS2. Accumulation of SOCS2 protein causes an increased binding of SOCS to pY487 site, resulting in competitive inhibition of STAT5 binding and downstream signalling. Further, SOCS2 increases ubiquitination of GHR, which causes an increase in proteasomal degradation of GHR, consequently decreasing the levels of GHR in the cell (Greenhalgh et al., 2005; Vesterlund et al., 2011). Other residues important

for SOCS binding are tyrosine residue (Y595) and threonine residue (P495). A mutation in P495 residue (GHRP495T) is linked to a ~50% increase in susceptibility to lung cancer. This mutation results in increased GHR signalling and reduced GHR degradation due to downregulation of SOCS2 binding (Chhabra et al., 2018). These observations highlight the importance of internalisation and degradation of GHR in normal cellular physiology.

1.4.3.1 Insulin-like Growth Factors

GH exerts its physiological actions through stimulation of IGF1 production in multiple tissues, but primarily in the liver. IGF1, in turn, represses the production of GH by a negative feedback loop, which functions by stimulating somatostatin release in the hypothalamus and inhibiting *GH* gene transcription in the pituitary (Boguszewski, 2003). There are six major binding proteins that preferentially bind to *IGF1* mRNA, which are known as Insulin like Growth Factor Binding Proteins (IGFBP-1 to -6). IGFs and IGFBPs are also locally produced in many other tissues where autocrine or paracrine mechanisms take place, and where IGFs actions are also regulated by IGFBP proteases (Samani et al, 2007). These binding proteins serve different functions, the most common of which is to increase the half-life of circulating IGF1. IGFs are often found to be associated with a high molecular weight complex (a ternary complex), consisting of IGFBP-3 and the IGF acid labile subunit (IGFALS) in plasma (Bach, 2018). IGF1 mediates its functions when the ternary complex dissociates; IGFBP-IGF are removed from the circulation and cross the endothelium to reach the target tissues, where IGF1 binds to the IGF1 receptor (IGF1R) and initiates downstream cell signalling pathways. Similar to GH levels, IGF1 levels rise during juvenile life and then start to decline after puberty. The IGF1R is a hetero-tetramer formed by two identical a-subunits and two identical b-subunits, and is a member of the tyrosine kinase growth factor receptor family (Hakuno and Takahashi, 2018; Rother and Accili, 2000). It has high sequence similarity with the insulin receptor (IR), and cross-talk between these receptors and their signalling pathways has been demonstrated (Křížková et al., 2016). IGF1 and insulin can bind to each other's receptor but with lower affinity. Hybrid receptors of IGF1R and IR are often formed and interestingly exhibit higher binding affinity for IGF1 than for insulin (Pandini et al., 2002). These hybrid receptors are often implicated in diseases (Jin et al., 2019; Nagle et al., 2018).

Another important member of GH/IGF axis is insulin-like growth factor-2 (IGF2). It is a mainly neonatal growth factor, unlike its counterpart IGF1. Circulating IGF2 levels are highest in the foetal circulation but are still considerably high in adulthood (Rother and

Accili, 2000). Like IGF1, *IGF2* is also found associated with binding proteins (IGFBP-1 to -6). These binding proteins play several roles, including extending the half-life of IGF2 protein and translation of *IGF2* mRNA (Daughaday et al., 1989; Firth and Baxter, 2002).

The receptor for IGF2 (IGF2R) has a different structure to the IGF1R. It is a type I transmembrane glycoprotein with a large extracellular domain consisting of three ligand-binding regions: one for IGF2 and two for proteins containing mannose-6-phosphate (M6P) and the dormant form of transforming growth factor beta (TGF-β). Although IGF2R exhibits structural and biochemical differences with the IGF1R and IR, IGF2 can bind to all three receptors. In addition, the extracellular domain of the IGF2R dissociates from the cell membrane as a soluble fragment, circulating in the blood with the ability to bind to IGF2 and facilitates its degradation. The IGF2R does not transduce an intracellular signal, but rather acts to reduce the bioactivity of IGF2 by sequestering it away from the IGF1R. Consequently, IGF2R provides extra control for the circulating levels of IGFs, functioning as a "sink" that controls the local bioavailability of IGF ligands for binding to the IGF1R. Like GH, aberrant IGF system signalling is also associated with diseases (Bergman et al., 2013; Brouwer-Visser and Huang, 2015).

1.4.3.2 Prolactin, a GH-related hormone

Prolactin (PRL) is a lactogenic GH-related hormone that is expressed in the pituitary, placenta, and endothelial cells, and is involved in a wide range of physiological roles. It plays a significant role in reproduction; it promotes neurogenesis and regulates adipose tissue metabolism and alveolar development of mammary glands during lactation (Carré and Binart, 2014; Crowley, 2015; Grattan, 2015; Marano and Ben-Jonathan, 2014; Trott et al., 2012). PRL also plays a major homeostatic role in the body, being involved in lipid metabolism, immune system regulation, maintenance of osmotic balance, and angiogenesis (Ben-Jonathan et al., 2006; Bole-Feysot et al., 1998; Carré and Binart, 2014; Goffin et al., 1998; Marano and Ben-Jonathan, 2014; Yang and Friedl, 2015). PRL has been found to have pro-angiogenic effects; however, a 17 kDa proteolytic fragment of PRL has been demonstrated to possess anti-angiogenic activity (Ge et al., 2007; Yang and Friedl, 2015). Aberrant PRL expression and signalling have been implicated in autoimmune diseases, breast and prostate cancer, and metabolic disorders, such as obesity (Ben-Jonathan et al., 2006, 2008; Brandebourg et al., 2007; Clevenger et al., 1997, 2003; Goffin et al., 2011; Harvey et al., 2008; Montgomery, 2001)

The *PRL* gene has been mapped to chromosome 6 in 6p22.2-p21.3. From an evolutionary point of view, the prolactin gene (*PRL*) shares 24% amino acid sequence identity and 50% amino acid similarity with the *GH1* gene, and 24% amino acid identity and 47% amino acid similarity with the *CSH1* gene (Goffin et al., 1996; Mertani et al., 1998).

PRL binds to another type I cytokine receptor that has a similar structure to the GHR: the PRL receptor (PRLR). Structurally, the 5-prime-untranslated region of the PRLR gene contains 2 alternative first exons (found within 800 bp of each other), a noncoding exon 2, and exons 3-10 encoding the *PRLR* gene product, with exon 10 encoding most of the intracellular domain (Hu et al., 1999). The *PRLR* gene maps to chromosome 5p13.2, reasonably close to the *GHR* gene (Gross, 2013). These two genes are thought to have evolved from a common precursor and are both homologous to receptors for members of the cytokine superfamily. This is supported by the fact that GH binds to the prolactin receptor, which forms the basis of the induction of lactation by GH in humans, and that the PRLR can form hetero-multimers with GHR (Frank et al, 2016).

1.5 GH/IGF1 axis and disease

Due to effects of GH/IGF1 axis in multiple facets of growth and metabolism, deficiency or excess of these hormones, as well as compromised signalling pathways involving this axis, have been implicated in multiple growth disorders. Deficiency in GH during childhood leads to short stature. Decreased GH can also affect adults, resulting in a condition known as adult GH deficiency. Patients suffering from this disorder have increased fat mass, decreased bone mineral density, excessive fatigue, depression, and cardiac dysfunction (Boer et al, 1995). Congenital functional GHR deficiency due to inactivating mutations in the GHR causes Laron syndrome, and results in decreased IGF1 levels and short stature (Berg et al., 1993; Laron & Kauli, 2016; Steuerman et al, 2011). Patients affected by this syndrome have reduced incidence of cancer and diabetes (Guevara-Aguirre et al., 2011; Steuerman et al., 2011).

An excess of GH in the body can also lead to disorders, both in children and in adults. In children, increased GH has been associated with abnormally tall stature, known as gigantism. This condition pre-disposes the patients to a variety of serious diseases, such as cancer and decreased life span. In adults, GH excess (and related conditions) is termed acromegaly. This is usually caused by hypersecretion of GH from a pituitary adenoma, and is characterised by

overgrowth of face skeletal tissues and extremities, headache, bone deformities, and hypertension. This disorder manifests slowly but is potentially life threatening (Kannan and Kennedy, 2013; Melmed, 2006).

Since the functions of GH/IGF axis extends beyond just normal growth, this axis has been linked to other serious disorders. This axis has been linked to multiple cancer types, such as breast, endometrial, prostrate, and colorectal cancer. There are numerous experimental studies in animal and human cell models that back this observation (Chhabra et al., 2011; Perry et al., 2017). Overexpression of genes related to this axis have been found to decrease response to chemotherapeutic drugs and radiation treatment (Bougen et al., 2012; Clayton et al, 2011; Perry et al, 2017; Weroha & Haluska, 2012). Since GH plays a major role in cell proliferation and differentiation, loss of GH transcriptional control could potentially lead to tumorigenesis and malignancy by increased cellular proliferation and survival. Studies demonstrating inhibition and reduction of oncogenicity in cell-lines and tumour models upon treatment with pegvisomant, an antagonist of the GHR, further implicate GH and related genes in cancer (Divisova et al., 2006; Evans et al., 2016; Lu et al., 2019).

1.5.1 Nuclear translocation of classical cell-surface receptors

Advances in cellular and molecular techniques have revitalised interest in the unconventional behaviour of cell-surface receptors including that of nuclear translocation. Throughout the 1980s-1990s, an increasing number of reports identified the presence of various classic cell-surface receptors in the nucleus of both normal and malignant cells (Lobie et al., 1994a; Maher, 1996; Podlecki et al., 1987). This phenomenon was proposed to be associated with diseases, such as cancer (Conway-Campbell et al., 2007; Pollak, 2012; Simpson et al., 2017; Wang and Hung, 2009). However, at this time it was difficult to progress these observations, given the limitations of available methodologies. Advances in molecular biology techniques, such as chromatin immunoprecipitation (ChIP), and new research approaches that allow for increased insight into the complexities of cellular signalling have enabled us to revisit this process of aberrant nuclear localisation. As many of the nuclear translocation studies involve the imaging of fixed cells, one ought to proceed with caution in interpretation of this data considering the artefacts presented by protein subcellular localisation imaging techniques.

1.5.1.1 The growth hormone receptor translocates to the nucleus

The GHR potentially functions as a "moonlighting protein," which means that it serves a dual purpose. Traditionally, GHR functions by binding of its ligand, GH, at the cellular surface, leading to stimulation of a cascade of signalling pathways that drive gene transcription. In the past few decades, however, an alternative mechanism for GH/GHR activity has been reported, which may mediate some of the functions of GH, not necessarily the ones modulated by the cell-surface interaction. Numerous studies have reported the nuclear translocation of GHR upon stimulation with GH, both *in vivo* and *in vitro* (Conway-Campbell et al., 2007; Lincoln et al., 1998; Lobie et al., 1992, 1994a; Mertani et al., 1998). This has also been reported in porcine hepatocytes and transgenic zebrafish (Figueiredo et al., 2016; Lan et al., 2017). It has also been reported that GHR nuclear translocation is mediated by the importin (IMPα/β) pathway (Conway-Campbell et al., 2007, 2008; Figueiredo et al., 2016; Lan et al., 2017).

Coerced localisation of GHR into the nucleus of Ba/F3 proB cells by insertion of a nuclear targeting sequence (NLS) dramatically increased its sensitivity to GH (Conway-Campbell et al., 2007). This sensitivity caused the cells to proliferate autonomously. The full functional consequences of nuclear import of GHR remain to be determined, but there is a correlation between nuclear GHR and increased proliferation leading to oncogenic progression. This could be because GH is proliferative in its effects, and nuclear GHR exhibits an increased sensitivity to GH. It has been recently demonstrated that pegvisomant can block nuclear import of GHR (Lan et al., 2019). Several studies have highlighted the significance of pegvisomant in anti-cancer therapy (Divisova et al., 2006; Evans et al., 2016; Thankamony et al., 2009). These observations strongly implicate GHR nuclear localisation in tumorigenesis.

Conway-Campbell *et al* discovered the presence of a functional transactivation domain in GHR, which implies that nuclear GHR may function as a transcription factor, either directly or indirectly (Conway-Campbell et al., 2008). They also showed a GH- dependent interaction between the GHR extracellular domain and a protein called coactivator activator (CoAA). CoAA is a potent nuclear receptor coactivator protein which is found to be overexpressed in a wide range of cancers (Kai, 2016; Sui et al., 2007). This study further corroborates the significance of nuclear translocation of GHR and its potential implication in cancer.

1.5.1.2 Nuclear translocation of other cell-surface receptors

Research over the past few decades has identified a new paradigm in the field of signal transduction, which is the nuclear translocation of cell-surface receptors. Potential destinations for internalised cell-surface receptors include the plasma membrane for receptor recycling, lysosomes for degradation, and the nucleus. Approximately 18 tyrosine kinase receptors (a class of cell-surface receptors) have been reported to translocate into the nucleus, including the epidermal growth factor receptor (EGFR), fibroblast growth factor receptor 1 (FGFR1), vascular endothelial growth factor receptor 1 (VEGFR1), platelet derived growth factor receptor beta (PGDFR-β), IGF1R, and PRLR (Carpenter and Liao, 2013; Maher, 1996; Papadopoulos et al., 2018; Shay-Salit et al., 2002).

One of the well-studied examples of this phenomenon is the EGFR. Nuclear localisation of EGFR has been shown in multiple developmental and malignant cell-types and tissues (Brand et al., 2011; Li et al., 2008; Lin et al., 2001; Marti et al., 2001; Psyrri et al., 2005; Wang and Hung, 2009). EGFR has been reported to directly interact with DNA and function as a transcription factor (Brand et al., 2011; Lin et al., 2001; Rakowicz-Szulczynska et al., 1986). Studies have shown that EGFR interacts with transcription factors, such as STAT5, STAT3, and E2F1, and regulates gene expression (Hung et al., 2008; Lo et al., 2005; Wang et al., 2006). Nuclear EGFR also plays a role in regulation of DNA repair machinery, often in response to damaging stimulus like radiation or cisplatin (Chou et al., 2014; Liccardi et al., 2011). Upregulated nuclear EGFR levels are found in multiple cancer-types (Li et al., 2009; Traynor et al., 2013; Xia et al., 2009).

Numerous studies have established the presence of IGF1R in the nucleus of normal and cancer cells upon treatment with IGF1, and also its potential role in regulating gene expression (Aleksic et al., 2010, 2018; Poreba and Durzynska, 2020; Sarfstein et al., 2012). Sarfstein *et al.* showed that nuclear IGF1R acts as a transcription factor by auto-regulating *IGF1R* gene transcription in breast cancer cells through an estrogen receptor-mediated pathway (Sarfstein et al., 2012). Additionally, Aleksic *et al* reported the presence of IGF1R binding sites in the vicinity of the transcription start sites of oncogenic genes, such as *JUN* and *FAM21*, in prostate cancer cells. Additionally, they showed direct binding of nuclear IGF1R to DNA and interaction with RNA polymerase II, which promotes tumorigenesis by upregulating the expression of *JUN* and *FAM21* (Aleksic et al., 2018). IGF1R has also been shown to function as a transcriptional activator of cancer-related

transcription factors, T cell factor/lymphoid enhancer factor (TCF/LEF), resulting in increased expression of *cyclin D1* and *axin2* genes (Warsito et al., 2012). Several mechanisms of IGF1R nuclear translocation are being investigated currently. These mechanisms include, firstly, the insulin receptor substrate-1 (IRS1)-mediated pathway, and, secondly, the incorporation of small ubiquitin-like modifiers (SUMOylation). IGF1R does not possess a nuclear localisation sequence (NLS), and it does not interact with β-importin, which rules out importin-mediated pathways. IRS-1, however, contains an NLS and it had been shown that IGF1R is responsible for activation of IRS1 nuclear translocation. It is possible that IRS1 plays a role in nuclear import of IGF1R. SUMOylation is considered to be crucial for IGF1R nuclear translocation. Mutation in SUMO-1 binding sites prevented the nuclear translocation of IGF1R and the consequent transcriptional activity. SUMOylation of IGF1R has been implicated in increased cellular proliferation and cancer.

The GHR-related receptor, PRLR, translocates into the nucleus upon treatment with PRL (Rao et al., 1995). It has also been reported that nuclear PRLR functions as a transcriptional protein, causing an effect on gene expression. The PRLR has, further, been found to localise with chromatin directly, and to interact with transcription factors in the nucleus, such as STAT5A and HMGN2 (Fiorillo et al., 2011). Since the PRLR can form hetero-multimers with GHR (Liu et al., 2016a), it would be interesting to determine whether GHR/PRLR hetero-multimers can translocate into the nucleus.

1.5.1.3 Biological/clinical implications of nuclear translocation of cell-surface receptors and future directions

There are significant clinical implications of understanding the consequences of nuclear import of receptors in different cell models. The consequences of nuclear import of receptors include their potential transcription factor activity, promotion of tumorigenesis, and resistance to radiotherapy. As described in previous sections, evidence suggests that increased levels of nuclear cell-surface receptors, such as GHR, EGFR, PRLR, and IGF1R, are associated with cancer progression, as well as poor prognosis following cancer treatment. Since there is a strong correlation between receptors in nucleus and increased carcinogenesis, prevention of receptor nuclear import may be a potential therapeutic adjunct to already

existing treatments. This could be studied by blocking of nuclear localisation signals and/or rendering receptor binding sites of nuclear import proteins unavailable, so as to prevent the entry of the receptor into the nucleus. The latest bioinformatics methods can aid greatly in exploring the mechanism behind this nuclear receptor import, and in the putative target discovery of drugs. This could add significantly to our understanding of complex interactions between cellular signalling cascade and regulation of gene expression, which can, in turn, be useful for designing therapeutic measures for diseases, such as cancer.

1.6 Conclusions and thesis direction

There is now mounting evidence that spatial organisation within the nucleus is key to genome function. Genetic variation can impact on 3D genome structure and, consequently, on development and disease. Chapter 3 identified SNPs associated with the *GH* locus genes that can potentially alter regulation of gene expression, both proximally (*cis*) and distally (trans), which, in turn, can affect GH function and contribute to disease.

The GH receptor is a cell surface receptor that has recently been shown to translocate to the nucleus and to localise with chromatin, but the function of nuclear localisation is unclear. Nuclear localisation of other cell-surface receptors, such as the EGF receptor, has shown that these receptors can function either directly as a transcription factor, or can interact with the transcription factors/complex and drive changes in gene expression. GHR nuclear translocation has been linked to increased tumorigenesis, which suggests that GHR potentially interacts with transcription factors and/or other proteins upon migration into the nucleus. In Chapter 4, an integrated approach was taken to determine if the GHR interacts with proteins in the nucleus (IP-mass spectrometry), and explore its subsequent effects on gene expression (Clariom D microarray) in an endometrial cancer cell-line, RL95-2.

Regulation and balance of transcription of both coding and non-coding genes is essential for mediation of cellular processes. Dysregulation of non-coding genome expression can upset this balance and can lead to disease. It has been reported that hormones can stimulate changes in transcription and binding of non-coding RNA. Chapter 5 is based on an exploration of the time-dependent effects of GH treatment on the regulatory mechanisms of coding and non-coding RNA that are associated with modulation of GH functions.

1.7 Aims

Three specific aims were investigated in this thesis.

1. To determine whether the *GH* locus and locus control region potentially functions as a complex regulatory hub. In this aim, SNPs located within the *GH* locus were identified and assessed for association with changes in expression of genes in *cis* and in *trans*.
2. To determine the genomic and proteomic consequences of nuclear translocation of the GHR. This study incorporated immunoprecipitation-mass spectrometry to identify proteins which the GHR interacts with in the nucleus and expression analysis to identify GH-induced changes in gene expression.
3. To assess GH-dependent changes in coding and non-coding mRNA transcription profiles and to construct a regulatory network of mRNA, miRNA and lncRNA.

CHAPTER 2

Materials and Methods

2.1 Materials

Table 2.1. . List of chemicals, reagents and suppliers

Material	Source
Acrylamide/Bis solution (40%)	Bio-Rad laboratories, Inc., Hercules, CA, USA
Agarose (Ultra-pure)	Life Technologies, Carlsbad, CA, USA
Ammonium persulfate (APS)	Serva Electrophoresis GmbH, Heidelberg, Germany
Bromophenol Blue	Sigma Chemical Company, St Louis, MO, USA
Bovine serum albumin (BSA)	Immuno Chemical Products Ltd, Auckland, New Zealand
Clarity Western ECL Substrate (Cat No. 1705060)	Bio-Rad laboratories, Inc., Hercules, CA, USA
Halt Protease Inhibitor Cocktail (100X)	Life Technologies, Carlsbad, CA, USA
Dimethyl-sulphoxide (DMSO)	Sigma Chemical Company, St Louis, MO, USA
Ethylenediaminetetraacetic acid (EDTA)	Sigma Chemical Company, St Louis, MO, USA
Ethanol (absolute, analytical grade)	Scientific Supplies Ltd, Auckland, New Zealand
Fetal bovine serum (FBS)	Gibco New Zealand Ltd., Auckland, New Zealand
Glycine	Applichem GmbH, Darmstadt, Germany
Horse serum	Gibco New Zealand Ltd., Auckland, New Zealand
Human recombinant pituitary growth hormone	Harbor-UCLA Medical Centre, Torrance CA, USA
MEGM medium	Lonza, Basel, Switzerland
Methanol	Scientific Supplies Ltd, Auckland, New Zealand

Mercaptoethanol	Sigma Chemical Company, St Louis, MO, USA
Nitrocellulose membrane	Bio-Rad laboratories, Inc., Hercules, CA, USA
Paraformaldehyde	Sigma Chemical Company, St Louis, MO, USA
Pierce BCA Protein Assay Kit (Cat No. 23225)	Invitrogen Life Technologies, Carlsbad, CA, USA
Resazurin sodium salt	Sigma Chemical Company, St Louis, MO, USA
RPMI 1640 medium	Invitrogen Life Technologies, Carlsbad, CA, USA
Seeblue plus2 protein marker	Invitrogen Life Technologies, Carlsbad, CA, USA
Sodium dodecyl sulfate (SDS)	Invitrogen Life Technologies, Carlsbad, CA, USA
SureBeads Protein G Magnetic Beads	Bio-Rad laboratories, Inc., Hercules, CA, USA
Tetramethyl-ethylendiamin (TEMED)	Sigma Chemical Company, St Louis, MO, USA
Tris	Serva Electrophoresis GmbH, Heidelberg, Germany
Triton X-100	Sigma Chemical Company, St Louis, MO, USA
Trizol	Life Technologies, Carlsbad, CA, USA
Trypan Blue	Gibco New Zealand Ltd., Auckland, New Zealand
Trypsin	Gibco New Zealand Ltd., Auckland, New Zealand
Tween 20	Serva Electrophoresis GmbH, Heidelberg, Germany

Table 2.2. List of antibodies and the suppliers

Name	Catalogue Number	Source
Anti-GAPDH antibody	Ab36840	Abcam, Cambridge, UK
Anti-β-ACTIN antibody	A1978	Sigma-Aldrich, MO, USA
Anti-EGFR antibody	2232-S	Cell Signalling Technology, Danvers, MA, USA
Anti-GHR antibody (extracellular domain)	Ab89400	Abcam, Cambridge, UK
Anti-GHR antibody (intracellular domain)	Sc-137185	Santa Cruz Biotechnology, CA, USA
Anti-HMGN1 antibody	720387	Life Technologies, Carlsbad, CA, USA
Anti-IGF1R antibody	3027	Cell Signalling Technology, Danvers, MA, USA
Anti-mouse IgG, HRP-linked antibody	A9044	Sigma-Aldrich, MO, USA
Anti-rabbit IgG, HRP-linked antibody	7074	Cell Signalling Technology, Danvers, MA, USA
Anti-SUMO1 antibody	332400	Life Technologies, Carlsbad, CA, USA
Rabbit anti-phospho-STAT5	71-6900	Invitrogen Life Technologies, Carlsbad, CA, USA
Stat5 (C-17)-G	sc-835-G	Santa Cruz Biotechnology, CA, USA

Table 2.3. List of buffers and solutions

Buffers and solutions	Composition
CELL CULTURE	
Cell culture medium (RPMI-1640)	475 mL RPMI medium, 25 mL heat inactivated FBS, 5 mL glutamine, 5mL penicillin/streptomycin
Cell culture medium (MEGM)	475 mL MEGM medium, 25 mL heat inactivated horse serum, 5 mL glutamine, 5mL penicillin/streptomycin

10x PBS	80 g NaCl, 2 g KCl, 14.4 g Na_2HPO4, 2.4 g KH_2PO4, 1 L Milli-Q water. Adjust pH to 7.4
10x trypsin/EDTA solution (0.25%)	2.5 g Trypsin, 0.372g EDTA, 0.35 g $NaHCO_3$, 1 L autoclaved H_2O. Adjust pH to 7.2.
Freezing medium	1 mL DMSO, 4 mL heat inactivated FBS, 5 mL RPMI medium
20x resazurin dye solution	0.1 g resazurin sodium salt in 100 mL PBS, and filtered through a 0.22 um filter, and then were stored at 4 ºC.
Protein extraction	
Cell lysis buffer	2 mL 10% SDS, 2 mL glycerol, 1.2 mL 0.5M Tris-HCl pH 6.8, 1 tablet complete minitab protease inhibitor, 4.8 mL Milli-Q water
	SDS-PAGE
4% Stacking gel	500 µl 40% acrylamide, 1.26 mL 0.5 M Tris-HCl pH 6.8, 50 µl 10% SDS, 3.18 mL Milli-Q water, 5 µl TEMED, 25 µl 10% APS
12% Separating gel	3 mL 40% acrylamide, 2.5mL 1.5M Tris-HCl pH8.8, 100 ul 10% SDS, 4.35 mL Milli-Q water, 5ul TEMED, 50ul 10% APS
6x SDS loading dye	6 mL glycerol, 3 mL 1 M Tris-HCl, pH 6.8, 1.2 g SDS and 5 mg bromophenol blue
1x SDS running buffer	3.03 g Tris, 14.41 g glycine, 10mL 10% SDS, 1L Milli-Q water
1x Transfer buffer	3.03 g Tris, 14.41 g glycine, 200 mL Methanol, 1 mL 10% SDS, 800 mL Milli-Q water
Western blot	
0.1% PBS-Tween	100 mL 10x PBS, 900mL Milli-Q water, 1 mL Tween-20
Blocking buffer	5 g non-fat dry milk powder, 100 mL 0.1% PBS-Tween
	IMMUNOFLUORESCENCE
Fixing medium	4 g Paraformaldehyde in PBS (w/v)
Permeabilisation buffer	0.5 mL Triton X-100, 100 mL 0.1% PBS
Blocking buffer	1 g BSA, 100 mL 0.1% PBS
	RNA EXTRACTION
RW1 buffer	Washing buffer containing a guanidine salt
RPE buffer	RNA precipitation buffer

2.2 Methods

2.2.1 Cell culture

2.2.1.1 Mammalian cell-lines

The human mammary epithelial cell-line, MCF-10A, and human endometrial carcinoma cell-line, RL952, were obtained from the American Type Culture Collection (ATCC) (Manassas, VA, USA).

2.2.1.2 Passaging and harvesting of cell-lines

MCF-10A cells were maintained in MEGM™ Mammary Epithelial Cell Growth Medium BulletKit™ (Lonza,) supplemented with 100 µg/ml streptomycin (Sigma-Aldrich), 100 U/ml penicillin (Sigma-Aldrich), and 5% horse serum (ThermoFisher Scientific). RL95-2 cells were maintained in in RPMI-1640 growth medium (Gibco), supplemented with 5% fetal bovine serum (Thermo Scientific), 100 µg/ml streptomycin (Sigma-Aldrich), 100 U/ml penicillin (Sigma-Aldrich), and Glutamax. Both cell-lines were grown at 37°C in a humidified 5% CO2 incubator to a confluence of 70-80%, before passaging into new cell culture vessels to continue the stock cultures.

For passaging cells, the media was aspirated out, followed by washing of cells with PBS and addition of 2-3 mL of Trypsin/EDTA for detachment of cells. The cells were incubated at 37°C for 3-5 min and then supplemented with serum-containing media (15 mL) to neutralise the trypsin and pipetted well. The cell suspensions were centrifuged at 1000 rpm for 5 min to pellet the cells. The supernatant was discarded, pelleted cells were re-suspended in fresh media and transferred to the incubator for maintenance and growth.

2.2.1.3 Cell counting

Cell concentration was measured using a haemocytometer. 10 µl of the cell suspension was transferred to an Eppendorf tube and mixed with 90 µl of serum-free culture medium. 10 µl each of this diluted solution was transferred onto both sides of the haemocytometer and counted. The number of cells /ml were calculated using the formula below.

Cells/ml= (Number of cells in 4 quadrants/4) × dilution factor × 10000

2.2.1.4 Storage of cell-lines

Counted cells were centrifuged (1000 rpm, 5 min) and re-suspended in the freezing medium (40% FBS, 10% DMSO in serum-free medium without antibiotics). 1 ml aliquots of the cell suspension were gently transferred into cryogenic vials (Nalgene, Rochester, NY, USA) and into an appropriate freezing chamber. The chamber was placed into a -80°C freezer for 24 h to allow gradual cooling and freezing to prevent cell damage. After 24 h, frozen cells were transferred to liquid nitrogen for long-term storage.

2.2.1.5 Revival of cell-lines from liquid nitrogen storage

Cell aliquots stored in liquid nitrogen were thawed rapidly in 5 mL of serum-supplemented culture medium and centrifuged (1000 rpm, 5 min). The supernatant was discarded and the pelleted cells were re-suspended in fresh culture medium, and transferred into a 75 cm^2 tissue culture flask (Greiner Bio-One, Germany). The cells were then cultured at 37°C in a humidified 5% CO_2 incubator.

2.2.2 RNA extraction

RL95-2 or MCF-10A cells were plated in 10cm cell culture dishes at the density of 5 x 10^6 cells per dish. Cells were then serum-starved overnight before treating them with 250 ng/mL or 500 ng/mL GH for indicated times. Pelleted cells were stored in 1 mL of Trizol® reagent (#15596-026, Life Technologies, Carlsbad, CA). RNA was extracted using a protocol that integrated the use of Trizol® reagent with spin columns (RNeasy® Mini Kit, #74104, Qiagen, Hilden, Germany). 0.2 mL of chloroform was added to every 1 mL of Trizol solution used and it was shaken vigorously for 15 sec and incubated at room temperature (RT) for 5 min. Samples were then centrifuged for 15 min at 12,000g at 4°C.

The aqueous phase was transferred to another tube and was mixed with an equal volume of 100% ethanol. 700 µL of this mixture was added to an RNeasy® spin column and centrifuged (12,000g, 30 sec, RT). The flow-through was discarded and the process was repeated to process the entire sample volume. Wash buffer RW I (700 µL) was added to the column and centrifuged (12,000g, 30s, RT), flow-through was discarded and the spin column was transferred to a new collection tube. Buffer RPE (500µL) was added to each spin column and centrifuged (12,000g, 30s, RT), discarding the flow-through. This step was repeated but the column was centrifuges for 2 min. The spin column was then placed in an RNase free tube. Nuclease free water (15 µL – 30 µL) was added directly onto the membrane, incubated for 1

min at RT and the RNA was eluted by centrifugation (12,000g, 2 min, RT). Eluted RNA was quantified, quality checked (Bioanalyser and Nanodrop™) and stored at -80°C until required for further analysis.

2.2.2.1 RNA quality evaluation and quantification

Extracted RNA was quantified using 2 µL of each sample in a micro-volume spectrophotometer (Nanodrop ND-1000, Thermo Scientific, and Wilmington, DE). The Nanodrop measures absorbance at 260, 280 and 230 nm. The ratio of absorbances at these wavelengths is used to determine the purity of nucleic acids. A 260/280 ratio of ~2.0 is generally considered to be pure for RNA. Absorbance at 230 nm indicates presence of other contaminants such as phenol. A 260/230 ratio of 2.0-2.2 is considered to be good quality RNA.

The integrity of the extracted RNA was determined using an Agilent bioanalyser (Model 2100, Agilent Technologies, Santa Clara, CA, USA) and an RNA 6000 Nano LabChip kit according to the Manufacturer's instructions. RNA integrity number (RIN) values range from 1 to 10, with 1 being the most degraded and 10 being the least degraded RNA. Only RNA samples with a with a RIN number >9 were used.

2.2.3 Protein extraction

Cells were plated in T75 cell culture flasks or 10 cm dishes with a cell density of 5 x 10^6 cells per flask/dish, serum-starved overnight then treated according to the different experiments. Following treatment, the cells were placed on ice and ice-cold lysis buffer (50 mM Tris-HCL (pH 7.4),1% Nonidet P-40; 150 mM NaCl, 1 mM EDTA, 1 mM NaF, 1 mM phenylmethylsulfonylfluoride, 1 mM Na3VO4, and protease inhibitor cocktail) was added to each flask/dish. The cells were collected using a cell scraper and transferred to a fresh tube, followed by centrifugation (20,000g, 30 min, and 4°C). The pellet was discarded and supernatant containing the proteins was transferred to a new tube and stored at -80°C until required for further analysis.

To extract proteins present in different subcellular fractions (cytoplasmic and nuclear), the NE-PER fractionation kit (ThermoFisher Scientific) was used according to the manufacturer's instructions. After treatment (described above), cells were harvested by centrifugation at 500g for 5min at 4°C. Cells were washed with PBS, following which ice-cold CER I reagent from the kit was added to the pellet (200 µL), the tube was vortexed for

15 sec and incubated on ice for 10 min. Ice-cold CER II (11 µL) was then added and the tube was vortexed for 5 sec, followed by incubation on ice for 1 min. The tube was then centrifuged at maximum speed for 5 min and the supernatant was transferred to a fresh, pre-chilled tube. This supernatant is the cytoplasmic extract. The pellet was then re-suspended in ice-cold NER (100 µL) and vortexed for 15 sec every 10 min on ice, for a total of 40 min. The tube was then centrifuged at maximum speed for 10 min and the supernatant (nuclear extract) was transferred to a clean, pre-chilled tube. The extracts were stored at -80°C.

2.2.3.1 Protein quantification

Extracted proteins were quantified using the BCA Protein Assay Kit (Pierce, Thermo Scientific, IL, USA). The standards were prepared according to the manufacturer's instructions using the lysis buffer as the diluent. Working reagent was prepared by combining BCA Reagent A with BCA Reagent B in the ratio of 50:1. In a 96-well plate, 10 µL of the protein standards and samples were added in triplicates followed by the addition of 200 µL of the working reagent into each well. The plate was covered in tin foil to prevent exposure to light and agitated gently for 3 min to mix the reagents. The plate was then incubated in the dark for 30 min at 37ºC. Readings were recorded using a Biosynergy2 plate reader at the absorbance wavelength of 562 nm. Measurements for the protein standards were used to plot a standard curve to calibrate the readings. Linearity of the curve (ideal range = R^2 >0.99) are a measure of the accuracy of the test. This curve was used for calculation of the total protein concentration in the lysates.

2.2.3.2 SDS-PAGE and western blotting

Stacking (4%) and resolving (12%) gels were cast according to the recipes described in Table 2.3 of this Chapter. The electrophoresis tank was assembled, 1× SDS-PAGE running buffer was added to the tanks, and the combs from the wells were removed. 6× loading dye (with a reducing agent) was added to the lysates and incubated at 100ºC for 10 min. Lysates were then loaded into the wells. Electrophoresis was carried out at a constant voltage of 120 V for ~ 1-1.5 hours.

Proteins resolved by SDS-PAGE were then transferred onto nitrocellulose membranes at a constant voltage of 100V for 1.5 hours. Membranes were blocked with 5% BSA for 1-2 hours at RT. After washing with PBS-T (PBS + Tween 20) solution three times for 10 min, membranes were incubated with the primary antibodies overnight at 4°C. Membranes were then washed three times with PBS-T for 10 min, and incubated with a horseradish

peroxidase–conjugated secondary antibody for 1 hour at RT. Membranes were then washed three times for 10 min to reduce background and non-specific binding. Subsequently, the proteins were visualised using Clarity Western Peroxide Reagent (Bio-Rad) and a Bio-Rad Chemidoc MP system.

2.2.3.3 Immunoprecipitation

Cell lysates were pre-cleared with magnetic protein-G conjugated beads to reduce non-specific binding. Pre-cleared lysates (1 mg/ml) were incubated with 10 µg of indicated antibody or the isotype control (IgG) antibody overnight at 4°C on a rotating wheel to form protein-antibody complexes. 100 µL of protein-G conjugated beads were washed using 1X PBS solution and then incubated with the protein-antibody complexes at 4°C for 2-4 hours on a rotating wheel. The beads were separated from the solution using a magnetic rack and the solution was discarded. Beads were then washed 5-8 times with 1X PBS solution to reduce non-specific binding complexes.

For use in western blotting, the beads were re-suspended in 3× loading dye and incubated at 95°C for 10 min. The beads were separated using a magnetic rack and the solution is transferred to a fresh tube. This solution contains the immunoprecipitated proteins. The proteins were then detected by western blotting as described above under reducing conditions, or were used for mass spectrometry analysis.

A total of 1mg/ml of protein lysate was used per replicate for mass spectrometry. Beads were washed twice with freshly prepared 100 mM ammonium hydrogen carbonate (AMBIC) solution. Following the second wash step, beads were transferred to a fresh tube and snap-frozen. These tubes were then stored at -80°C and subsequently sent for mass spectrometric analysis.

2.2.4 Immunofluorescence

Cells (2×10^4) were plated on coverslips in 6-well plates and serum-starved overnight. Following stimulation with GH at the indicated times and concentrations, cells were fixed with 4% paraformaldehyde at 37°C for 10 min. Fixed cells were washed with PBS three times and permeabilised with 1% Triton X-100 for 30 min. After washing, the coverslips were blocked with 5% BSA in PBS for 2 hours at room temperature, washed with PBS and incubated with primary or mouse IgG control antibodies overnight at 4°C, followed by secondary antibodies labelled with a fluorophore for 1 hour at 37°C. Coverslips were washed,

stained with DAPI and mounted. Cells were visualised using confocal laser scanning microscopy (Zeiss LSM 800 Airyscan confocal microscope) with ×63 oil immersion objectives. Image analysis was performed using ZEN Blue and ImageJ software.

2.3 Microarray data analysis

R version 3.6.2 and RStudio version 1.2.5019 were used for all R scripts involved in data processing and statistical analysis. For Chapters 4 and 5, I used maEndToEnd (Baszczyński and Goldstein, 1967) and limma (Ritchie et al., 2015) for microarray data analysis. Following installation of maEndToEnd package and other related packages from Bioconductor (3.8) repository in R, raw data in the form of .CEL files was imported into R. These files were then read into the system to extract the data. Following this, a raw expression set was created and subjected to quality control. After the import and quality control, this data was adjusted to eliminate background noise and probe intensities which arise due to non-specific hybridization. The next step was normalisation across arrays to accurately compare measurements across different arrays by eliminating sources of variation, such as reverse transcription efficiency, batch effects and varying laboratory conditions (Irizarry et al., 2003). This is followed by a summarisation step because each transcript is represented by multiple probes and the normalised intensities of all the probes would have to be summarised into a quantity that represents an amount proportional to the amount of RNA transcript. This data was then annotated with genomic information, such as gene symbols and identifiers. Linear models, using the package limma, were then used to identify differentially expressed genes between control and treatments, based on our experimental design. Empirical Bayes variance moderation method was applied to calculate moderated t-statistics (single time point) or F-test (time-course). This method is used to account for the fact that the number of arrays in microarrays are often small which makes it difficult to estimate variance in the generated dataset. Finally, the results containing differentially expressed genes and associated statistical information (P values, log2 fold changes and multiple testing adjustment) were extracted and used for further analysis, such as pathway enrichment.

CHAPTER 3

3D Interactions with the Growth Hormone Locus in Cellular Signalling and Cancer-Related Pathways

Chapter published in Journal of Molecular Endocrinology

Jain L, Fadason T, Schierding W, Vickers MH, O'Sullivan JM, Perry JK. 3D interactions with the growth hormone locus in cellular signalling and cancer-related pathways. *J Mol Endocrinol*. 2020; 64(4):209-222

3.1 Introduction

Growth hormone (GH) is a peptide hormone released in a pulsatile fashion from the anterior pituitary. GH is critical for mediating normal postnatal longitudinal growth in childhood and puberty, and regulating metabolism (Bonert and Melmed, 2017). Secretion of GH is positively modulated by GH releasing hormone (GHRH), ghrelin, and negatively by somatostatin. Insulin-like growth factor is a key mediator of GH actions and compromised GH/IGF1 signalling is associated with several well characterised growth disorders and is linked to an altered susceptibility to cancer, diabetes and cardiovascular disorders (Gadelha et al., 2018; Guevara-Aguirre et al., 2018; Hannon et al., 2017). In particular, altered GH expression has been linked to melanoma, breast, endometrial, liver and colorectal cancer by evidence from *in vitro*, animal, and clinical studies (Brittain et al., 2017; Chhabra et al., 2011; Perry et al., 2017).

The *GH* gene cluster is located on the long arm of chromosome 17 (17q23) and is composed of five homologous genes, *GH1* (also known as *GH-N*), *GH2* (also known as *GH-V*) and chorionic somatomammotropin genes (*CSH1*, *CSH2* and *CSHL1*) (Liao et al., 2018). Tissue-specific expression of the *GH* locus genes is regulated by a locus control region which overlaps the *CD79B* and *SCN4A* genes that are located upstream of the *GH1* gene. *GH1* is expressed primarily in the pituitary and other extra-pituitary tissues, whereas the rest of the locus genes are expressed predominantly in the syncytiotrophoblast layer of the placenta (Su et al., 1997). Studies have demonstrated that coordinated regulation of *GH* locus genes is mediated by complex chromatin looping and epigenetic mechanisms (Ganguly et al., 2015; Kimura et al., 2007; Tsai et al., 2016).

GH effects are mediated through activation of downstream signalling cascades following binding to the GH receptor (GHR), and through stimulation of secretion of secondary peptide mediator molecules, in particular, IGF1(Bonert and Melmed, 2017; Dehkhoda et al., 2018; Waters, 2016). Key signalling pathways pertaining to GH-GHR signal transduction include JAK2 signalling via STATs (1, 3 and 5), the MAPK pathway, JNK pathway, mTOR (mammalian Target of Rapamycin) and PI3K pathway (Carter-Su et al., 2016; Lu et al., 2019). These signal transduction pathways mediate GH effects by altering gene transcription

profiles, through direct stimulation of transcription and by modifying chromatin (Rotwein and Chia, 2010).

Enhancers and promoters are physically brought together to facilitate the regulation of gene expression by complex three dimensional mechanisms that are dependent on a multitude of factors (Sanyal et al., 2012; Schierding and O'Sullivan, 2015). This physical contact/interaction between enhancers and promoters can be captured by proximity ligation techniques such as Hi-C (Eijsbouts et al., 2019; Kong and Zhang, 2019; Lieberman-Aiden et al., 2009). A genomic variant associated with allele-specific changes in the expression of a gene is known as an expression quantitative trait locus (eQTL). Notably, the regulation of gene promoters can be mediated through both proximal and distal regulatory elements (*cis* and *trans* interactions, respectively), with the latter including interactions between different chromosomes (Gibcus and Dekker, 2013; Schierding et al., 2016). These interactions can associate with either higher (enhancer) or lower (insulator/silencer) expression.

As described above, coordinated regulation of *GH* locus genes is known to be mediated by complex chromatin looping (Kimura et al., 2007; Tsai et al., 2016). In light of recent studies which show regulation of gene networks by alteration of chromosomal interactions in the nucleus (Lanctôt et al., 2007), it is possible that some GH functions may be mediated by spatial interactions between regions of the *GH* gene locus and distal loci. Changes in chromatin organisation, structure and interactions play a crucial role in the regulation of gene expression (Dekker et al., 2013; Fadason et al., 2017; Schierding and O'Sullivan, 2015; Schierding et al., 2016). Therefore, the study of polymorphisms related to genes in this axis could potentially lead to elucidation of novel regulatory networks involving genes associated with GH/IGF axis function.

It was hypothesised that regulatory regions within the *GH* locus coordinate expression of a gene network that extends the impact of the *GH* locus control region. This study integrated 3-dimensional genome organisation and tissue-specific gene expression data to identify functional cis and trans spatial eQTLs that involved the *GH* locus. Regions within the *GH* locus were identified that regulate multiple genes involved in key cellular signalling and cancer-related pathways, many of which are related to GH-related signalling pathways.

3.2 Materials and Methods

3.2.1 Mapping of SNPs across the *GH* locus

Common single nucleotide polymorphisms (SNPs) (dbSNP147; Minor allele frequency ≥1%) located across the *GH* gene locus including its control region were collated (Chr17:62080000-61920000; GRCh37/hg19) (Supplementary Table 1). SNP density across the *GH* locus was calculated using a sliding window (500bp window, 100bp step size) in RStudio (Version 1.1.414).

Known enhancer sites across the *GH* locus were obtained from GeneHancer (Fishilevich et al., 2017), which collates information from the ENCODE project (Dunham et al., 2012), Ensembl regulatory build (Zerbino et al., 2018) and FANTOM5atlas of active enhancers (Lizio et al., 2015). Topologically associating domains (TADs) within the seven cell-lines *(i.e*. GM12878, HMEC, KBM7, HUVEC, IMR90, K562 and NHEK) were determined using the 3D Genome Browser at 1kb resolution and Hi-C data from Rao *et al.* (Rao et al., 2014).

3.2.2 Identification of eQTLs within the *GH* locus and their genome-wide targets

GH locus SNPs were analysed using the CoDeS3D (Contextualize Developmental SNPs using 3D Information) algorithm (GitHub, https://github.com/Genome3d/codes3d-v1) (Fadason et al., 2017). CoDeS3D integrates genome spatial connectivity data (*i.e.* maps of loci that physically interact, captured by Hi-C (Rao et al., 2014)), and links it to eQTL data obtained from the Genotype-Tissue Expression (GTEx) database (v7) (Ardlie et al., 2015)s. These results are corrected for false discovery using the Benjamini-Hochberg correction procedure (Fadason et al., 2017). The CoDeS3D database was loaded with Hi-C data for GM12878, IMR90, HMEC, NHEK, K562, HUVEC, and KBM7 human cell-lines (Rao et al., 2014). *Cis*-eQTLs were defined as involving SNPs that were located <1 Mb from the affected gene (or eGene), whereas *trans*-eQTLs were defined as involving SNPs located >1 Mb from the affected eGene on the same chromosome, or on a different chromosome (Figure 3.1A) (Fadason et al., 2017).

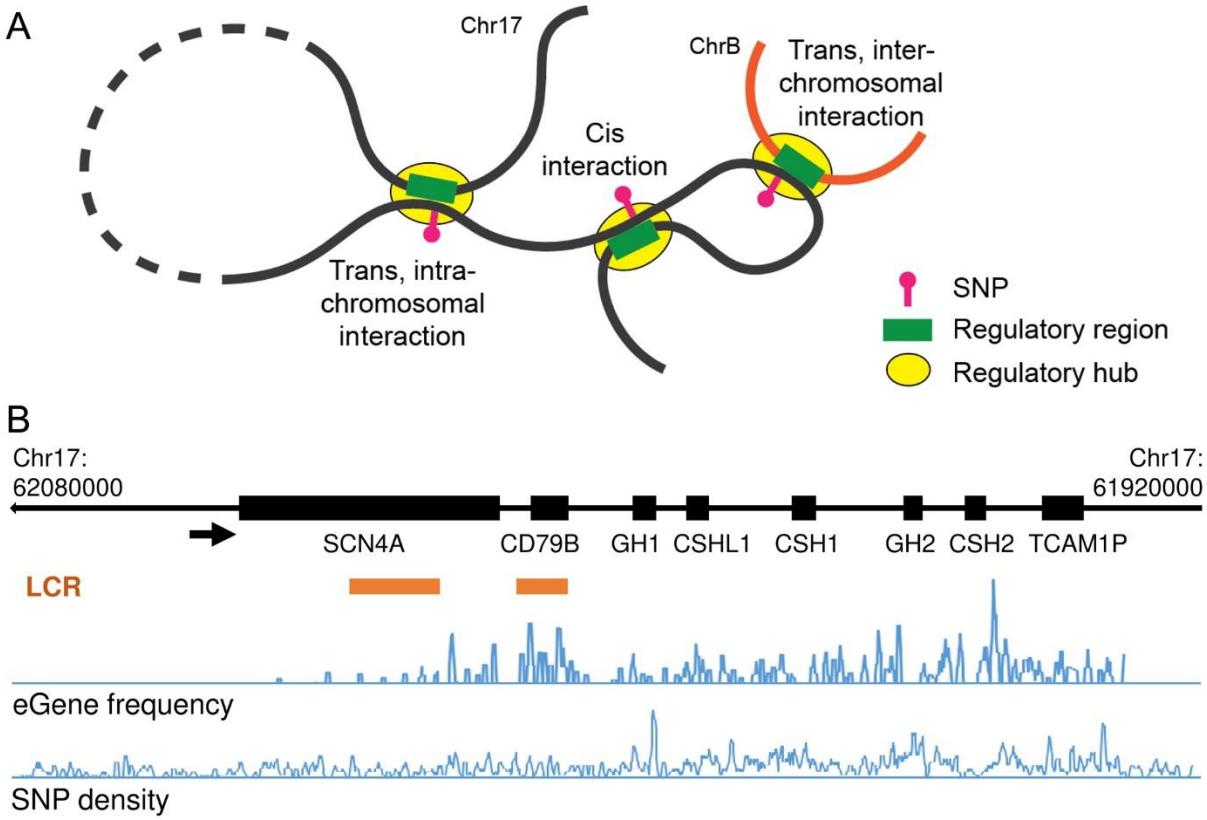

Figure 3.1. eQTL interactions across the *GH* locus
(A) Schematic diagram representing the types of eQTL interactions. Cis interactions occur when a region containing a SNP physically contacts a nearby region/gene (<1 Mb away) whereas trans interactions occur between a SNP containing region and a distal region/gene (>1 Mb apart). These regions can either be located on the same chromosome (intra-chromosomal) or on different chromosomes (inter-chromosomal). (B) Pattern of eQTL regulatory interactions for common SNPs (sourced from dbSNP147) located across the *GH* gene locus (GRCh37/hg19- Chr17:62080000-61929000). Orange boxes indicate the approximate position of the locus control region. Tracks below show the density of the 529 common SNPs that were analysed across the region and the frequency of eGenes identified as being associated with those SNPs

3.2.3 Pathway enrichment analyses

To identify enrichment of eGenes within biological pathways, the eGenes set (identified by CoDeS3D) was analysed using g:ProfileR package in R (https://biit.cs.ut.ee/gprofiler/). The reference gene sets were from KEGG (KEGG FTP Release 2019-09-30), Reactome (annotations: ensemble classes: 2019-10-2) and WikiPathways (20190910). An adjusted *P* value <0.05 was significant following Benjamini-Hochberg correction for false discovery rate (FDR) (Raudvere et al., 2019).

3.2.4 Data visualisation

Figures were drawn in RStudio (Version 1.1.414) using the following packages/libraries: Circlize (Gu et al., 2014) and ggplot2 (Gómez-Rubio, 2017).

3.3 Results

3.3.1 Common polymorphisms across the *GH* gene locus are associated with expression of multiple downstream genes

CoDeS3D was used to analyse 529 common SNPs (dbSNP 147) across the *GH* locus (Chr17:62080000-61920000; GRCh37/hg19) and identify SNP-eGene pairs in which the SNP was associated with the eGene expression level (Figure 3.1B; Supplementary Table 1) (Fadason et al., 2017).

181 SNPs that interacted with 292 genes in 48 different tissues were identified (Table 3.1, Supplementary Table 2). There were 2141 SNP-eGene associations (FDR <0.05, Benjamini-Hochberg correction) in *cis* (<1 Mb distance between the SNP and eGene) and 708 associations in *trans* (where the SNP and eGene are >1 Mb apart or on different chromosomes).

Table 3.1. Overview of connections in cis and trans identified by CoDeS3D

SNPs analysed	SNPs with eQTLs	eGenes	Cis eGenes	Trans eGenes	Cis connections	Trans connections
529	181	292	32	260	2141	708

The distribution of the eQTL SNP-eGene interaction frequency was compared with the query SNP density across the selected region (Figure 1B). Notably, the number of SNPs per 500bp sliding window did not correlate (R^2=0.1) with the number of functional eQTL-eGene interactions (Figure 3.2). Thus, the identification of regions with functional eQTL-eGene interactions was not an artefact of regions of higher SNP density across the *GH*

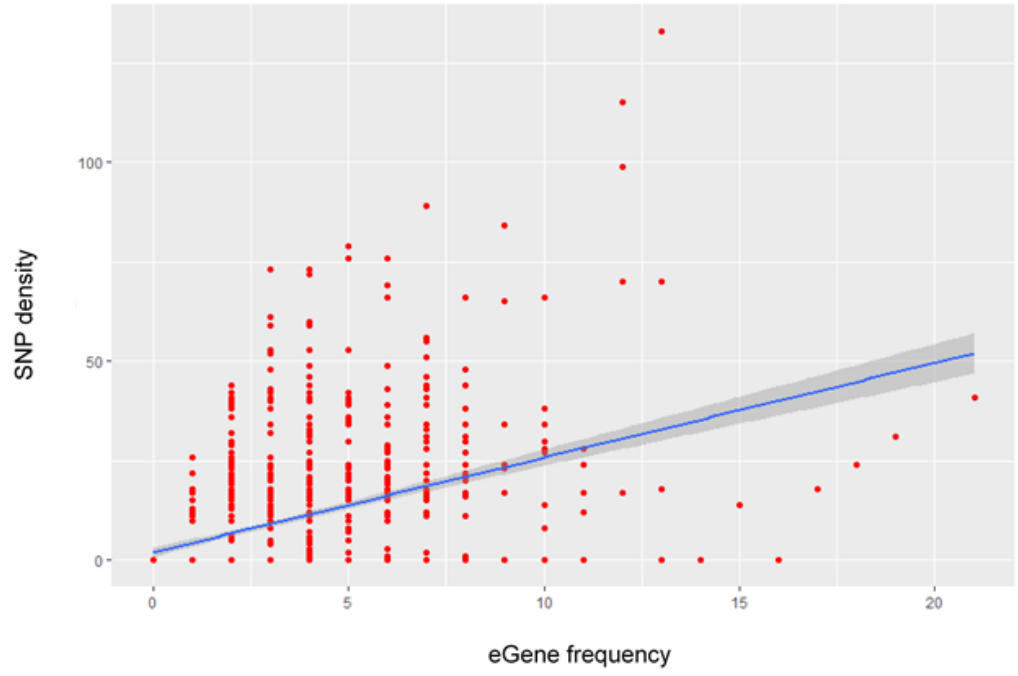

Figure 3.2. Correlation plot between SNP density across *GH* locus region and frequency of identified eGenes demonstrating that there was no correlation between them (R^2=0.1).

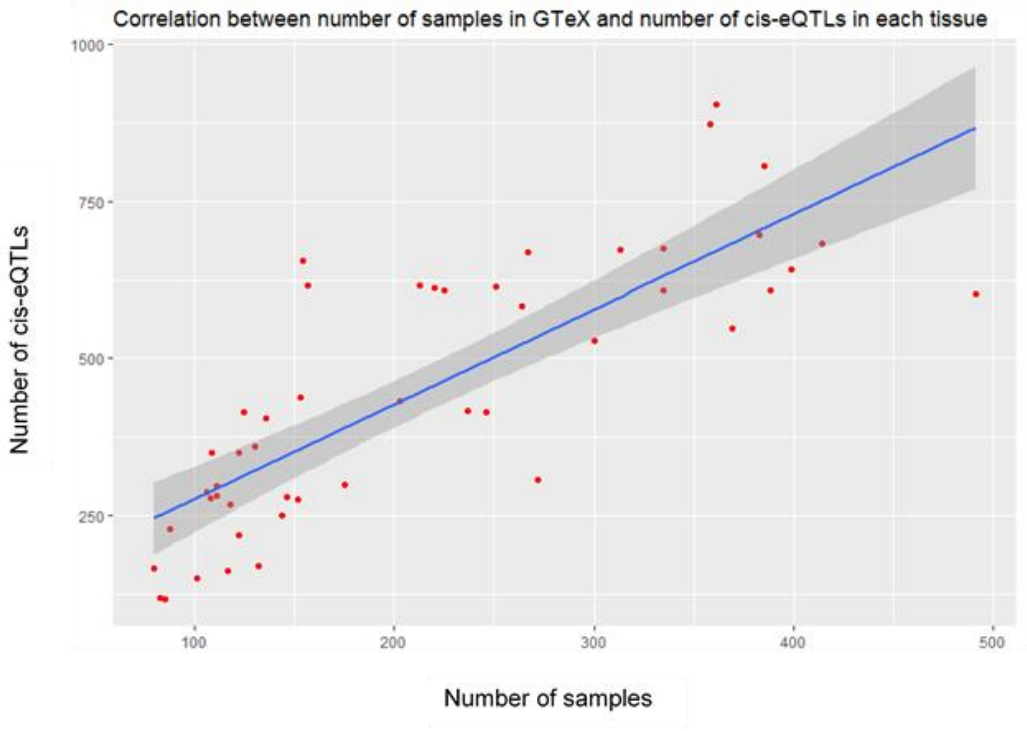

Figure 3.3. Correlation plot between number of samples present in GTEx per tissue and the number of cis-eQTLs in the respective tissue demonstrating a strong correlation between the two (R^2=0.64).

locus. However, the number of *cis*-eQTLs in each tissue correlated positively (R^2=0.64) with the number of samples present in GTEx for the respective tissue (Figure 3.3). Notably, there are some tissues that are outliers in the correlations, including the pituitary, cerebellum, oesophagus, pancreas, breast and adipose tissues. Interestingly, most of these tissues have key functions that relate to GH action (Bartke, 2011; Bonert and Melmed, 2017; Clasen et al., 2014; Duan et al., 2015; Pekic et al., 2017; Tarnawski et al., 2015).

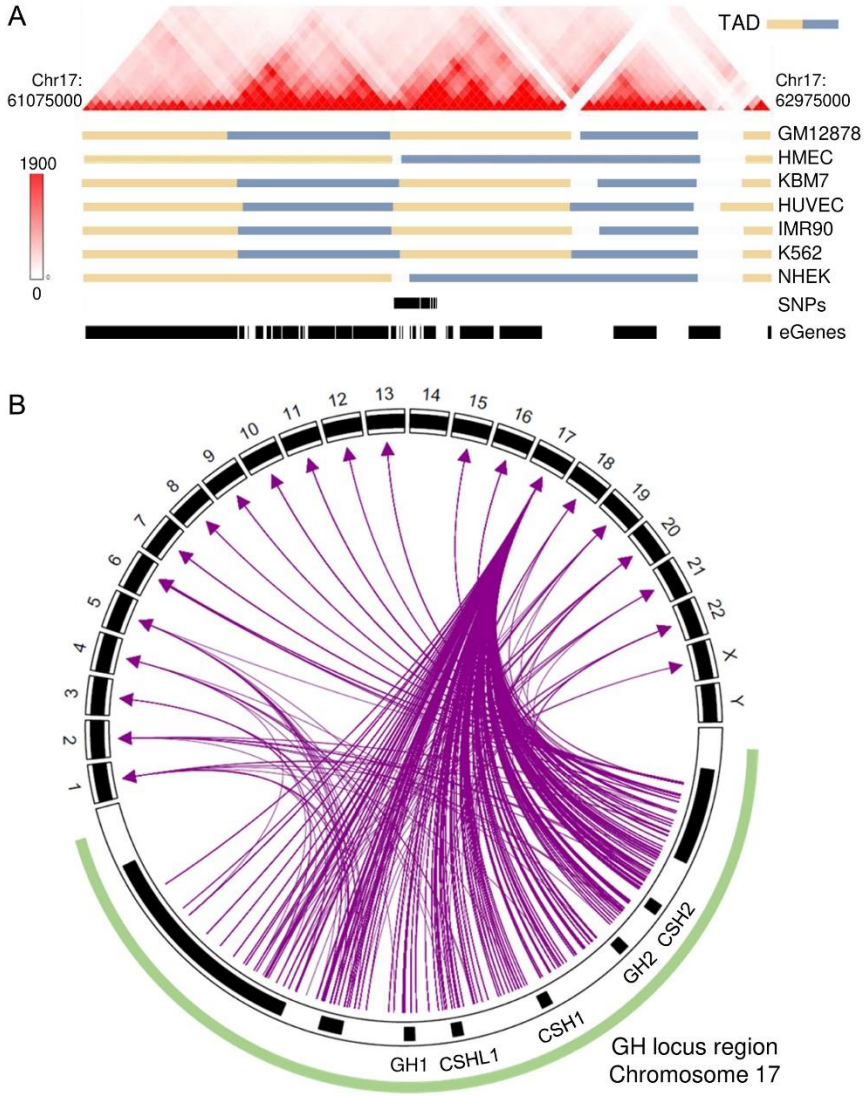

Figure 3.4. Structural analysis of the *GH* locus and eQTLs associated with *GH* locus SNPs
(A) Distribution of topologically associating domain (TAD) structures with cis-eGenes (i.e. eGenes located <1 Mb from the SNP). The Hi-C heat map was generated with the 3D Genome Browser using data from Rao *et al.* (Rao et al., 2014) for all the seven cell-lines (GM12878, HMEC, KBM7, HUVEC, IMR90, K562 and NHEK). Tracks show TAD structures, the region containing the SNPs and identified eGenes. Blue and yellow bars represent different TAD regions. (B) A circos plot

illustrating all the connections from *GH* locus SNPs to genes present on different chromosomes which includes chromosome 17. These eQTLs were not just limited to chromosome 17 but extended to multiple chromosomes. The circus plot has been split into two sections with the green curve representing the *GH* locus region across which the SNPs were analysed.

Within the nucleus, chromosomes are arranged through a hierarchy of structures that include topologically associating domains (TADs). TADs are defined as regions where spatial contacts are enriched (Tang et al., 2015). To determine whether the observed SNP-eGene connections crossed TAD boundaries, a Hi-C heat map was generated at 1 kb resolution using Hi-C data captured within the GM12878, HMEC, KBM7, HUVEC, IMR90, K562 and NHEK cell-lines (Rao et al., 2014). The Hi-C analysis clearly showed that there were numerous SNP-eGene connections observed both within the individual TAD containing the *GH* locus region and across the TAD boundaries (Figure 3.4A, Figure 3.5). This is consistent with SNP-eGene connections not being limited to occurring within TADs (Chen et al., 2018; Ciabrelli and Cavalli, 2015; Ulianov et al., 2016). Connections were also observed with eGenes present on different chromosomes (Figure 3.5B) consistent with the concept that physical interactions between two genomic regions are not restricted by proximity in linear distance.

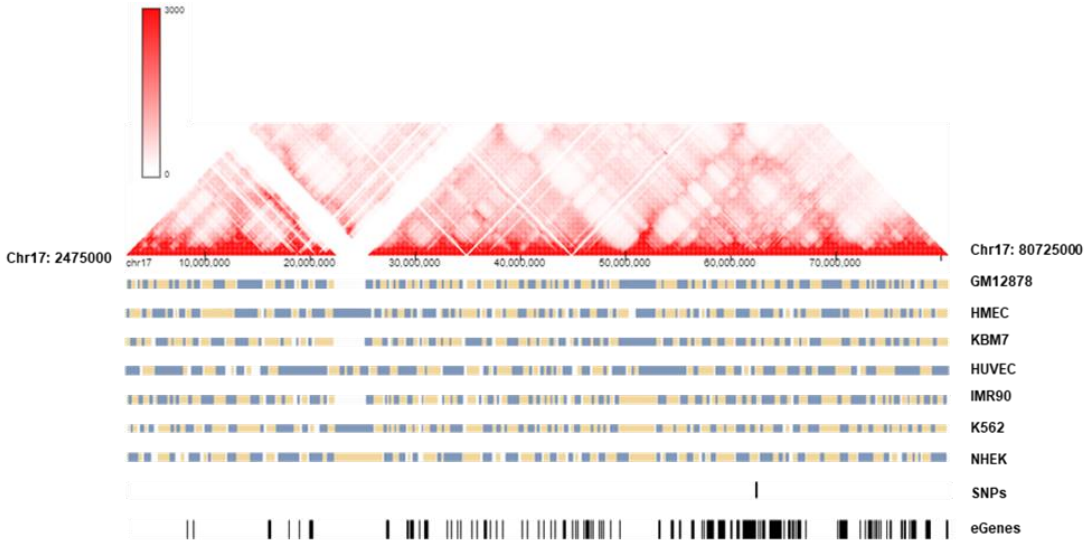

Figure 3.5. Distribution of topologically associating domain (TAD) structures with eGenes across chromosome 17
The Hi-C heat map was generated with the 3D Genome Browser using data from Rao *et al* for all the seven cell-lines (GM12878, HMEC, KBM7, HUVEC, IMR90, K562 and NHEK). Tracks show TAD structures, the region containing the SNPs and identified eGenes. Blue and yellow bars represent different TAD regions.

As the SNPs which were used for our study were part of the *GH* region (which is comprised of *SCN4A*, *CD79B*, *GH1*, *CSHL1*, *CSH1*, *GH2*, *CSH2* and *TCAM1P* genes), it was analysed whether there were any connections going back into these genes. SNPs (n=73) located across the entire *GH* region had connections with *CD79B* in 16 tissues, one SNP with *CSHL1*, three SNPs with *GH2* in pituitary tissue, 114 SNPs with *CSH2* in three tissues and 173 SNPs with *TCAM1P* gene in 24 tissues. As the locus control region is critical for regulation of these genes, the analysis was focused on interactions in this region. 27 SNPs from across the *GH* locus control region connected to *CD79B* in 13 tissues, one SNP to *CSHL1* and one SNP to *GH2* in pituitary tissue, 19 SNPs to *CSH2* in three tissues (testis, cerebellum and cerebellar hemisphere) and 33 SNPs to *TCAM1P* gene in 18 tissues (Supplementary Table 2). This is consistent with Tsai *et al.* 2016 who used chromatin conformation capture (3C) data from human pituitary and placental tissues to demonstrate that the *GH* locus control region regulates these genes (Tsai et al., 2016). Notably, only a few eQTLs with *CSHL1* and *GH2* and no eQTLs with the *GH1* gene were identified, which may reflect the age, sex (65.8% male, 68.5% above 50 yrs of age) and tissue distribution (i.e. no placental samples) that was used to characterise expression profiles within the GTEx database (Ardlie et al., 2015).

CD79B displays a very complex eQTL-pattern. Eight *CD79B* SNPs (rs1051684, rs12603821, rs1051688, rs12451467, rs2070776, rs2005132, rs2320125 and rs8077653) connect to 66 eGenes in 48 different tissues. Out of these genes, 47 eGene contacts were in *trans* (>1 Mb away, distal) and 19 were in *cis* (<1 Mb away, nearby). Out of 47 *trans* eGenes, 24 genes were distal but present on chromosome 17 (up to 59 Mb downstream and 18 Mb upstream from *CD79B*). These eight SNPs were also associated with increased expression of the *CD79B* gene in 11 tissues and a decrease in *CD79B* expression in atrial appendage of the heart. Seventy-three output SNPs from across the analysed region were found to be associated with altered *CD79B* expression in 16 tissues. Some of these SNPs were linked to an increase in expression of *CD79B* whereas some linked to downregulation of this gene in a tissue-specific manner. This differential expression across tissues could imply disruption of binding sites of tissue-specific transcription factors by these or linked SNPs. For example, rs3815358 and rs12452767 occur in a CTCF binding site. CTCF is an important transcriptional regulator protein (Splinter et al., 2006). Disruption of these CTCF binding sites can block or cause inefficient binding of transcription factors, or may impact on other regulatory processes (Ohlsson et al., 2001).

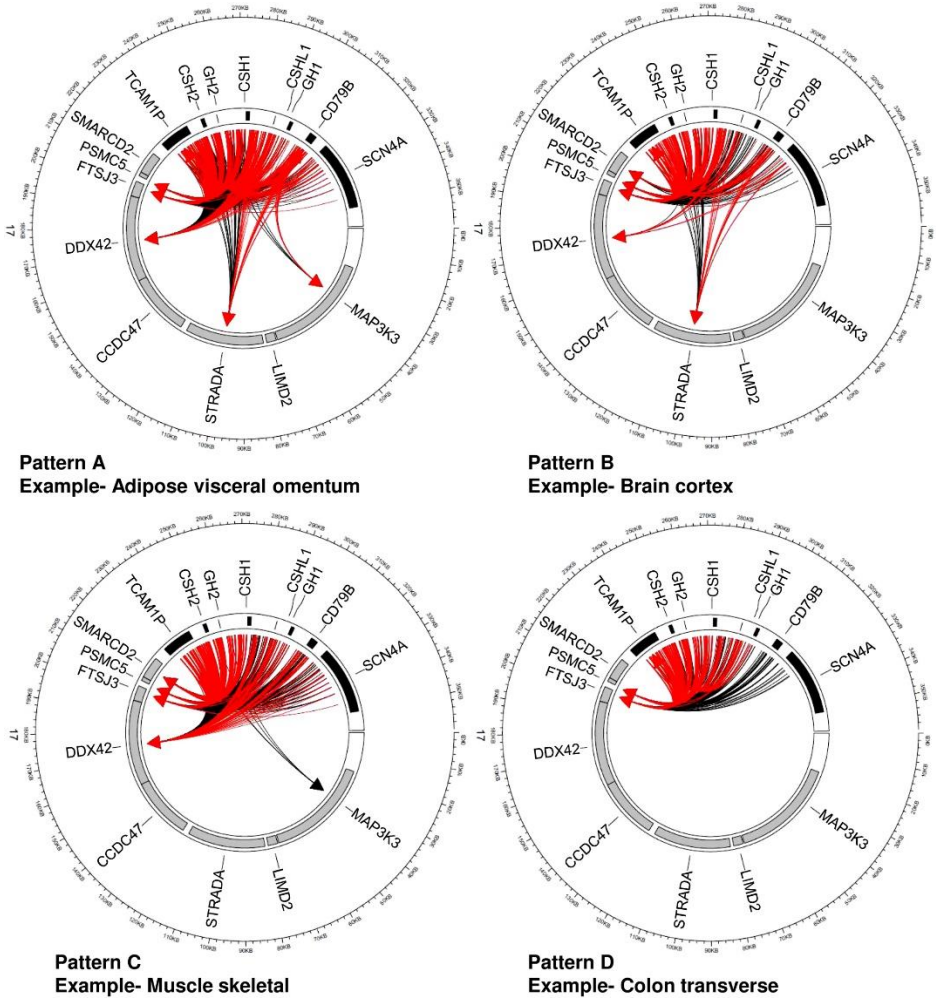

Figure 3.6. Pattern of cis-eQTL connections in different tissues
Identified contacts with genes present in the local environment of the *GH* gene locus appear to cluster together in four different pattern types as shown in the figure. Representative figures for patterns A, B, C and D are the connections observed in omental visceral adipose, cortex, skeletal muscle and transverse colon respectively. Rectangles in black represent the genes of *GH* locus across which SNPs were analysed, those in grey are the associated cis-eGenes. Links in red represent upregulation of the eGene associated with the SNP, whereas the links in black represent downregulation.

To assess the *cis* connections with the *GH* locus SNPs, eQTLs with genes immediately downstream of the *GH* locus- *SMARCD2*, *PSMC5*, *FTSJ3*, *DDX42*, *CCDC47*, *STRADA*, *LIMD2* and *MAP3K3* were investigated. Four strikingly different normalised effect size (NES) patterns in eQTL associations were observed, involving these genes across all the tissues (Figure 3.6). Eleven tissues were classified as pattern A, twelve as pattern B, four as pattern C and fifteen as pattern D. It was interesting to observe such a marked difference between these patterns; however, the functional significance of these differences remains to

be determined. Collectively, our results are consistent with the presence of regulatory sites located within the *GH* locus.

3.3.2 Regions in the *GH* locus interact with multiple genes (*cis* and *trans*) in a tissue-specific manner

Within the *GH* locus, *CSH2* exhibited the maximum number of associations with 114 SNPs in three tissues - testis, brain cerebellum and brain cerebellar hemisphere. In addition, the *ICAM2* and *FTSJ3* genes associated with 108 SNPs in one tissue (skeletal muscle) and 103 SNPs across all 48 tissues respectively (Supplementary Table 3). *CSH2* is part of the *GH* locus and has a high sequence similarity with other *GH* genes.

While the *cis*-eGenes mostly had eQTLs with multiple SNPs in several tissues, *trans*-eGenes exhibited more tissue-specificity (Supplementary Table 4A).139 eGenes formed an eQTL with only one SNP whereas 55 had associations with two SNPs. Conversely, seven genes (*SCPEP1, B3GNTL1, NCOR1, RGS9, VMP1, LINC00511* and *CACNG4*) had eQTLS with multiple SNPs (Supplementary Table 4B). For example, *CACNG4* (Calcium Voltage-Gated Channel Auxiliary Subunit Gamma 4) was found to be associated with 40 SNPs in the atrial appendage.

Interestingly, differences in the proportions of *cis* and *trans* eQTLs were observed in different tissues. For example, the pituitary, which had the highest number of *cis*-eQTLs compared to all the tissues, did not have highest number of trans-eQTLs. Brain-related tissues had the most eQTL-associations in *trans* (Supplementary Table 5), whereas the pituitary had the maximum number of eQTLs with genes in *cis* (Supplementary Table 6). This is particularly notable considering the role of the pituitary in *GH*-related function.

Some SNPs such as rs3815358 connected with multiple eGenes (27 eGenes) in a variety of tissues (37 tissues) whereas a few SNPs such as rs11869827 only associated with one or two eGenes (Supplementary Tables 7 & 8). The maximum number of eGenes which SNPs in our data were associated with was 27. More than 50% of the SNPs exhibited eQTL associations in all 48 tissues which could imply that these SNPs could potentially have a role in regulation of normal cell functioning (Supplementary Table 9).

3.3.3 Genes targeted by *GH* eQTLs are potentially involved in growth hormone functions

To explore if the set of eGenes (Supplementary Table 10) were enriched for canonical pathways, the eGene sets were analysed using g:Profiler (Raudvere et al., 2019) (Supplementary Tables 11 & 12). The top ten most significant (adjusted *P* value) enriched pathways are summarised in Figure 3.7. This analysis demonstrated enrichment for a subset of these genes in numerous GH-related cellular signalling pathways such as the PI3K-Akt, MAPK, mTOR, prolactin, insulin and ErbB signalling pathways (Supplementary Tables 11 & 12). Another pathway which is potentially involved in regulation of GH signalling is the Wnt pathway, which was also among the enriched pathways (Supplementary Table 13). There was also a significant representation of these genes in carcinogenic pathways. Five cancer terms were enriched in the KEGG subset and seven in the WikiPathways subset. There were four cancer types in common in both sets; these were hepatocellular carcinoma, colorectal cancer, breast cancer and non-small cell lung carcinoma. Altered GH signalling has previously been established to be associated with multiple cancer states including the ones identified (Chhabra et al., 2011; Clayton et al., 2011; Perry et al., 2017). There was also eGene enrichment in other pathologic conditions as well like Alzheimer's and Huntington's disease. Altogether, these observations are consistent with the hypothesis that genetic variation in the *GH* locus is associated with modulation of GH function.

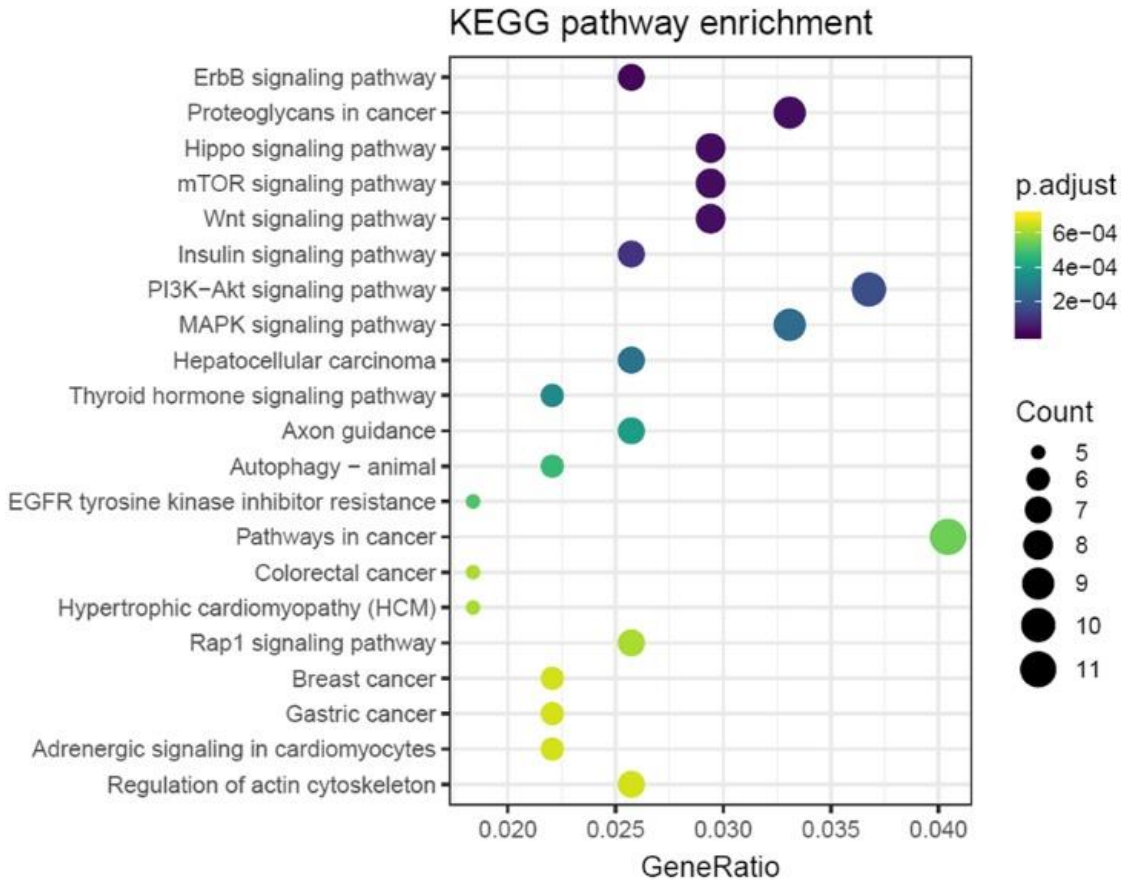

Figure 3.7. Dot plot showing top 20 pathways enriched by KEGG using g:Profiler
The plot is presented in decreasing order of adjusted *P* value. The size of the circle represents the number of genes enriched in the pathways.

3.3.4 Identified eQTLs co-localise with GWAS signals

To determine if the identified eQTLs were associated with a disease or a population, the list of SNPs was cross-referenced with the genome wide association studies (GWAS) Catalog (Buniello et al., 2019) and found that 5 variants (rs2005172, rs2070776, rs2532111, rs28386778 and rs2854160) exhibited GWAS associations with certain traits or populations (Table 3.2). rs2005172, rs2070776, rs2532111 are associated with fat-free mass, height and waist-hip ratio, respectively, in individuals of European ancestry. SNP rs2854160 is associated with height in East Asian and African population whereas rs28386778 is linked to prudent dietary patterns in 141 individuals. This is important in the context of GH biology since regulation of height is one of the major functions modulated by this hormone. Excess or deficit GH can lead to multiple growth disorders such as dwarfism, gigantism and acromegaly. SNP rs2070776, which is linked to height in European population, connects to

26 eGenes, (11 genes in *cis* and 15 in *trans*) across 39 different tissues. 6 of the *trans*-eGenes are located on different chromosomes. One of these identified eGenes, *NF1* (Neurofibromin 1), which is more than 1 Mb distance away from rs2070776, is known to be strongly related to human height (Kehrer-Sawatzki et al., 2017). The height-related SNP rs2854160 in individuals of East Asian and African descent connects to 13 eGenes in 42 tissues in which there are 11 *cis*-eGenes and 2 *trans*-eGenes (on the same chromosome). These two SNPs are located in different regions in the *GH* locus.

Table 3.2. Subset of identified SNPs in the GWAS catalogue

SNPs	Pubmed ID	Disease/Trait	Sample size	Context	P value
rs2005172	30593698	Fat-free mass	70,700 European ancestry female individuals, 85,261 European ancestry male individuals	Intron variant	8E-14
rs2070776	20881960	Height	133,653 European ancestry individuals	Stop gained	9E-09
rs2070776	25282103	Height	253,288 European ancestry individuals	Stop gained	6E-41
rs2532111	30595370	Waist-hip ratio	approximately 458,000 European ancestry individuals	3 prime UTR variant	6E-10
rs28386778	28644415	Prudent dietary pattern	141 individuals	Regulatory region variant	0.000006
rs2854160	25429064	Height	36,227 East Asian ancestry individuals	Intergenic variant	2E-12
rs2854160	21998595	Height	20,427 African ancestry individuals	Intergenic variant	5E-08

3.3.5 Regulatory potential of *GH* eQTLs are supported by functional studies

Next, experimental evidence for the SNP eQTL-eGene connections was sought from published literature. van Arensbergen *et al*. recently developed a functional screening method known as Survey of Regulatory Elements (SuRE) which identifies SNPs that impact on regions of regulatory significance (van Arensbergen et al., 2019). Using this method, a total of 5.9 million SNPs were surveyed in two cell-lines - human erythroleukemia cells (K562) and human hepatocellular carcinoma cells (HepG2), to identify SNPs that alter the activity of putative regulatory elements(van Arensbergen et al., 2019). 33 of the SNPs identified in our study were demonstrated to be of regulatory importance by van Arensbergen *et al.* 16 SNPs were identified in HepG2 cells, 16 in K562 cells, whereas one SNP, rs12451467 was found to be significant both cell-lines (Supplementary Table 14). Notably, our results also revealed Hi-C interactions between 17 SNPs with eGenes in K562. For example, SNP rs2584608 has regulatory interactions with 12 eGenes (including the chromatin-remodelling factor SMARCD2 and papillary thyroid carcinoma biomarker LIMD2), across 48 different tissues (Supplementary Table 15).

To further strengthen the functional application of our approach, known enhancer/promoter regions were analysed using data retrieved from GeneHancer. There were 16 enhancer/promoter regions in the region used for our study (Chr17:62080000-61920000; GRCh37/hg19) which target 13 of the identified eGenes (Supplementary Table 16). These regulatory regions encompass 7 of the output SNPs (rs12452767, rs2286564, rs2286565, rs2457681, rs2665808, rs2854184, rs34684062, rs3815358, rs6171 and rs8080613), out of which, two SNPs (rs2286564, rs2286565) were shown to have functional regulatory relevance in the data shown above (Supplementary Table 14). Collectively, these observations provide support for the functional roles of the eQTLs identified in this study.

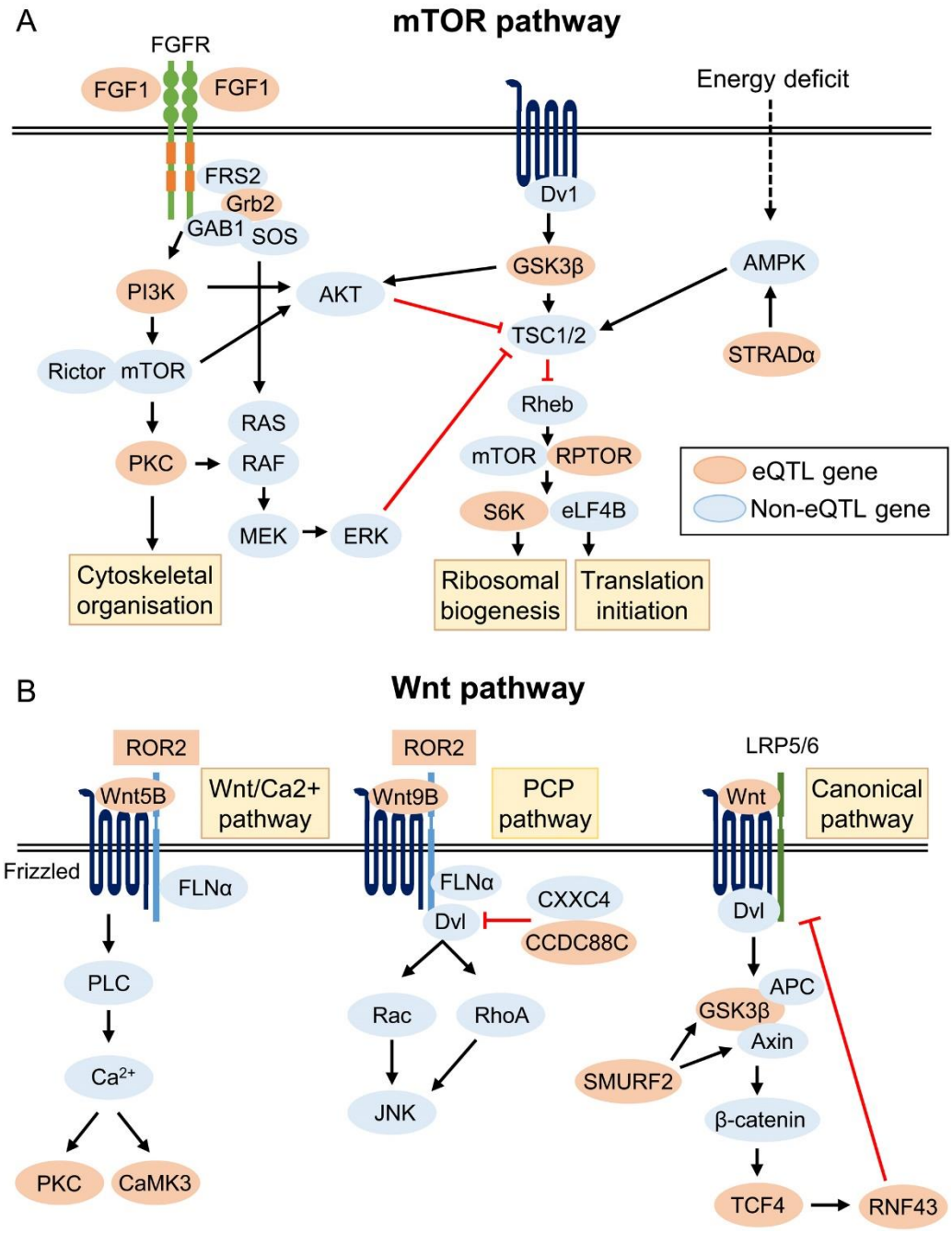

Figure 3.8. eGenes identified by CoDeS3D enriched in the mTOR (A) and Wnt (B) signalling pathways.
These figures illustrate where these pathway related eGenes fit in the cellular signalling cascade. Genes/proteins in orange boxes represent eQTL pathway-related genes whereas those in blue boxes are non-eQTL genes. The eGenes and tissues in which these eQTLs were identified and the SNPs they connect to are summarised in Supplementary Tables 17 and 18.

3.4 Discussion

Common SNPs within the *GH* locus were analysed, and the *GH* locus was identified as a potential regulatory hub for genes located not only on chromosome 17 but also on other chromosomes. Many of the identified genes are a part of GH related cellular signalling pathways including pathways in cancer. These results suggest that some GH functions could potentially be mediated by interaction of regulatory regions within the *GH* locus region.

On examining the specifics of these eQTL-eGene interactions, interesting gene expression patterns were observed in a distinct tissue-specific manner. 114 SNPs from the entire *GH* region were found to be associated with altered regulation of gene expression of *CSH2* gene in the testis and brain tissues (cerebellar hemisphere and cerebellum). It should be noted that samples named brain cerebellum and brain cerebellar hemisphere in GTEx are considered duplicates, as they were the same tissue taken at different times post-mortem (Ardlie et al., 2015). The presence of the eQTL in both indicates that the mRNA was not subject to rapid degradation due to senescence. Most of the *CSH2* associated eQTL SNPs were found clustered just downstream of *CSH2* gene. Although *CSH2* mRNA expression is extremely low in the testis and brain (GTEx), CSH2 protein expression is observed in some germ cell tumours of the testis (Berger et al., 1999). CSH2 has also been linked to disease states including fetal growth disorders pre-eclampsia, and choriocarcinoma (Kim, 2003; Liu et al., 2011; Männik et al., 2010).

CD79B is an immune-related gene which is important for initialising signal transduction activated by the B-cell receptor complex (Alfarano et al., 1999). Polymorphisms in this gene have been linked to several types of cancer. Mutations in *CD79B* have been linked to different types of diffuse large B-cell lymphoma (DLBCL) (Frick et al., 2018; Schrader et al., 2018) and to osteosarcoma (Mirabello et al., 2011). Identified SNPs in *CD79B* and *CSH2* have enhancer marks in multiple human tissues and cell-lines. Analysis of known regulatory regions from GeneHancer showed that GH17J063930 is an enhancer for both *CD79B* and *CSH2, and* GH17J063877 is an enhancer for *CSH2.* This super-enhancer region encompasses 4 SNPs (rs2286564, rs2286565, rs3815358 and rs12452767), and the *CSH2* enhancer overlaps with one SNP, rs2457681. Consistent with this, all four SNPs had eQTLs with *CD79B* and *CSH2* in our analysis, and rs2457681 connected with *CSH2* (Supplementary Table 2). The study b*y van Arensbergen et al.*, which validated SNPs important for regulatory activities, identified two of the SNPs from this enhancer region (rs2286564 and rs2286565)

as variants which are of regulatory importance (van Arensbergen et al., 2019). This is consistent with our hypothesis that the entire *GH* locus potentially serves as a complex regulatory region.

Pathway analysis (g:Profiler) of the identified eGenes (292 genes) identified enrichment of these genes in various cellular signalling pathways including classical *GH*-related pathways such as the PI3K-Akt, mTOR, MAP kinase, prolactin and insulin signalling pathways. The role of mTOR signalling in growth and development and mediation of GH function is well established. mTOR is a protein kinase that regulates cell metabolism, proliferation and survival and is important for mediating GH-related pathways, including those related to insulin and IGF actions (Bartke, 2011; Hayashi and Proud, 2007; Saxton and Sabatini, 2017), so it was interesting to find that eGenes associated with SNPs from the locus were enriched in the mTOR pathway (Supplementary Table 17). Another pathway enriched in this analysis was the Wnt signalling cascade, which mediates embryonic development pathways, cell polarity, migration and division (Supplementary Table 18). The enriched eGenes were part of both canonical and non-canonical Wnt signalling (Figure 3.8). Since, there is significant crosstalk between Wnt and other signalling pathways like mTOR, MAP kinase and P13K, it is possible that there was just an overlap between GH-mediated signalling and the Wnt pathway. However, several studies suggest a potential connection between GH and Wnt signalling (Osmundsen et al., 2017; Vouyovitch et al., 2016). For example, Vouyovitch *et al.* demonstrated that GH regulates the expression of the secreted protein, Wnt4, a Frizzled receptor ligand involved (Vouyovitch et al., 2016). This suggests that GH may mediate some of its functions through direct impact on gene expression levels of the Wnt signalling cascade components, and this may be mediated by spatial interaction of these regions with the *GH* locus. To confirm this, putative regulatory regions can be confirmed with reporter assays, and single-nucleotide editing techniques such as CRISPR/Cas coupled with proximity ligation assays like Hi-C and gene expression studies may substantiate the impact of allele variants on the expression of Hi-C linked genes *in vitro* or in animal models. However, studies of this nature are complex as regulatory regions act in a combinatorial manner and therefore, modification of isolated SNPs may have minimal impact.

Pathway enrichment analysis identified that eGenes associated with *GH* locus SNPs were overrepresented in pathological conditions such as cancer, growth disorders and Alzheimer's disease. There is extensive literature supporting a role for GH in cancer. Altered GH signalling and increased expression of *GH* and the genes it is associated with functionally,

such as *IGF1* is linked to progression of numerous cancer types (Lu et al., 2019; Perry et al., 2013; Simpson et al., 2017). Autocrine GH increases the size of hepatocellular tumour xenografts and is associated with a worse relapse-free and overall survival in patients with hepatocellular carcinoma (Kong et al., 2016). GH decreases expression of the tumour suppressor gene *p53* in the colon and contributes to the development of colorectal carcinoma (Chesnokova et al., 2016). Elevated GH expression is also observed in colorectal cancer and is positively associated with tumour size and lymph node metastasis (Wang et al., 2017). Similarly, there are multiple studies which have established the key role of GH in breast cancer, endometrial and non-small cell lung carcinoma (Chhabra et al., 2011, 2018; Lu et al., 2019; Pandey et al., 2008; Perry et al., 2013). Identified eGenes in our data show significant enrichment in all of these cancer types which suggests that the eGene set is strongly associated to GH-related cancer types, not purely by chance. This data suggests that genes in and around *GH* locus and those associated with polymorphisms across the locus have a direct or indirect association with cancer, possibly by alteration of gene regulatory networks.

This study also overlapped with other recent studie,s which identified genes associated with Alzheimer's disease and the eGenes in our data. There were five genes (*ACE*, *PRKCA*, *ERN1*, *GSK3B* and *MAPT*) from one study (Grimm et al., 2019) and two genes from another (*KRTAP5-AS* and *PSMC5*) (Kikuchi et al., 2019) which were shared with the identified eGene set. This could possibly indicate a link between *GH* gene locus mediated/coordinated gene regulation and Alzheimer's disease. However, significant experimental evidence is needed to substantiate this.

This study shows that there is a physical contact between SNPs across the *GH* locus (including the locus control region) and *GH2* and *CSHL1* genes. In addition, an association of these SNPs with change in expression levels *GH2* and *CSHL1* was identified in pituitary tissue. This result can be linked back to the studies from the Liebhaber and Cooke Laboratories (Ho et al., 2008; Kimura et al., 2007; Tsai et al., 2016), which show that genes of the *GH* locus (*GH1*, *GH2*, *CSH1*, *CSH2* and *CSHL1*) are expressed in the placenta or pituitary in a tissue-specific manner and are coordinated by the locus control region overlapping genes *CD79B* and *SCN4A.* This regulation is under epigenetic control and is affected by chromatin looping that brings these regions into close proximity with the target promoters.

GH is crucial for modulation of normal growth and metabolism in the human body, through stimulation of signalling pathways and regulation of multiple growth factors. It is well established that GH-related downstream pathways such as mTOR, and other pathways identified here, such as Wnt signalling, are also pivotal in central biological roles affecting growth, development and metabolism (Liu and Sabatini, 2020; Taciak et al., 2018). The impact of GH on critical signalling cascades and their associated functions is likely a consequence of the interplay of genetic, molecular and evolutionary factors. Collectively, our observations support the hypothesis that the *GH* locus (which includes *GH1*, *GH2*, and *CSH1*, *CSH2* and *CSHL1* and its locus control region spanning genes *CD79B* and *SCN4*) functions as an extended regulatory region that coordinates expression of genes located both within and outside of the locus, which are in key GH-linked pathways and contribute to GH function. This is particularly important to consider in the context of evolutionary biology, as it is unlikely that co-regulation of these pathways has evolved by chance. Instead, it is consistent with the premise that genetic regions that are critical in coordinated regulation of aligned biological processes are linked as part of their maintenance. In the context of the human *GH* locus it is notable that most non-primate mammals only have a single *GH* gene and the *GH* cluster arose from gene duplication independently in New World and Old World Monkeys, and thus varies considerably in structure (González Alvarez et al., 2006; Wallis and Wallis, 2006). Comparative studies of the 3D interactions between the *GH* locus of primate and non-primate mammals, with genes in pathways identified in our study may highlight the potential role of co-regulation of these pathways in evolution.

CHAPTER 4

Growth Hormone Receptor Interacts with Transcription Factors upon GH-Induced Nuclear Translocation

This chapter has been formatted for publication

4.1 Introduction

Recent advances in cellular and molecular technology have revitalised interest in the unconventional behaviour of cell-surface receptors such as nuclear translocation. Throughout the 1980s-1990s, an increasing number of reports identified the presence of various classic cell-surface receptors in the nucleus of both normal and malignant cells (Lobie et al., 1994; Maher, 1996; Podlecki et al., 1987). This unusual phenomenon may also have an association with disease, in particular cancer (Conway-Campbell et al., 2007; Wang and Hung, 2009).

Human growth hormone (GH), a peptide hormone synthesised by the somatotrophic cells of the anterior pituitary gland, is released in a pulsatile fashion from the pituitary, with secretion modulated by GH releasing hormone (GHRH), somatostatin, and ghrelin (Bonert and Melmed, 2017). Classically, GH exerts its impact on cell growth and differentiation by binding and activation of the GH receptor (GHR) (Waters, 2016), followed by stimulation of secondary peptide mediator molecules, such as insulin-like growth factor-1 (IGF1), in an endocrine, paracrine, and autocrine manner (Bonert and Melmed, 2017; Harvey et al., 2015; Nilsson et al., 1990). The GHR belongs to the cytokine type I receptor family, which is characterised as single-pass transmembrane receptors containing a minimum of one classic cytokine receptor homology (CRH) domain. GH binding to the GHR is followed by conformational changes in the intracellular domain of the receptor, which leads to activation of non-receptor protein tyrosine kinases, JAK2, and c-SRC, thus initiating an entire cascade of downstream signal transduction (Carter-Su et al., 2016; Dehkhoda et al., 2018; Lu et al., 2019; Waters, 2016). These signal transduction pathways mediate effects by alteration of gene transcription profiles through direct stimulation of transcription and by regulation of histone modifications (Perry et al., 2006; Rotwein and Chia, 2010).

The GHR is internalised following activation by GH (Lobie et al., 1994b; Sachse et al., 2001). The major target for the internalised GH/GHR complex is the lysosomes, where it is degraded and ensures cellular GHR levels are maintained at a normal physiological state (Van Kerkhof et al., 2000; Strous and Van Kerkhof, 2002). GHR internalisation is facilitated by pathways such as the clathrin and/or caveolin systems (Lobie et al., 1999; Sachse et al., 2001; Yang et al., 2004). Recent findings have demonstrated that the GHR is targeted for proteasome-dependent degradation by SOCS2-mediated ubiquitination. *SOCS2* mRNA

expression is induced by GH stimulation, and the protein is a crucial negative regulator of GHR expression and consequential signal transduction (Greenhalgh et al., 2002). SOCS2 decreases GHR levels in the cell by increasing proteasomal degradation of GHR (Greenhalgh et al., 2005; Vesterlund et al., 2011).

In addition to classic cell surface GHR-mediated signalling, studies from the Waters lab have demonstrated that the GHR is rapidly imported into the nucleus of cells upon stimulation with GH, in both normal and diseased cell types (Conway-Campbell et al., 2007; Lincoln et al., 1998; Lobie et al., 1992, 1994a; Mertani et al., 1998). This has subsequently been demonstrated to occur in cell-lines from multiple species, including rodents, pigs and fish (Conway-Campbell et al., 2007; Figueiredo et al., 2016; Lan et al., 2017). This nuclear localisation has been found to be associated with increased proliferation of cells, especially cancer cells, leading to increased tumorigenesis in human cell-lines and transgenic zebrafish (Conway-Campbell et al., 2007; Figueiredo et al., 2016). Intriguingly, numerous other classic cell-surface receptors that were previously thought to signal exclusively at the cell-surface have also been demonstrated to localise to the nucleus, and some have been shown to directly influence gene expression. For example, nuclear epidermal growth factor receptor (EGFR) functions as a transcription factor for multiple genes (reviewed in (Brand et al., 2011; Bryant and Stow, 2005; Carpenter and Liao, 2013; Shah et al., 2019)).

The consequences of nuclear GHR import remain to be determined, despite a correlation with cell proliferation. Here, it was investigated whether the GHR interacts with proteins in the nucleus following activation and nuclear translocation. Integration of GHR immunoprecipitation-mass spectrometry data identified multiple proteins which immunoprecipitated with the GHR, including two transcription factors, HMGN1 and SUMO1. Target genes of HMGN1 and SUMO1 were found to be differentially expressed following GH treatment. This study provides further insight into the significance of nuclear import of GHR upon treatment as an auxiliary regulatory mechanism for modulation of GH activity.

4.2 Materials and Methods

4.2.1 Cell-lines and reagents

The human endometrial carcinoma cell-line RL95-2 was maintained in RPMI-1640 growth medium (Life Technologies) at 37°C, 5% CO_2 supplemented with 5% fetal bovine serum (Thermo Scientific), 100 μg/ml streptomycin (Sigma-Aldrich), 100 U/ml penicillin (Sigma-Aldrich), and Glutamax. The human mammary epithelial cell-line MCF-10A was maintained in MEGM^{TM} Mammary Epithelial Cell Growth Medium BulletKitTM (Lonza) supplemented with 100 μg/ml streptomycin (Sigma-Aldrich), 100 U/ml penicillin (Sigma-Aldrich), and 5% horse serum (Sigma-Aldrich) in a humidified chamber with 5% CO_2 at 37°C. Cells in the logarithmic phase of growth were used for hormone treatment, western blot, mass spectrometry, and microarray analyses.

Recombinant GH was purchased from the National Hormone and Pituitary Program (NHPP) and reconstituted in sterile phosphate-buffered saline (PBS), pH 7.4. Antibodies used in the experiments were Phospho-STAT5 (pTyr694) (Life Technologies, 716900), STAT5 (C-17) (Santa Cruz Biotechnology, sc-835), β-ACTIN (Sigma-Aldrich; A1978), GAPDH (Life Technology), GHR extracellular domain (Abcam, ab89400) GHR intracellular domain (Santa Cruz Biotechnology, sc-137185), epidermal growth factor receptor (EGFR) (Cell Signalling, 2232-S), insulin-like growth factor-1 receptor (IGF1R) (Cell Signalling, 3027), HMGN1 (Life Technologies, 720387), SUMO1 (Life Technologies, 332400), anti-mouse secondary (Sigma-Aldrich, A4416) and anti-rabbit (Sigma-Aldrich, A3687) secondary antibodies.

4.2.2 Western blotting

Cells were plated in T75 cell culture flasks at 5 x 10^6 cells per flask, serum-starved overnight then treated with GH at different concentrations (0, 50, 100, 250 and 500 ng/ml). For whole cell lysates, cells were lysed in 50 mM Tris-HCL (pH 7.4),1% Nonidet P-40; 150 mM NaCl, 1 mM EDTA, 1 mM NaF, 1 mM phenylmethylsulfonylfluoride, 1 mM Na_3VO_4, and protease inhibitor cocktail (Life Technologies). For subcellular fractionation, cytoplasmic and nuclear fractions were extracted using the NE-PER fractionation kit (ThermoFisher Scientific), according to the manufacturer's instructions.

Proteins were then resolved by SDS-PAGE (12% resolving gel) and transferred to nitrocellulose membranes. Membranes were blocked with 5% BSA for 1-2 hours at room

temperature. After washing with PBS-T (PBS + 0.1% Tween 20) solution, membranes were incubated overnight at 4°C, with the primary antibodies as indicated. Membranes were subjected to three PBS-T washes, followed by incubation with a horseradish peroxidase–conjugated secondary antibody for 1 hour at room temperature. Membranes were then washed three times, and the proteins were visualised using Clarity Western Peroxide Reagent (Bio-Rad) and Bio-Rad Chemidoc MP system.

4.2.3 Immunofluorescence

RL95-2 or MCF-10A cells (2 x 10^4) were plated on coverslips in 6-well plates and serum-starved overnight. Following stimulation with GH at the indicated times and concentrations, cells were fixed with 4% paraformaldehyde (w/v) at 37°C for 10 min. Fixed cells were washed with PBS three times and permeabilised with 1% Triton X-100 for 30 min. After washing, the coverslips were blocked with 5% BSA in PBS (pH 7.4) for 2 hours at room temperature, washed with PBS and incubated with primary or mouse IgG control antibodies overnight at 4°C, followed by secondary antibodies for 1 hour at 37°C. Coverslips were washed, stained, and mounted using SlowFade™ Diamond Antifade Mountant with DAPI (ThermoFisher Scientific). Cells were visualised using confocal laser scanning microscopy (Zeiss LSM 800 Airyscan confocal microscope) with ×63 oil immersion objectives. Image analysis was performed using ZEN Blue and ImageJ software.

4.2.4 Immunoprecipitation

RL95-2 cells were plated in T-175 cell culture flasks with 10 x 10^6 cells per flask and were serum-starved overnight. Following GH treatment (500 ng/ml) for 10 min, cells were either lysed to prepare whole-cell lysates (for mass spectrometry) or fractionated into nuclear and cytosolic lysates (method described above). Cell lysates were pre-cleared with magnetic protein-G conjugated beads. Pre-cleared lysates (1 mg/ml) were incubated with 10 µg of primary antibody or isotype control overnight at 4°C on a rotating wheel. 100 µL of protein-G conjugated beads were then incubated with lysate-antibody complexes at 4°C for 2-4 hours with rotation. The magnetic beads were separated using a magnetic rack, and the solution was discarded. Beads were then washed 5-8 times with PBS pH 7.4 to reduce non-specific binding. The immunoprecipitated proteins were then detected by western blotting under reducing conditions, as described above, or by mass spectrometry. For mass spectrometry, protein lysates for each replicate were diluted to a concentration of 1 mg/ml.

4.2.5 Mass spectrometry

Following immunoprecipitation, beads were washed twice with freshly prepared 100 mM ammonium hydrogen carbonate (AMBIC) solution. Protein complexes bound to the washed beads were eluted by incubation with 100 µl of 5% acetic acid for 2 min. Following this, the eluate was concentrated using a centrifugal vacuum concentrator, and diluted with bicarbonate solution to pH 8. This sample was subjected to reduction with dithiothreitol, alkylation with iodoacetamide, and digestion with 0.5 µg sequencing grade modified porcine trypsin (Promega, Madison, WI, USA). The digest was then desalted on HLB cartridges (Waters, Milford, MA, USA) by solid-phase extraction (SPE), and concentrated to a volume of ~15 µl using a centrifugal vacuum concentrator.

Samples were injected onto a 0.3 x 10 mm trap column packed with 3u Reprosil C18 media (Dr Maisch), and desalted for 5 min at 10 µl/min, before being separated on a 0.075 x 200 mm picofrit column (New Objective) packed in-house with 3u Reprosil C18 media. The following gradient was applied at 300 nl/min using a Exsigent NanoLC 400 UPLC system (Sciex): 0 min 5% B; 45 min 40% B ; 47 min 95% B ; 50 min 95% B; 50.5 min 5% B; 60 min 5% B, where A was 0.1% formic acid in water and B was 0.1% formic acid in acetonitrile.

The picofrit spray was directed into a TripleTOF 6600 Quadrupole-Time-of-Flight mass spectrometer (Sciex) scanning from 350-2000 m/z for 200 ms, followed by 45 ms MS/MS scans on the 40 most abundant multiply-charged peptides (m/z 80-1600) for a total cycle time of ~1.8 sec. The mass spectrometer and HPLC system were under the control of the Analyst TF 1.7 software package (Sciex).

4.2.6 Mass spectrometry data analysis

The resulting mass spectrometry data was searched against a database comprising Uniprot Human entries appended with a set of common contaminant sequences, using ProteinPilot version 5.0 (Sciex) (Seymour and Hunter, 2015). Search parameters were the following: Sample Type, Identification; Search Effort, Thorough; Cys Alkylation, Iodoacetamide; Digestion, Trypsin. The peptide summary exported from ProteinPilot was further processed in Excel to remove proteins with Unused Scores below 1.3, to eliminate inferior or redundant peptide spectral matches, and to sum the intensities for all unique peptides from each protein.

Following this, peptides that were present in even one replicate of IgG control samples were filtered out from the GH untreated and treated datasets.

The processed data was then loaded into an MSstats R package (Choi et al., 2014) for normalisation and differential expression analysis. Normalisation was performed to remove systematic bias between MS runs. The default 'equalizeMedians', which represents constant normalisation (equalising the medians) based on reference signals, was used. This was followed by summarisation of the data using a TMP (Tukey Median Polish) algorithm, followed by differential expression analysis between the control and treatment samples. *P* values were adjusted using the Benjamini-Hochberg correction method (Benjamini and Hochberg, 1995). Pathway analysis was performed using g:Profiler package in R (Raudvere et al., 2019).

4.2.7 RNA extraction, RT-PCR, and microarray

RL95-2 or MCF-10A cells were plated in 10cm cell culture dishes at 5 x 10^6 cells per dish, serum-starved overnight, then treated with 500 ng/mL GH for 90 min. Total RNA was extracted using TRIzol reagent (Life Technologies), and column-purified using the RNeasy mini kit (Qiagen) according to the manufacturer's instructions. The purity and concentration of extracted RNA were determined by $OD_{260}/OD_{280,}$ using a NanoDrop ND-1000 instrument (Thermo scientific). The integrity of extracted RNA was assessed by Bioanalyser, using an RNA 6000 Nano LabChip kit according to the manufacturer's instructions (Agilent).

Microarray analysis was performed by Auckland Genomics (University of Auckland), using ClariomD microarray. Sample labelling, microarray hybridisation, and washing were performed according to the manufacturer's instructions. The array images were acquired by means of the Affymetrix GeneChip Operating Software.

4.2.8 Microarray data analysis

Data normalisation and subsequent data processing were performed in R (R version 3.4.4). Differentially expressed genes at each time point were estimated by using R packages maEndToEnd (Baszczyński and Goldstein, 1967) and limma (Ritchie et al., 2015). Transcripts with a *P*-value < 0.01 was considered significantly differentially expressed. Differentially expressed genes were normalised at the transcript level, using the robust multi-array average method. A volcano plot for representation of differentially expressed genes was

generated using the ggplot2 R package (Gómez-Rubio, 2017). Pathway analysis of differentially expressed genes was performed with the g:Profiler package in R (Raudvere et al., 2019).

4.3 Results

4.3.1 The GHR translocates into the nucleus following stimulation with GH

It was investigated whether GHR translocates to the nucleus in the endometrial cancer cell-line, RL95-2, and a mammary epithelial cell-line, MCF-10A. First, it was confirmed that *GHR* mRNA was expressed in both cell-lines by semi-quantitative RT-PCR, with higher expression observed in the RL95-2 cell-line (Figure 4.1A). Both RL95-2 and MCF-10A cell-lines expressed low levels of prolactin receptor (PRLR) mRNA (Figure 4.1A), which is important to note, as human GH can also activate the human *PRLR* (Goffin et al., 1996).

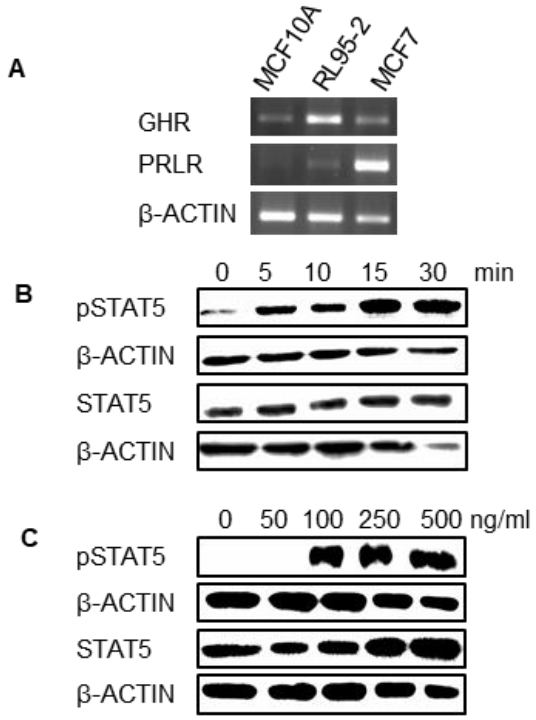

Figure 4.1. Characterisation of GH response in RL95-2 cells
(A) Semi-quantitative RT-PCR analysis of *GHR* and *PRLR* mRNA expression in cell-lines RL95-2, MCF-10A, and the breast cancer cell-line MCF7 (positive control). (B) GH treatment time-course in RL95-2 cells. Serum-starved RL95-2 cells were treated with 500 ng/ml recombinant human GH for 0,

5, 10, 15, and 30 min. (C) GH treatment concentration-response in RL95-2 cells. Serum-starved RL95-2 cells were treated with 0, 50, 100, 250, and 500 ng/ml GH for 15 min. Cell lysates were immunoblotted for phosphorylated STAT5 (pSTAT5) and total STAT. β-ACTIN was used as a loading control for all the experiments above.

To determine the response of RL95-2 and MCF-10A cells to GH, GH-dependent phosphorylation of the downstream signal transduction molecule, STAT5 was measured at 0, 5, 10, 15, and 30 min post-treatment, and in response to different GH concentrations (0, 50, 100, 250 and 500 ng/ml). GH treatment increased STAT5 activation in RL95-2 and MCF-10A cells, as demonstrated by western blot analysis of phoshoSTAT5 (Figure 4.1B and C, and Figure 4.2), with maximal stimulation after 15 min observed at a concentration of 500 ng/ml GH in RL95-2 cells (Figure 4.1C) and 250 ng/ml in MCF-10A cells (Figure 4.2).

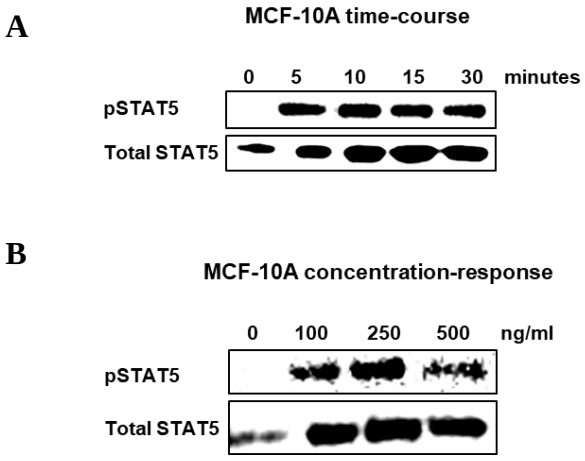

Figure 4.2. GH increases STAT5 phosphorylation in MCF-10A cells
(A) GH treatment time-course in MCF-10A cells. Serum-starved MCF-10A cells were treated with 500 ng/ml recombinant human GH for 0, 5, 10, 15, and 30 min. (B) GH treatment concentration response in MCF-10A cells. Serum-starved MCF-10A cells were treated with varying doses of recombinant human GH, 0, 100, 250, and 500 ng/ml for 15 min. Cell lysates were immunoblotted for phosphorylated STAT5 (pSTAT5) and total STAT.

It was further determined if the GHR translocated to the nucleus in RL95-2 and MCF-10A cells, using immunofluorescence and western blotting. Immunofluorescence consistent with nuclear GHR localisation was observed in RL95-2 cells 5 min after treatment with 500 ng/ml GH. Maximal localisation was observed at 10 min, with a significant decline at 15 min (Figure 4.3). A similar trend was observed for MCF-10A cells in immunofluorescence studies (Figure 4.4). Nuclear localisation of the GHR was also detected, using anti-GHR antibodies

that target the intracellular cytoplasmic and extracellular domains of the GHR in both cell-lines (anti-GHR_{IC} and anti-GHR_{EC} respectively) (Figure 4.5).

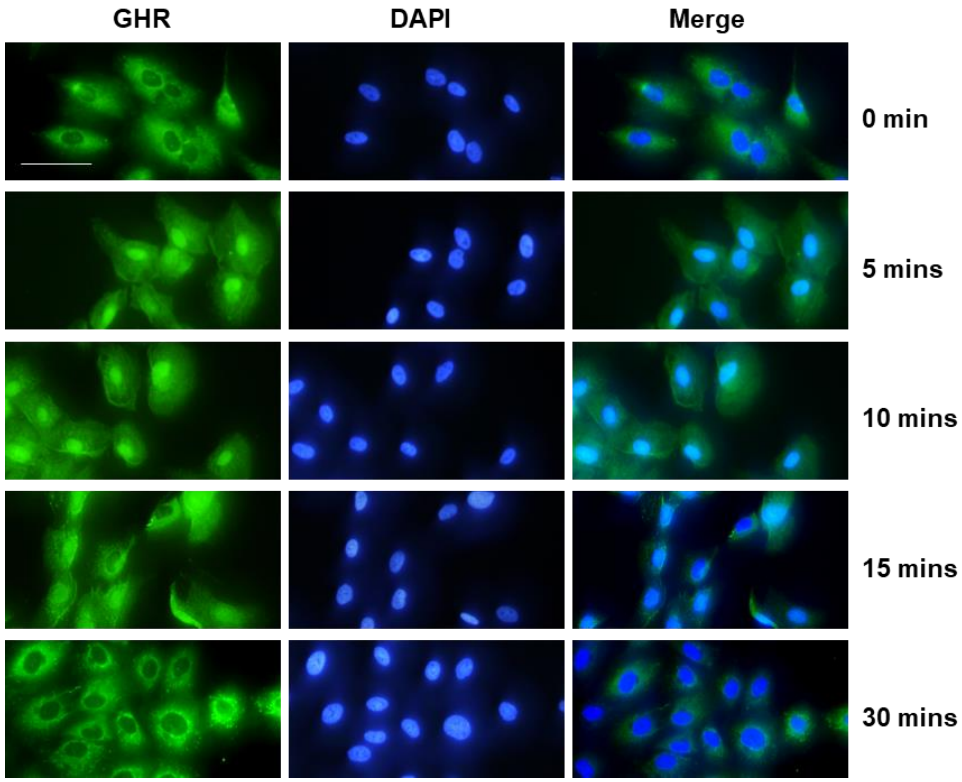

Figure 4.3. Nuclear localisation of the GHR in MCF-10A cells
MCF-10A cells were grown on coverslips, serum-starved overnight, and were treated with 500 ng/ml recombinant human GH for 0, 5, 10, 15, and 30 min. Cells were then fixed, permeabilised, blocked, and immuno-stained with the anti-GHR_{EC} antibody (Abcam, 89400) and the fluorescent secondary antibody. The slides were visualised using fluorescence microscopy. Green (alexa-fluor 488) represents GHR staining, and blue (DAPI) is a nuclear stain. The scale bar represents 100 µm.

Nuclear GHR translocation in RL95-2 cells was also confirmed by western blotting following nuclear fractionation. RL95-2 cells were treated with 500 ng/ml GH for 0, 5, 10, 15, and 30 min and the cytoplasmic and nuclear fractions isolated following cell lysis. GHR protein was detected in the nuclear fraction by western blotting analysis after 5 min treatment, and was not detected after 30 mins (Figure 4.3). In addition, reduced GHR protein levels were observed in the cytoplasmic fraction at 10 and 15 mins (Figure 4.3). Two isoforms of GHR (75 and 110 kDa) were detected using the anti-GHR_{IC} antibody. These GHR isoforms correspond to previously reported sizes for the glycosylated and unglycosylated forms of the receptor (Conte et al., 2002; van den Eijnden et al., 2006). Notably, only the 75 kDa, unglycosylated form of GHR, was identified in the nuclear fraction, using the anti-GHR_{IC} antibody (Figure 4.5C).

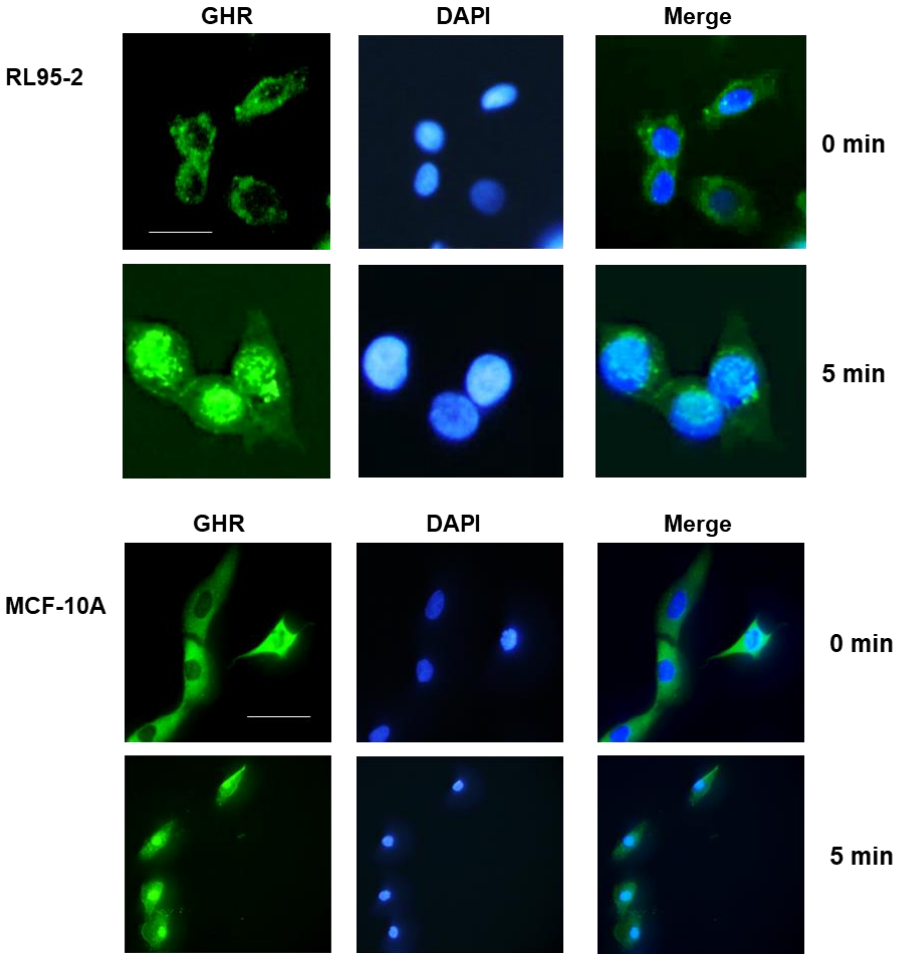

Figure 4.4. GHR nuclear localisation in RL95-2 and MCF-10A cells using the anti-GHR$_{IC}$ antibody
Cells were grown on coverslips then serum starved and were treated with 500 ng/ml recombinant human GH for 0 and 5 min. Cells were then fixed, permeabilised, blocked and immuno-stained with the anti-GHR$_{IC}$ antibody (sc-137185) and fluorescent secondary antibody. The slides were visualised using fluorescence microscopy. Green (alexa-fluor 488) represents GHR staining and blue (DAPI) is a nuclear stain. The scale bar represents 100 µm.

Collectively these results indicate that a 75 kDa GHR isoform rapidly translocates to the nucleus in RL95-2 and MCF-10A cells, and that nuclear localisation is short-lived, declining after 15 mins. Subsequent experiments focused on the RL95-2 cell-line, due to the higher levels of GHR expression that were observed in this cell-line.

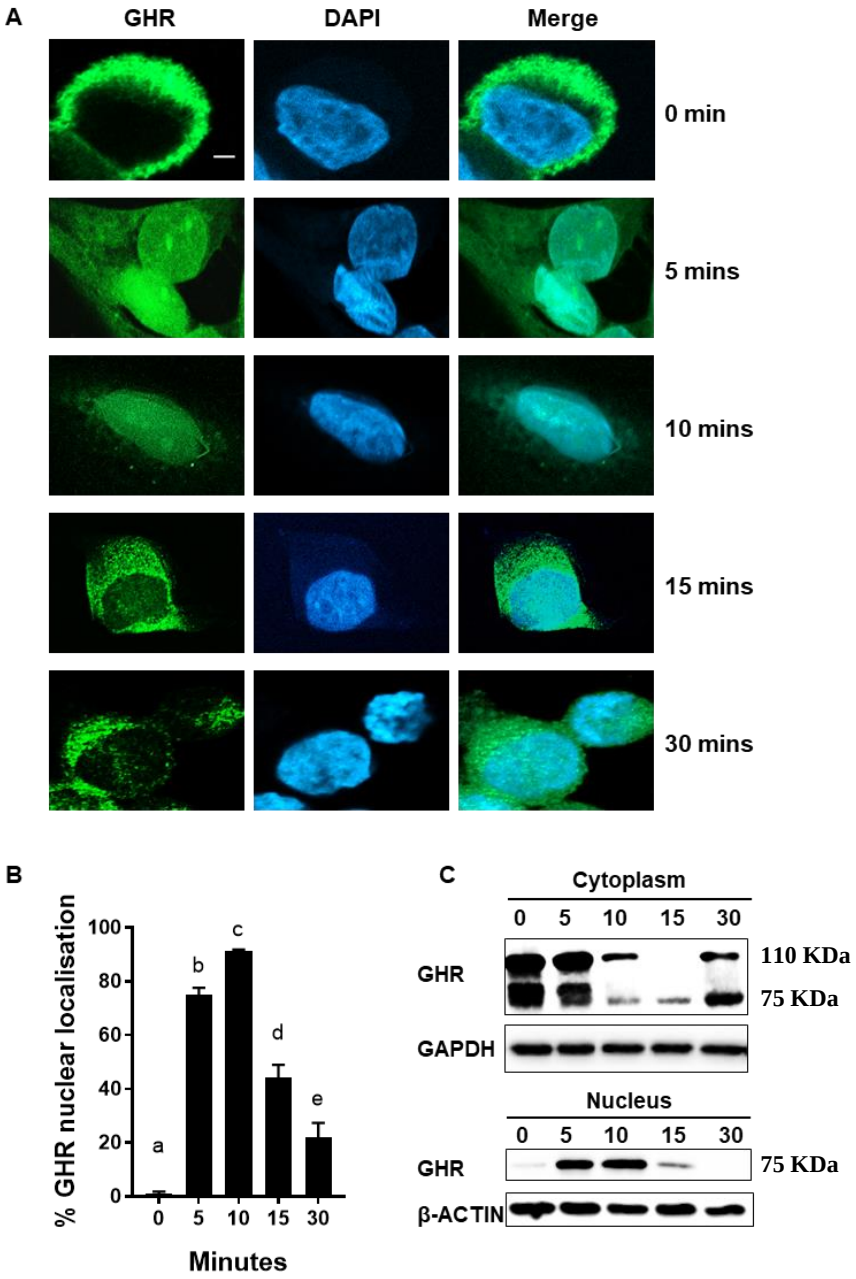

Figure 4.5. Subcellular localisation of GHR in RL95-2 cells following GH treatment
(A) RL95-2 cells were grown on coverslips, serum-starved and treated with 500 ng/ml recombinant human GH for 0, 5, 10, 15, and 30 min. Cells were then fixed, permeabilised, blocked, and immuno-stained with anti-GHR antibody and fluorescent secondary antibody. Slides were visualised using confocal laser scanning microscopy. Green colour (Alexa-Fluor 488) represents GHR staining, and blue colour (DAPI) is nuclear stain. The scale bar represents 100 µm. (B) Percentage of cells with nuclear GHR localisation at different time points. Approximately, 750 cells were counted at each time point to obtain this data. Data is presented as mean ± SEM. Groups that have different letters are significantly different from each other (P<0.001, One-way ANOVA). This figure is a representative of three individual experiments. (C) GHR time-course in nuclear and cytoplasmic fractions in RL95-2 cells. Serum-starved RL95-2 cells were treated with 500 ng/ml recombinant human GH for 0, 5, 10, 15, and 30 min. Cell lysates were fractionated into cytoplasm and nucleus, and immunoblotted with the anti-GHR_{EC} antibody (ab89400), with β-ACTIN as a loading control for nuclear samples, and GAPDH for cytoplasmic samples.

4.3.2 GHR interacts with known proteins in the nucleus

To determine whether the GHR interacts with proteins in the nucleus, immunoprecipitation and western blotting was carried out. Initially, proteins that are known to either interact or crosstalk with the GHR at the cell membrane were investigated. RL95-2 cells were treated with/without 500 ng/ml GH for 5 min, nuclear and cytoplasmic fractions were isolated, and proteins immunoprecipitated using a combination of an equal concentration of anti-GHR_{IC} and anti-GHR_{EC} antibodies (Figure 4.6). Eluates were then analysed by western blot.

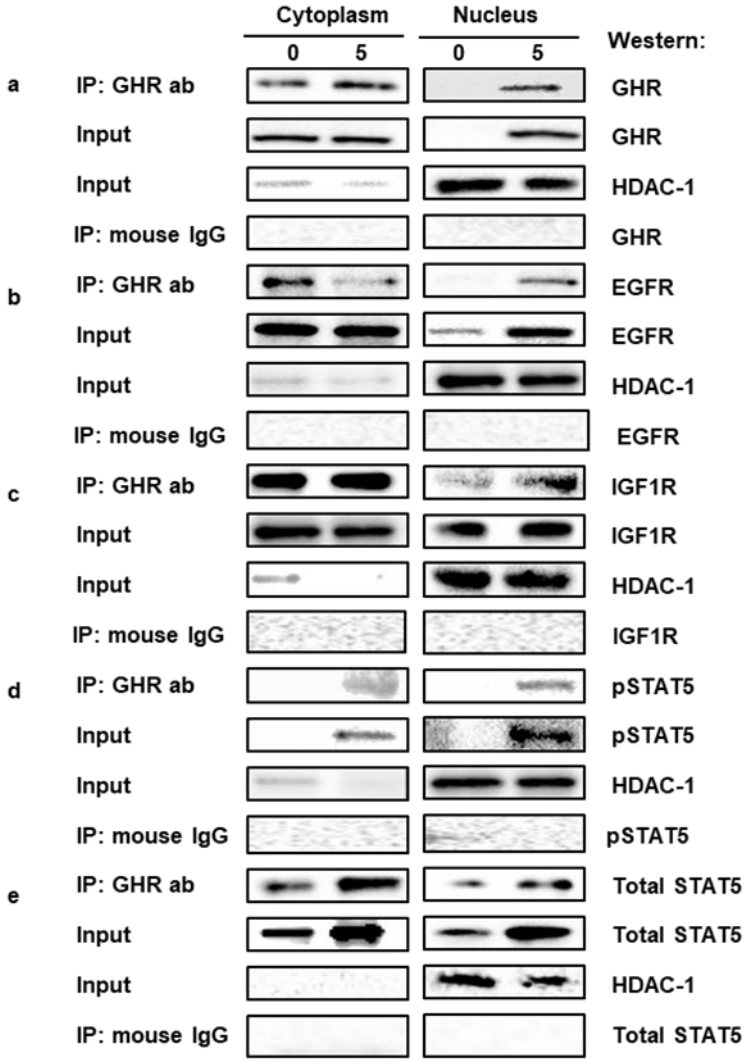

Figure 4.6. Co-immunoprecipitation of proteins associated with the GHR in cytoplasmic and nuclear extracts
Serum-starved RL95-2 cells were treated with 500 ng/ml recombinant human GH for 0 and 5 min. Cells were then lysed, and cytoplasmic and nuclear fractions isolated. Proteins were immunoprecipitated with the combination of anti-GHR_{EC} and anti-GHR_{IC} antibodies (ab89400 and sc-137185, respectively). Eluates were separated by SDS-PAGE and immunoblotted for different

proteins. (A) GHR, (B) EGFR, (C) IGF1R, (D) pSTAT5 and (E) Total STAT. HDAC-1 was used as a loading control for nuclear fractions. IP: immunoprecipitation**.**

First, RL95-2 nuclear and cytoplasmic protein lysates were western blotted for the GHR, to confirm that the receptor was enriched in the nucleus at 5 minutes after GH treatment (Figure 4.6A). Next, known GHR binding partners (*i.e.* EGFR, IGF1R, total STAT and phosphorylated STAT5) were investigated for nuclear localisation. Notably, all of these proteins were detected in the nuclear fraction after treatment with GH. Prior to immunoprecipitation, EGFR was found to be present in the cytoplasm and nucleus, with enrichment in the nuclear fraction observed following GH treatment (Figure 4.6B). In addition, EGFR co-immunoprecipitated with the GHR in both cytoplasmic and nuclear fractions using the combination of intracellular and extracellular anti-GHR antibodies. After GHR co-immunoprecipitation, decreased EGFR protein was observed in the cytoplasm following GH treatment, with enrichment in the nucleus. This suggests that GH treatment may result in increased nuclear localisation of EGFR, and that the GHR associates with EGFR in the nucleus.

In contrast, IGF1R appeared to be present in the cytoplasm and nucleus before and after GH treatment (5 min), and no obvious enrichment was observed following treatment. IGF1R co-immunoprecipitated with the GHR in the cytoplasm and nucleus, but seemed to be enriched in the nucleus only after stimulation with GH (Figure 4.6C). Phospho-STAT5 and total STAT were also found to be immunoprecipitated with GHR both in the cytoplasm and the nucleus, whereas co-immunoprecipitation of total STAT with GHR was GH-dependent only in the nucleus (Figure 4.6D and 4.6E). Phospho-STAT5 was enriched in the cytoplasm after treatment with GH, which correlates with GH/GHR signalling. However, it is not clear from these results if GHR translocates into the nucleus with these proteins bound to it, or it forms a complex with it in the nucleus. Mass spectrometry analysis of proteins that co-immunoprecipitate with the GHR

To identify novel nuclear GHR binding partners co-immunoprecipitation coupled with mass spectrometry was performed on RL95-2 cells treated with GH. RL95-2 cells (5×10^6) were serum-starved and treated with 500 ng/ml GH for 7 minutes. GHR pull-down was carried out on total cell lysates using the anti-GHR$_{IC}$ and anti-GHR$_{EC}$ antibodies, and was confirmed for both antibodies by immunoprecipitation and western blotting (Figure 4.6 and Figure 4.8).

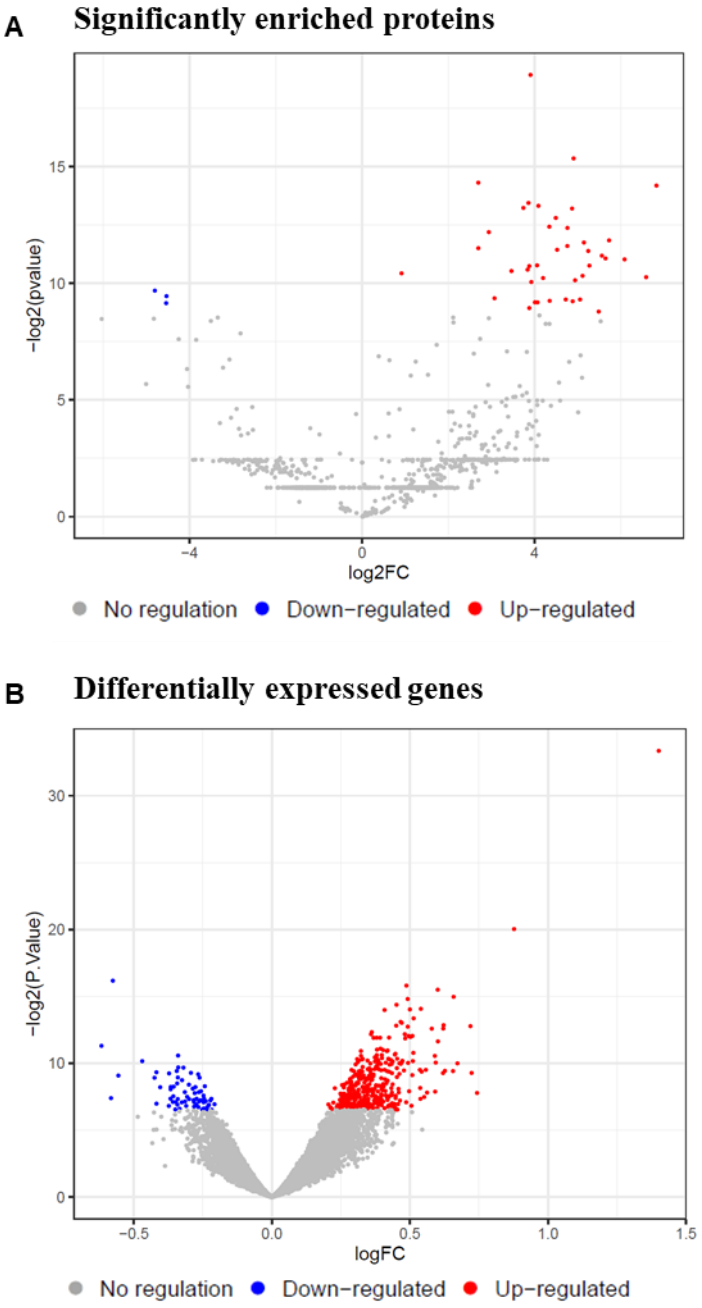

Figure 4.7. Significantly enriched proteins and genes following GH treatment
(A) Volcano plot showing proteins, identified by co-immunoprecipitation with a combination of anti-GHR_{EC} and anti-GHR_{IC} antibodies coupled with mass spectrometry, that were differentially expressed following treatment with 500 ng/ml GH for 5 mins compared to control. Blue dots represent proteins that are downregulated in treated samples compared to the control. Red dots represent proteins that are either upregulated compared to the control following GH treatment, or are present only in the GH-treated group. (B) Volcano plot representation of differentially expressed genes between control and treatment with 500 ng/ml GH for 90 minutes, obtained using a Clariom D microarray. Blue dots represent downregulated genes, whereas red dots represent significantly upregulated genes in the set.

For mass spectrometry, total cell lysates were split and immunoprecipitated separately with each antibody, then eluates were combined and processed for mass spectrometry. An anti-

mouse IgG antibody was used as a control. Analysis of these samples using mass spectrometry identified 44 significantly enriched proteins between control and treatment groups (P value <0.05, FDR <0.05) (Table 4.1; Figure 4.7A). There were 21 proteins which were present in both control and GH-treated groups, out of which 18 were upregulated and 3 were downregulated in the GH-treated cells. An additional 23 proteins were exclusively present in the GH-treatment group (Supplementary Tables 1-3).

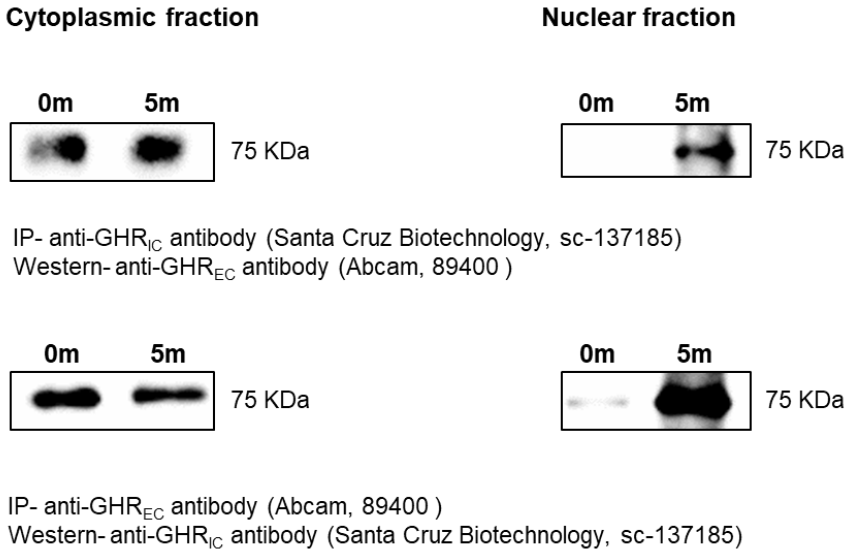

Figure 4.8. Validation of GHR antibodies
Serum-starved RL95-2 cells were treated with 500 ng/ml recombinant human GH for 0 and 5 min. Cell lysates were fractionated into cytoplasmic and nuclear fractions, immunoprecipitated, and immunoblotted for GHR using the anti-GHR$_{EC}$ or anti-GHR$_{IC}$ antibodies (ab89400 and sc-137185, respectively), as indicated.

Pathway analysis was conducted using the g:Profiler R package, to determine if the proteins were enriched in specific biological pathways. Results from this analysis demonstrated that the majority of these proteins were enriched in cellular processes, such as translation of proteins, ribosomal pathways, cytoskeletal organisation and trafficking, cell-cycle maintenance and, splicing and ubiquitination pathways (Supplementary Tables 4, 5 and 6). STRING analysis (Szklarczyk et al., 2018) was conducted to identify whether some of the proteins identified by mass spectrometry were enriched in a complexes. But this analysis did not yield any significant results.

4.3.3 Integration of microarray and mass spectrometry data identifies that GH increases mRNA levels of HMGN1 and SUMO1 gene targets

As the GHR translocates to the nucleus upon activation, the ChIP-Atlas database was used to investigate whether any of the proteins identified by mass spectrometry were transcription factors (Oki et al., 2018). This analysis identified 2 out of 44 proteins as transcription factors or transcription regulators. These were HMGN1 and SUMO1, which have 1145 and 1735 gene targets respectively. Next, it was investigated whether GH treatment led to changes in expression of HMGN1 and SUMO1 gene targets in RL95-2 cells. Cells were treated with/without 500 ng/mL GH for 90 min, and gene expression was analysed using Clariom D microarrays. This analysis identified 416 differentially expressed (DE) genes following GH stimulation (Figure 4.7B, Supplementary Table 7). Pathway analysis demonstrated that these genes were found to be significantly enriched in pathways such as VEGF/VEGFR, mRNA processing, TGF-β, adipogenesis, nuclear receptor, and EGF/EGFR signalling (Supplementary Tables 8-10).

The known gene targets of HMGN1 and SUMO1 were then intersected with the differentially expressed genes obtained from our microarray analysis. This analysis identified 45 genes in our dataset as being a target of HMGN1 and 124 genes as targets of SUMO1 (Supplementary Table 11). Bootstrapping of these values indicated that this intersection is not likely to be aleatory (*P* value < 0.001). In addition, 17 genes were identified which were targets of both HMGN1 and SUMO1 (Figure 4.9).

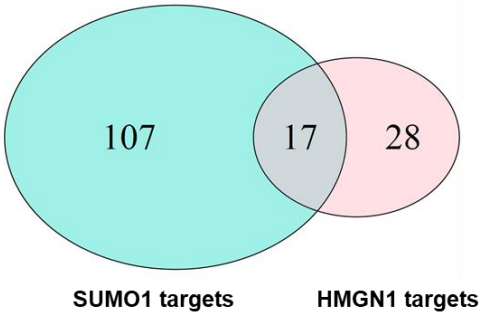

Figure 4.9. GHR nuclear localisation is associated with differential expression of HMGN1 and SUMO1 target genes

Venn diagram representing the intersection of known gene targets of HMGN1 and SUMO1, obtained from the ChIP-atlas database with differentially expressed genes from microarray analysis.

Protein	log2FC	SE	T value	P value	FDR
MBNL2	3.901149	0.005514	707.4399	2.00E-06	0.001838
AK2	4.898611	0.023969	204.3758	2.39E-05	0.011012
PSMD6	2.69112	0.018853	142.7415	4.91E-05	0.012203
H2BFS	6.818602	0.049932	136.5585	5.36E-05	0.012203
RCC2	3.853199	0.036491	105.5923	8.97E-05	0.012203
TRIP6	4.084511	0.040401	101.0999	9.78E-05	0.012203
LAMC1	3.735546	0.038108	98.02539	1.04E-04	0.012203
EIF5A	4.863538	0.050104	97.06798	1.06E-04	0.012203
HNRNPU	4.48702	0.053031	84.611	1.40E-04	0.014276
RALB	4.336918	0.058492	74.14498	0.000182	0.015763
SUMO1	4.756519	0.06531	72.83031	0.000188	0.015763
CAD	2.935254	0.042887	68.44213	0.000213	0.016361
LASP1	5.721248	0.094541	60.51588	0.000273	0.019081
TMPO	5.138264	0.087587	58.66484	0.00029	0.019081
CTSD	4.753834	0.085346	55.70074	0.000322	0.019081
RUVBL2	2.69139	0.049943	53.8896	0.000344	0.019081
KHSRP	4.518443	0.085827	52.64569	0.000361	0.019081
EEF1A1	5.238331	0.101242	51.74064	0.000373	0.019081
RCN1	5.553274	0.115133	48.23356	0.00043	0.0208
RPS27A	5.638342	0.121754	46.30923	0.000466	0.021007
PLP2	6.080029	0.133187	45.65019	0.00048	0.021007
ACO2	4.052897	0.097012	41.77743	0.000572	0.022425
CALM3	5.265623	0.12658	41.5992	0.000577	0.022425
CRIP2	3.871649	0.093684	41.32662	0.000585	0.022425
PLIN3	3.833868	0.097879	39.16932	0.000651	0.023962
MATR3	3.459674	0.090139	38.38151	0.000678	0.023995
ATP1A1	0.911469	0.024568	37.09965	0.000726	0.024729
EEF1D	5.107703	0.142805	35.76695	0.000781	0.025583
MYL6	6.57996	0.187728	35.05051	0.000813	0.025583
TOMM70	4.192039	0.121155	34.60065	0.000834	0.025583
TPM4	4.934182	0.147437	33.46629	0.000892	0.026462
DDX21	3.920215	0.120238	32.60367	0.000939	0.027008
RPS18	-4.80298	0.167505	-28.6736	0.001214	0.033847
ITIH3	-4.53357	0.171508	-26.4336	0.001428	0.038528
SEC31A	3.066999	0.119854	25.58955	0.001524	0.038528
CLTB	4.716849	0.187677	25.13285	0.001579	0.038528
PFDN2	5.049446	0.200972	25.12517	0.00158	0.038528
SHTN1	4.344467	0.176356	24.6346	0.001644	0.038528
SLC1A5	4.877745	0.19927	24.47803	0.001665	0.038528
RPN1	4.003379	0.16609	24.10362	0.001717	0.038528
FLNA	4.062703	0.168714	24.08042	0.00172	0.038528
HSPA8	-4.54099	0.190697	-23.8126	0.001759	0.038528
HMGN1	3.872981	0.175147	22.1128	0.002039	0.043622
RPL35	5.481249	0.261127	20.99077	0.002262	0.047294

Table 4.1. Significantly differentially expressed proteins between control and GH treated samples, identified by mass spectrometry analysis (FDR <0.05)

To further confirm the likelihood that GH treatment alters the transcriptional profile of HMGN1 and SUMO1 target genes, the RegulatorTrail database was used to predict potential transcription regulators from gene expression data. RegulatorTrail is a database that integrates transcriptomic and epigenomic data to identify transcriptional regulators involved in the regulation of a set of genes (Kehl et al., 2017). Analysis of the list of differentially expressed genes through this database generated a list of proteins that are likely involved in transcriptional regulation of these genes. Only two proteins from the mass spectrometry data were present in the list: HMGN1 and SUMO1. This suggests a strong likelihood that these two proteins are potentially involved in direct/indirect gene transcription following stimulation with GH and nuclear translocation of GHR.

As the co-immunoprecipitation mass spectrometry was carried out on whole cell lysates, co-immunoprecipitation and western blotting were performed on cytoplasmic and nuclear fractions to determine whether HMGN1 and SUMO1 were associated with the GHR in the cytoplasm or the nucleus. HMGN1 was localised exclusively in the nucleus of RL95-2 cells, and was found to be associated with the GHR following 5 min of GH treatment, as demonstrated by co-immunoprecipitation using a combination of equal amounts of anti-$GHR_{\mathbf{IC}}$ and anti-$GHR_{\mathbf{EC}}$ antibodies as bait (Figure 4.10A). Similar results were observed when using the anti-HMGN1 antibody as a bait, with GHR enriched in the nucleus following co-immunoprecipitation (Figure 4.10C). SUMO was detected in the cytoplasm and nucleus. However, no association between SUMO1 and GHR was observed in either anti-SUMO1 or anti-GHR immunoprecipitation (Figure 4.10B and C). Collectively, these results suggest that the GHR and HMGN1 are localised in a complex in the nucleus that is involved in GH-dependent coordinated regulation of GHR/HMGN1 gene target transcription. Whether the GHR is associated with SUMO in the nucleus remains to be confirmed.

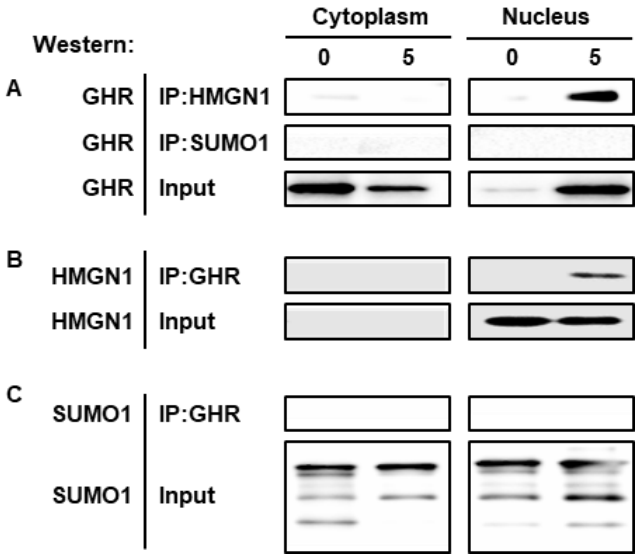

Figure 4.10. Validation of mass spectrometry data by co-immunoprecipitation followed by western blotting
Serum-starved RL95-2 cells were treated with 500 ng/ml recombinant human GH for 0 and 5 min. Cells were then lysed, and cytoplasmic and nuclear fractions isolated. Proteins were immunoprecipitated with a combination of anti-GHR_{EC} and anti-GHR_{IC} antibodies (ab89400 and sc-137185). Eluates were separated by SDS-PAGE and immunoblotted for different proteins. (A) HMGN1, (B) GHR and (C) SUMO1.

4.4 Discussion

Here, it is shown that the GHR rapidly translocates into the nucleus in RL95-2 cells following stimulation by GH, and co-localises with a transcription regulator, HMGN1, in the nucleus. This co-localisation correlates with the expression of known gene targets of HMGN1 in this cell-line. Nuclear translocation is not unique to the GHR; over the past few decades it has been demonstrated that multiple cell-surface receptors are located in the nucleus following a suitable stimulus. Receptors such as the GHR, EGFR, IGF1R, and PRLR to name a few, previously believed to exclusively signal at the cell-surface have now been shown to localise in the nucleus (Bryant and Stow, 2005; Conway-Campbell et al., 2007; Lobie et al., 1994a; Wang and Hung, 2009). Of these, nuclear translocation of the EGFR is the best characterised (Wang and Hung, 2009). Several studies have confirmed the presence of GHR in the cell nuclei of cell-lines from different species, with translocation often occurring within 5 to 10 minutes following GH stimulation (Conway-Campbell et al., 2007; Figueiredo et al., 2016; Lan et al., 2017; Lobie et al., 1994a). In accordance with these studies, rapid nuclear

translocation of the GHR was observed in RL95-2 and MCF-10A cells within 5 min of GH treatment, with maximal localisation at 10 min.

Immunoprecipitation using anti-GHR antibodies, combined with mass spectrometry, identified 44 protein-binding partners of the GHR in RL95-2 cells. Pathway analysis identified an enrichment in eukaryotic translation, ribosomes, cellular trafficking, cell-cycle, and splicing pathways. Our results also identified increased association of the GHR with proteins involved in trafficking following GH treatment. For example, LASP1, FLNA, PLIN3, CLTB, TRIP6, and SEC31A are important in regulation of the cytoskeleton, migration, and membrane trafficking (Galvez et al., 2012; Hirst et al., 2015; Orth et al., 2015; Savinko et al., 2018; Shibata et al., 2015; Willier et al., 2011). The activity of some of these proteins is also a key factor regulating the invasiveness of cancer cells. For example, LASP1 is an actin-binding structural protein that is upregulated in various cancers and promotes cancer cell proliferation, migration, and invasion (Hu et al., 2018). Nuclear localisation of LASP1 has previously been reported, and is associated with a reduced overall survival rate for patients with invasive breast cancer (Frietsch et al., 2010). TRIP6 was one of the proteins that only associated with the GHR after GH treatment. This protein is also implicated in cancer and regulates numerous cellular responses, such as actin cytoskeletal reorganisation, cell adhesion, and cell mobility (Willier et al., 2011). In addition, TRIP6 shuttles to the nucleus, where it acts as a transcriptional co-regulator. Mass spectrometry analysis also identified increased association of the GHR with proteins involved in ubiquitination following GH treatment. SUMO1, RPS27A, PSMD6, and TOMM70 are proteins involved in ubiquitination/ de-ubiquitination pathways. This is consistent with the significant role that the ubiquitin system plays in internalisation of GHR (Sachse et al., 2001). However, the complete picture is still unclear; the enrichment of numerous cytoskeletal organisation, cellular trafficking, and ubiquitin system-related proteins bound to nuclear GHR may support the import of GHR into the nucleus.

EGFR functions as a transcription factor in the nucleus (Brand et al., 2011) which led to the hypothesis that one of the functions of GHR nuclear localisation might be transcriptional regulation, which could either be through direct DNA binding, or by sequestration of other transcription factors and/or regulatory proteins. GHR is also known to cross-talk with the EGFR at the cell surface and that was supported by co-immunoprecipitation-western (Kostopoulou et al., 2017; Li et al., 2008; Qi et al., 2016). Intriguingly, it was also demonstrated in our study that a GHR-EGFR association occurs in the nucleus following GH

treatment. However, although EGFR was detected in the co-immunoprecipitation-mass spectrometry analysis following GH treatment, the association was not significant. In addition, other known binding partners, such as IGF1R and STAT5 were detected by co-immunoprecipitation-western after GH treatment but were not observed in the co-immunoprecipitation-mass spectrometry results. The reason is not clear, but it could either be because of the weak nature of the interaction, which led to the protein(s) being washed away during processing, or due to the lower detection threshold of the mass spectrometric method.

Importantly, our results indicated that the GHR is associated with two transcription regulators, HMGN1 and SUMO1, in the nucleus in response to GH treatment. An association between the GHR and these two transcription regulators has not been shown previously. HMGN1 is a transcription regulator that alters DNA helix:histone octamer interactions. Changing this interaction affects transcription of genes by altering chromatin conformations (Bannister and Kouzarides, 2011). SUMO1 is a member of the ubiquitin-like protein family, which is involved in multiple essential cellular functions including transcription, nuclear transport, DNA replication and repair, cell-cycle, and signalling. Although SUMO1 is better-known for its function as a post-translational modifier and does not possess specific DNA binding sequences, direct interaction of SUMO1 with DNA has been shown in some studies by using NMR spectroscopy and protein-DNA cross-linking experiments. Additionally, it has also been demonstrated that SUMO1 can cause an upregulation of the enzymatic turnover of thymine-DNA glycosylase by directly binding to the DNA regulatory domain of this enzyme in a sequence-independent manner (Eilebrecht et al., 2010; Ke et al., 2019; Priyanka et al., 2016; Smet-Nocca et al., 2011). Intersection of the HMGN1 and SUMO1 gene targets with genes that were found to be differentially expressed genes in response to GH treatment, identified 45 out of 416 differentially expressed genes that intersected with known gene targets of HMGN1, and 124/416 genes that intersected with SUMO1 gene targets. This was consistent with RegulatorTrail predictions (Kehl et al., 2017), which identified HMGN1 and SUMO1 as regulators of differential expression in our expression data. An association between the GHR and HMGN1 in the nucleus was confirmed by immunoprecipitation. However, no association between GHR and SUMO1 was observed, in either the nucleus or cytoplasm, which could possibly signify that GHR indirectly activates SUMO1 by interaction with other proteins present in the same complex, or in close vicinity. Collectively, our data are consistent with GHR being associated with HMGN1 in the nucleus following GH

treatment. This may then lead to regulation of gene transcription by HMGN1, either by direct DNA binding, or by binding other DNA binding transcription factors.

Using affinity chromatography and tandem mass spectroscopy, Conway-Campbell *et al.* identified that the GHR can bind to a transcriptional regulator, coactivator activator protein (CoAA), through the extracellular domain of the GHR, in a GH-dependent manner (Conway-Campbell et al., 2008). CoAA is a potent nuclear receptor coactivator protein that is found to be overexpressed in a wide range of cancers (Kai, 2016; Sui et al., 2007), and the authors suggest that CoAA contributes to the proliferative actions of nuclear GHR. Further, CoAA was not observed in our mass spectrometry dataset, which could possibly be due to the transient and cell-type specific nature of the protein-protein interactions, or the lower level of this protein in the cell. However, Conway-Campbell *et al.* also identified the translational regulator, EF1α, which is a specific GHBP-interacting protein and may be involved in GHR/GHBP nuclear import (Conway-Campbell et al., 2008). This regulator was also identified in our study, with co-immunoprecipitation-mass spectrometry.

Our data is consistent with an alternative mechanism through which GH mediates nuclear localisation of the GHR and interactions with transcription regulators, such as HMGN1 and SUMO1. This interaction could manifest in a few possible scenarios following GH stimulation. Firstly, the GHR could interact with one or both of these regulators in the cytoplasm and then translocate into the nucleus as a complex. Alternatively, the interaction may occur after they travel independently into the nucleus. In the nucleus, HMGN1 and SUMO1 may either bind to DNA directly while complexed with GHR, or these proteins may activate other transcription factors which are then involved in gene expression regulation, or the complex may dissociate in the nucleus allowing these factors to interact with DNA or other transcriptional regulators independently. Another possible scenario could involve interaction of GHR with other proteins leading to activation of HMGN1 and/or SUMO1 and subsequent regulation of transcription. Further exploration using cell-based reporter assays, knockout studies, and chromatin immunoprecipitation assays will significantly add to our understanding of the complex interactions between GHR nuclear localisation and regulation of gene expression.

Similarities can be drawn with other cell receptors. IGF1R interaction with SUMO-1 is crucial for its nuclear translocation (Sehat et al., 2010). Nuclear IGF1R has been reported to have a significant impact on tumorigenesis and DNA damage response (Poreba and

Durzynska, 2020; Simpson et al., 2017). Therefore, it is important to note that SUMOylation of IGF1R is linked to increased cellular proliferation in acute myeloid leukaemia, and that the SUMO1-IGF1R complex also mediates changes in cell cycle control (Lin et al., 2017; Zhang et al., 2015a). This highlights the importance of interaction of SUMO1 with IGF1R in the context of cell cycle regulation and proliferation, which could lead to pathologic effects. It is possible that SUMO1 might perform a similar role in GHR nuclear import, with implications in disease states. Our immunoprecipitation results indicate that GHR interacts with IGF1R in the cytoplasm, as well as in the nucleus. It has also been demonstrated previously that stimulation with GH results in the formation of a GHR-JAK2-IGF1R complex (Huang et al., 2004). Huang *et al.* also reported that a complex between GHR and soluble IGF1R dulls GH signalling in a human prostate cancer cell-line (LNCaP), as well as in mouse osteoblast cells (Gan et al., 2014). Hence, it is possible that the GHR complexes with IGF1R and localises to the nucleus following IGF1R-SUMO1 interaction.

In conclusion, this study demonstrates that GHR translocates into the nucleus in response to GH treatment in the RL95-2 cell-line. It was further shown that GHR co-localises with HMGN1 protein in the nucleus after GH treatment (5 min) and this co-localisation correlates with the differential expression of HMGN1 gene targets in this cell-line. Since HMGN1 is a transcription regulator that exerts an impact on chromatin conformation by altering the DNA-histone interaction, this study further suggests that GH may have a role in regulation of spatial genomic interactions resulting in transcriptional changes. There are potential significant clinical applications of understanding the consequences of nuclear import of GH receptor in different cell models. Since there is a strong correlation between GHR in nucleus and increased carcinogenesis, prevention of receptor nuclear import may be a potential therapeutic adjunct to already existing treatments. Further studies investigating the importance of GHR nuclear receptor import and interaction with these transcription regulators are warranted.

CHAPTER 5

Transcriptome Analysis and Construction of mRNA-miRNA-lncRNA Triple Network Following a Growth Hormone Treatment Time-Course in the MCF-10A Cell-Line

This chapter has been formatted for publication

5.1 Introduction

Cellular processes such as development, cell differentiation, and altered susceptibility to disease, are all associated with the coordination of expression profiles of coding and non-coding genes (Li and Liu, 2019; Nam et al., 2016). Alterations in these coordinated processes are now accepted as contributing to changes in transcriptional profiles that may lead to disease. Importantly, genetic changes (both coding and non-coding) that alter gene expression contribute to these disease risks (Esteller, 2011).

Non-coding RNAs can broadly be defined as transcripts that do not translate into proteins. This large class of RNAs was initially considered as "junk" RNA, which made it difficult to assign biological relevance. Although the true extent of the mammalian transcriptome remains to be determined, it is now widely accepted that a large proportion of non-coding RNAs have critical roles in many cellular functions (Deveson et al., 2017; Frith et al., 2005; Shabalina and Spiridonov, 2004). The non-coding genome comprises a variety of different types of RNAs, including short RNAs, such as microRNAs (miRNA), which are generally 19–22 nucleotides long; and longer RNAs, such as long non-coding RNAs (lncRNA), which can extend up to hundreds of bases (Bartel, 2004; Ulitsky and Bartel, 2013). Biological functions of these non-coding RNAs range from changes in chromatin conformation, epigenetic changes, and genomic imprinting, to transcriptional alterations (Birney et al., 2007; Gupta et al., 2010; Hanly et al., 2018; Loda and Heard, 2019). Recent studies have also identified a subset of lncRNAs that are translated into peptides (Housman and Ulitsky, 2016; Li and Liu, 2019).

Growth hormone (GH) plays an important role in multiple aspects of growth and metabolism (Bonert and Melmed, 2017; Chia, 2014; Lu et al., 2019). This includes roles in hepatic metabolism, immune function, bone and skeletal muscle growth, neurogenesis, lipolysis, insulin sensitivity, and reproduction (Bonert and Melmed, 2017; Devesa et al., 2016). Compromised GH signalling contributes to growth disorders and cancer, and has been implicated in microvascular complications associated with diabetes (Hannon et al., 2017; Perry et al., 2008, 2013).

The effects of GH on cell growth and differentiation are classically mediated through an interaction with a cell-surface GH receptor (GHR) (Waters, 2016). Binding of GH to the GHR activates signal transduction pathways critical for cell growth and survival, including

the JAK2/STAT, MAP Kinase, and PI3K/AKT pathways (Carter-Su et al., 2016; Lu et al., 2019). As described in Chapter 4, GH actions may also be mediated through nuclear translocation of the GHR (Conway-Campbell et al., 2007). The key downstream effects of GH occur through coordinated changes in gene transcription profiles, in particular altered mRNA transcription (Rotwein and Chia, 2010). This occurs through direct stimulation of transcription (e.g. through activation of the transcription factor STAT5) and through regulation of epigenetic modifications such as DNA methylation (Álvarez-Nava and Lanes, 2017; Carter-Su et al., 2016; Chia, 2014). GH-mediated changes in non-coding RNA transcription (miRNA and lncRNA) have also been increasingly reported (Chang et al., 2016; Melia et al., 2015; Perry et al., 2013; Zhang et al., 2015b). Furthermore, miRNAs and lncRNAs that regulate expression of GH or the GHR, and secretion of GH, have been identified (Du et al., 2019). Given the genetic intricacy and physiological importance of GH-mediated functions, it is likely that non-coding RNAs play a significant role in GH function and GH-linked disease progression. However, little is known regarding GH-regulated miRNA and lncRNAs in cell-lines.

The competitive endogenous RNA (ceRNA) hypothesis states that lncRNA, among other endogenous RNAs, can sequester miRNA by competing for shared binding sites resulting in altered expression of miRNA target genes (Salmena et al., 2011; Szczešniakjhk and Makałowska, 2016; Tay et al., 2014). Coordinated regulation of mRNA-miRNA-lncRNA networks has been shown to play a critical role in transcription, mRNA turnover and translation into proteins (Martirosyan et al., 2017a, 2017b; Miotto et al., 2019). Dysregulated ceRNA networks in disease states suggest an alternative gene regulatory pathway that potentially serves as a novel therapeutic recourse (Chen et al., 2017; Karreth and Pandolfi, 2013; Qi et al., 2015).

The present study was designed to evaluate coding and non-coding RNA expression profiles following GH treatment to elucidate a GH-induced ceRNA network and investigate the molecular mechanisms underlying GH actions. A time-course gene expression study was performed in the human mammary epithelial cell-line (MCF-10A) to identify differentially expressed transcripts following treatment with GH. Analysis of the correlation between coding and non-coding RNAs identified a coordinated mRNA-miRNA-lncRNA network. Our study identified a molecular ceRNA mechanism that could potentially be involved in modulation of GH activity.

5.2 Materials and Methods

5.2.1 Cell culture

Human MCF-10A cells were purchased from the American Type Culture Collection (ATCC). MCF-10A is a transformed, but otherwise normal mammary epithelial cell-line. Cells were maintained in a humidified chamber with 5% CO_2 at 37°C, in MEGM™ Mammary Epithelial Cell Growth Medium BulletKit™ (Lonza) media, supplemented with 100 µg/ml streptomycin (Sigma-Aldrich, St. Louis, MO, USA), 100 U/ml penicillin (Sigma-Aldrich) and 5% horse serum (Sigma-Aldrich). Cells in the logarithmic phase of growth were used for microarray analyses.

5.2.2 Cell treatment, RNA extraction and microarray analysis

MCF-10A cells were plated in triplicate at 5 x 10^6 cells in 10 cm cell culture dishes, serum-starved overnight, and then treated with 250 ng/mL recombinant GH (National Hormone and Pituitary Program, NHPP) for 0, 30, 90, 180 and 360 min. Total RNA was extracted using TRIzol reagent (Life Technologies) and column-purified using the RNeasy mini kit (Qiagen) according to the manufacturer's instructions. The purity and concentration of extracted RNA was determined by OD_{260}/OD_{280}, using a NanoDrop ND-1000 instrument (ThermoFisher Scientific). A Bioanalyser RNA 6000 Nano LabChip kit (Agilent) was used to assess the integrity of extracted RNA. Auckland Genomics (University of Auckland) performed the microarray analysis using the ClariomD microarray system (ThermoFisher Scientific).

5.2.3 Microarray data analysis

Raw .CEL files were processed and normalised (with the robust multi-array average method) using the R software packages (R version 3.4.4). Differentially expressed mRNA, miRNA, and lncRNA between control and treatment groups across different time points were determined using R packages maEndToEnd (Baszczyński and Goldstein, 1967) and limma (Ritchie et al., 2015). Significantly differentially expressed transcripts were estimated by applying an FDR *P*-value cut-off of < 0.05. Differentially expressed coding and non-coding RNAs were visualised by generation of heatmaps using the gplots package in R (Warnes et

al., 2020). A step-by-step description of the process involved in this analysis is provided in Section 2.3 of this thesis.

5.2.4 Pathway analysis

To identify significant pathways among the differentially expressed mRNAs, pathway analysis was performed using g:Profiler package in R (Raudvere et al., 2019).The threshold for identification as a significant pathway was an FDR adjusted *P*-value < 0.05. Significantly enriched pathways were represented in a dot plot, which was created in R, using the ggplot2 package (Gómez-Rubio, 2017).

5.2.5 Gene correlation analysis and construction of an mRNA-miRNA-lncRNA triple network

Gene expression correlation networks were constructed to identify interactions between coding and non-coding RNAs. Pearson's correlation coefficients (PCCs) were used to calculate the correlation between differentially expressed mRNA–miRNA and lncRNA-miRNA pairs. To facilitate visual representation of the results, only the correlations of miRNA with mRNA and lncRNA targets that intersected with differentially expressed mRNA and lncRNA were used to construct an mRNA-miRNA-lncRNA triple expression network. Cytoscape software (v2.8.3) was used to draw the mRNA–miRNA–lncRNA gene co-expression networks (Franz et al., 2016). Gene and lncRNA targets of miRNA were extracted from the miRNet database, version 2.0 (Chang et al., 2020). The line chart and UpSet plot were created in R, using the ggplot2 package and the ComplexHeatmap package, respectively (Gómez-Rubio, 2017; Gu et al., 2016).

5.3 Results

5.3.1 Microarray analysis identified significantly differentially expressed genes, miRNA and lncRNA

To understand whether GH treatment leads to changes in coordinated expression of coding and non-coding RNA in the MCF-10A cell-line, cells were treated with/without 250 ng/mL

GH for 0, 30, 90, 180 and 360 min, and gene expression was analysed using the Clariom D microarray system. This concentration of GH has previously been shown to

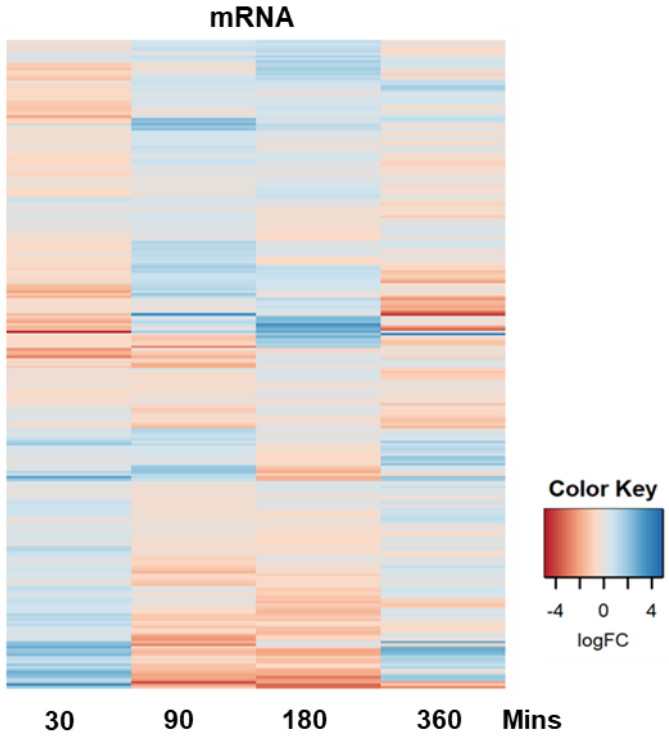

Figure 5.1. Differentially expressed miRNAs and lncRNAs over the time-course
Heatmap representing (A) miRNAs and (B) lncRNAs that significantly changed over the time-course (adjusted *P* value <0.05). MCF-10A cells were serum-starved overnight, then treated with 250 ng/ml GH for 30, 90,180 and 360 min. Blue represents a positive log fold change (logFC) and the red represents a negative log fold change.

stimulate optimally phosphorylation of the downstream signalling molecule, STAT5, in MCF-10A cells (Chapter 4). Changes in expression at each time point, expressed as log2 fold-change (LogFC), were small in absolute terms, with LogFC changes ranging from 0.03 (fold change= ~1) to 1.42 (fold change= ~2.7). However, when changes in expression were analysed across the time-course, significant changes in GH-induced transcription profiles were observed. For example, expression for a given gene may decrease at 30 min with GH treatment, but increase at 90 min and 180 min, then decrease again at 360 min. Analysis of the dataset with maEndToEnd and limma provides a robust way to identify these changes in expression profiles. Analysis through this workflow identified 259 mRNAs, 38 miRNAs, and 23 lncRNAs that were significantly differentially expressed across the time-course following GH treatment (adjusted *P*-value cut off <0.05) (Figures 5.1 and 5.2; Supplementary Table S1-S3).

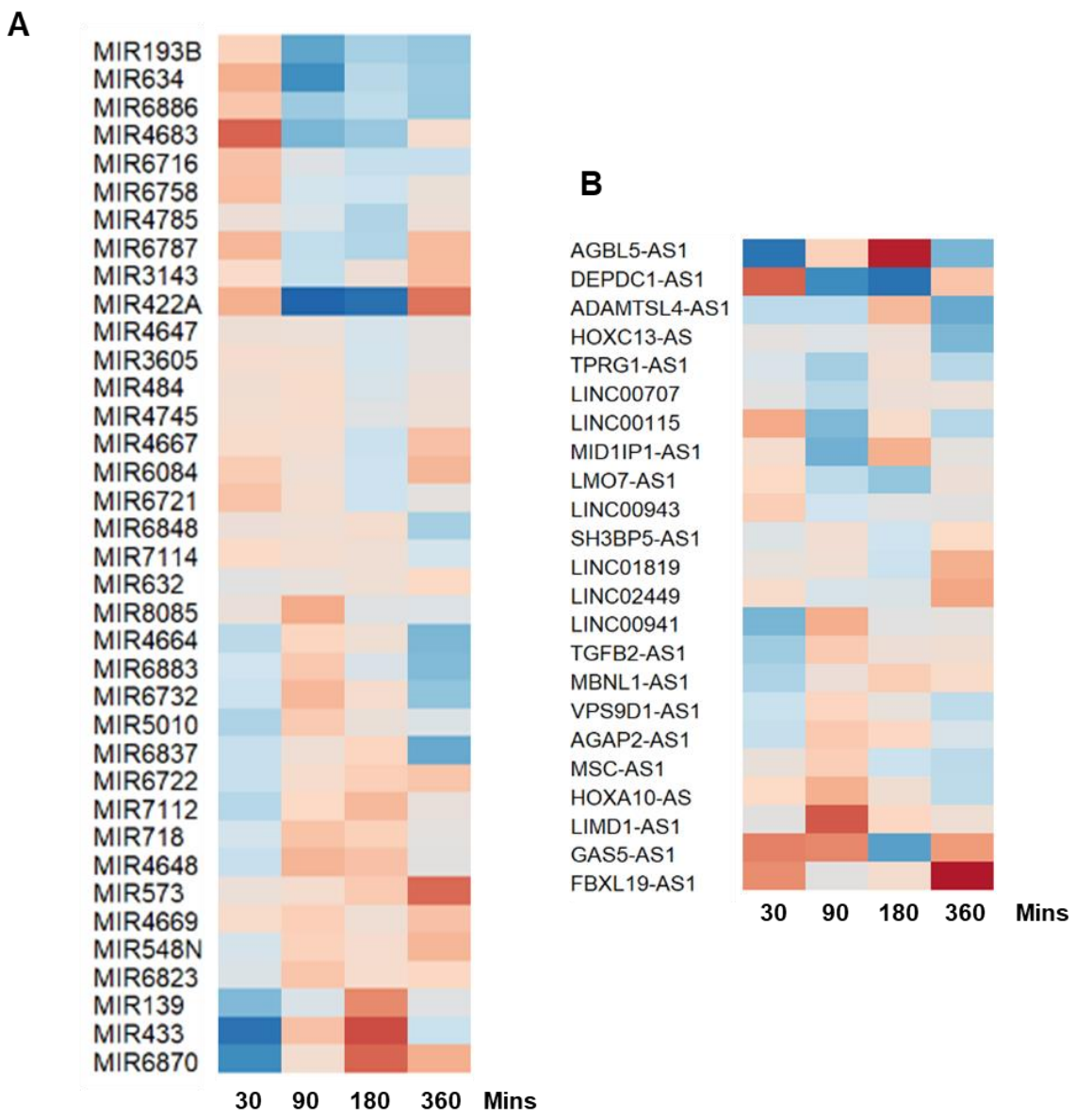

Figure 5.2. Differentially expressed miRNAs and lncRNAs over the time-course
Heatmap representing (A) miRNAs and (B) lncRNAs that significantly changed over the time-course (adjusted *P* value <0.05). MCF-10A cells were serum-starved overnight, then treated with 250 ng/ml GH for 30, 90, 180, and 360 min. Blue represents a positive log fold change (logFC), and the red represents a negative log fold change.

5.3.2 Functional enrichment analysis of significantly differentially expressed genes identifies known GH-related signalling pathways

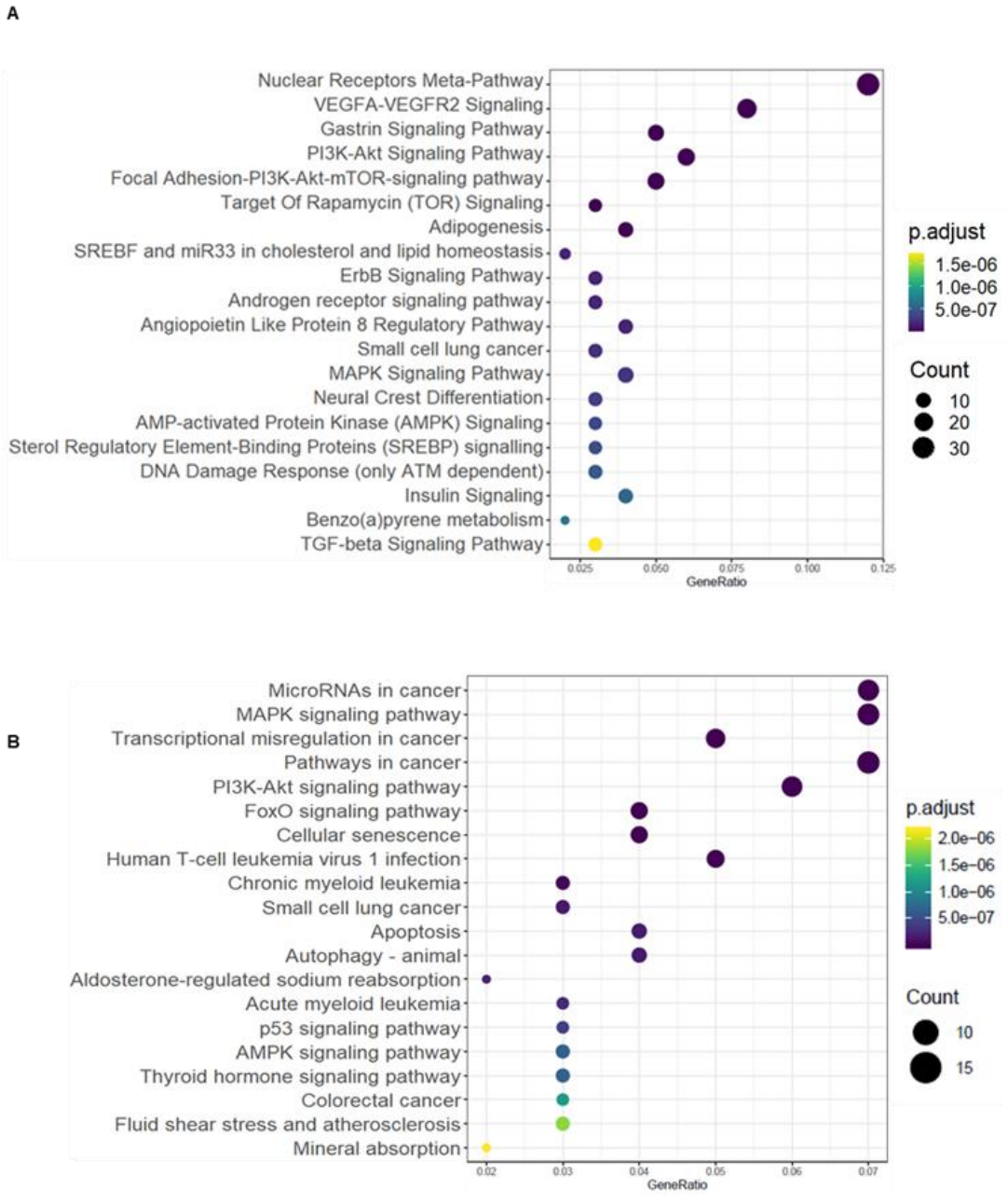

Figure 5.3. Pathway enrichment analysis of significantly differentially expressed genes following GH treatment in MCF10A cells
Dot plots showing top 20 pathways enriched by (A) WIKIPATHWAYS (B) KEGG using g:Profiler. The plot is presented in decreasing order of adjusted *P* values. The size of the circle represents the number of genes enriched in the pathways.

To determine whether differentially expressed mRNAsacross the time-course were enriched in signalling pathways, pathway analysis was performed using gProfiler restricted to KEGG and Wikipathways (Figure 5.3A and B; Supplementary Tables S4 and S5). This analysis showed that the differentially expressed genes were found to be significantly enriched in MAPK, VEGFA/VEGFR, ErbB, and PI3K-AKT signalling pathways. GH is known to mediate its cellular functions through these key signalling cascades.

5.3.3 Correlation analysis between differentially expressed mRNA, miRNA and lncRNA identifies several regulatory clusters which potentially mediate GH function

Previous studies have shown that miRNAs and lncRNAs can interact with each other and regulate the expression of mRNAs, resulting in coordinated regulation. To determine the correlation between differentially expressed coding and non-coding RNAs following GH treatment, a mRNA-miRNA-lncRNA ceRNA network was established. Pearson correlation coefficient values were calculated (Supplementary Table 6). To identify clusters of correlated mRNA, miRNA, and lncRNA, validated mRNA and lncRNA targets of differentially expressed miRNA (identified by microarray analysis) were extracted from the miRNet database (version 2.0). There were 1606 mRNA targets and 218 lncRNA targets of the differentially expressed miRNA. To connect these targets with our microarray analysis, this list was intersected with differentially expressed mRNA and lncRNA from the microarray dataset, which identified 83 mRNA targets and 7 lncRNA targets. Bootstrapping (n= 1000 iterations) confirmed that this intersection was more than expected by chance ($P < 0.001$).

Functional assessment was further refined by restricting the data to include only differentially expressed mRNA, miRNA, and lncRNA that were identified to be the targets of differentially expressed miRNA. Correlation between these RNAs was used to generate an mRNA-miRNA-lncRNA triple correlation ceRNA network using cytoscape (Figure 5.4). This

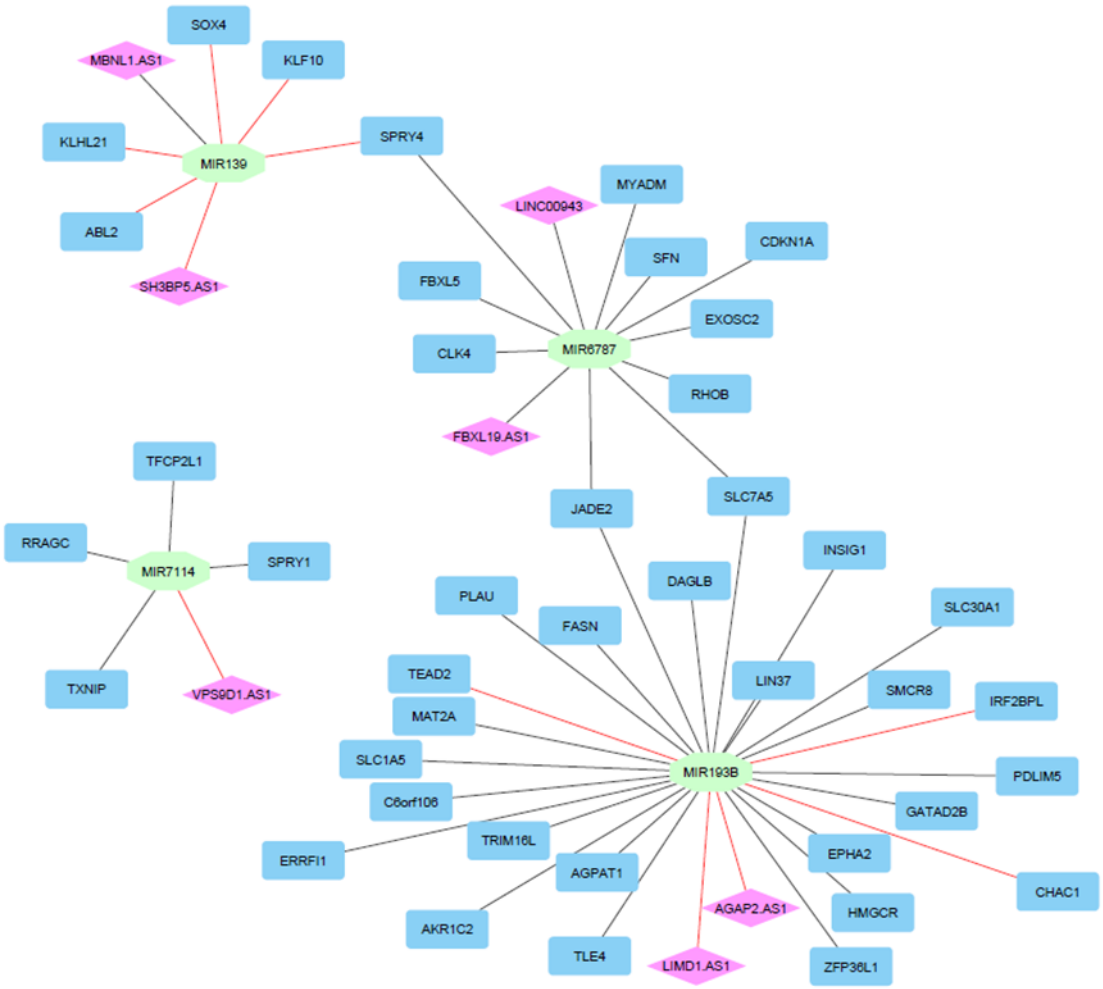

Figure 5.4. : LncRNA-miRNA-mRNA network competing endogenous RNA network
Green octagon nodes represent miRNA, blue rectangle nodes are mRNA and pink diamond nodes represent lncRNA. Red lines represent a negative correlation and black lines represent a positive correlation

network highlighted four distinct coding/non-coding gene clusters. Next, only the strongly correlated RNAs (>0.90) were selected to analyse the key regulatory coding and non-coding hubs in the network. These results demonstrated that there were three distinct clusters, one of which includes mRNA, miRNA, and lncRNA (Figure 5.5).

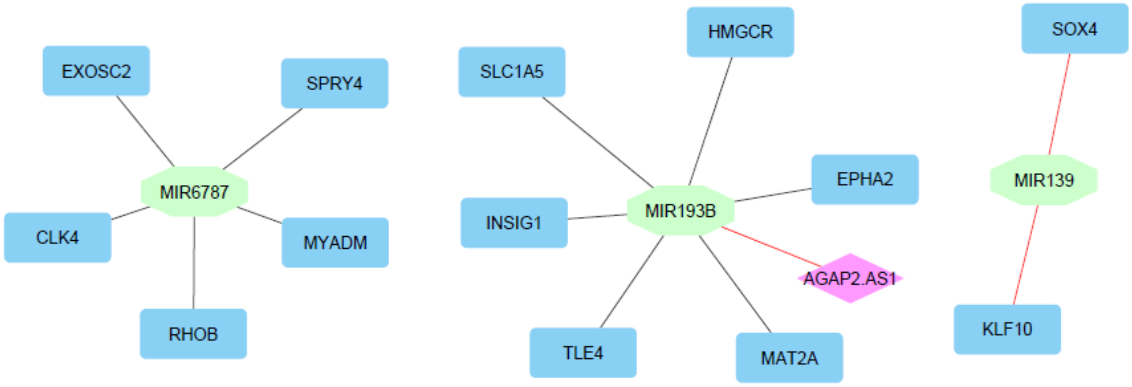

Figure 5.5. Filtered lncRNA-miRNA-mRNA network
lncRNA-miRNA-mRNA filtered for nodes with a correlation value >0.9. Green octagon nodes represent miRNA, blue rectangle nodes are mRNA and pink diamond nodes represent lncRNA. Red lines represent a negative correlation and black lines represent a positive correlation.

In the largest of the three clusters identified, *MIR193B* was positively correlated with six mRNAs: methionine adenosyltransferase 2A (*MAT2A)*, solute carrier family 1 member 5 (*SLC1A5)*, transducin-like enhancer protein 4 (*TLE4)*, ephrin receptor A2 (*EPHA2)*, 3-hydroxy-3-methylglutaryl-coA reductase (*HMGCR)*, and insulin-induced gene 1 (*INSIG1*). It negatively correlated, meanwhile, with the long non-coding RNA: ArfGAP with GTPase domain, ankyrin repeat and PH domain 2- antisense 1 (*AGAP2-AS1)*. *AGAP2-AS1* expression increased at 30 min, but decreased at 90 min, before increasing again at 180 and 360 min following GH treatment. In contrast, reciprocal trends were observed for all the other RNAs present in this cluster (Figure 5.6). Strikingly, pathway analysis indicated that the genes in this cluster were enriched in different GH-related metabolic pathways, such as lipid metabolism (*HMGCR* and *INSIG1*), glutamine metabolism (*TLE4*, *EPHA2*, and *SLC1A5*), and methionine metabolism (*MAT2A*) (Figure 5.7).

5.4 Discussion

This study integrated microarray expression data at different time points following GH treatment in the mammary epithelial cell-line, MCF-10A, to understand the transcriptome level regulation of coding and non-coding RNAs. Constructing a triple correlation ceRNA network comprising mRNA, miRNA and lncRNA highlighted the subtle effects of hormonal

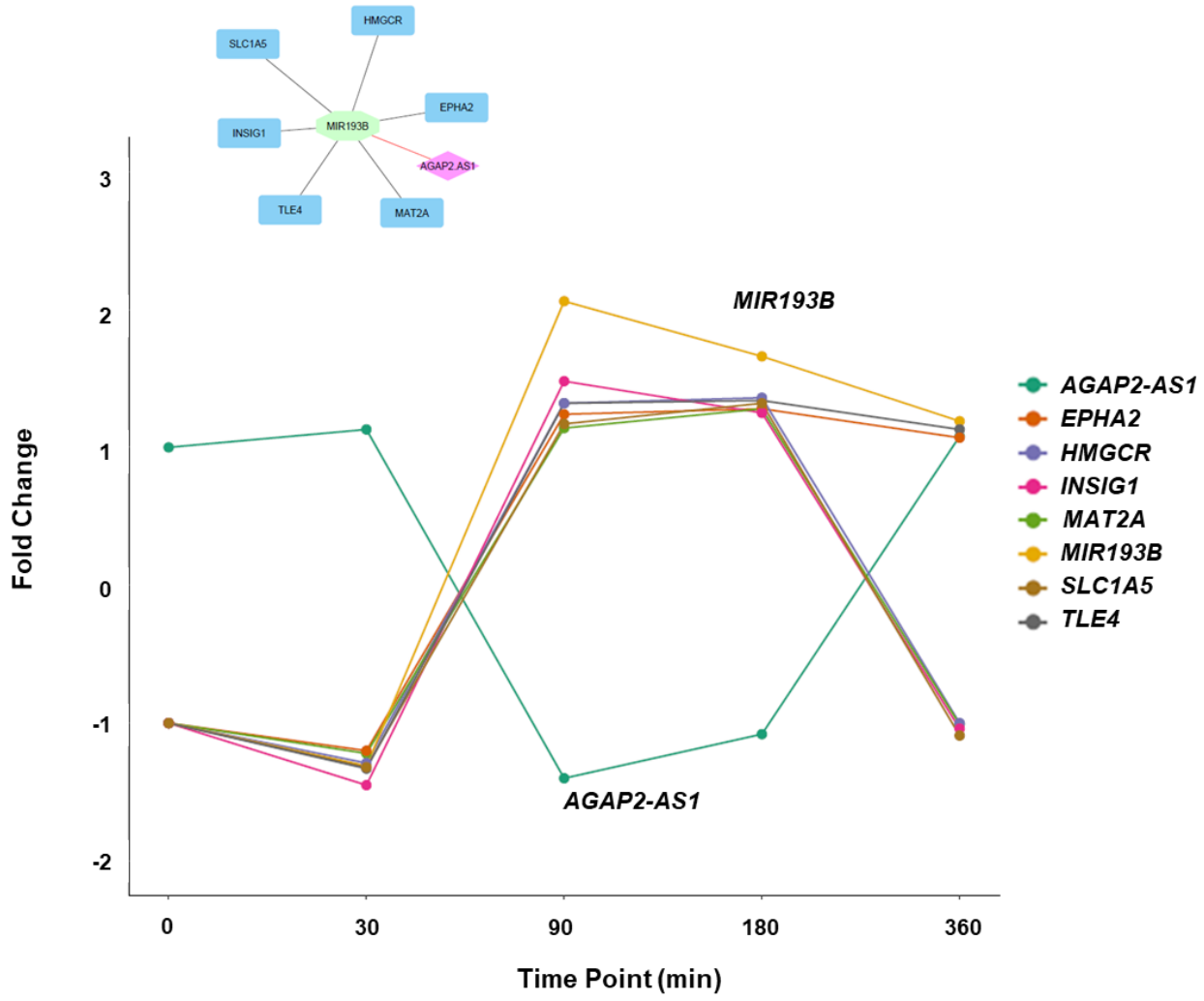

Figure 5.6. Time-course fold-change expression of the miRNA193B mRNA-miRNA-lncRNA cluster
GH treatment leads to decreased expression of *miRNA193B* and all the mRNA at 30 min followed by an increase at 90 min and the subsequent decrease through 180 and 360 min time points. This trend is reversed for the lncRNA, *AGAP2-AS1*.

treatment on the endogenous RNA in a normal mammary epithelial cell-line, as well as the complexity of GH biological regulatory mechanisms.

A major portion of the human transcriptome is composed of non-coding RNA (Dunham et al., 2012). Evolutionary studies emulated on a genome-wide scale have further corroborated the importance of non-coding RNAs by illustrating that, more often than not, alterations in non-coding gene expression is what separates different species from each other as opposed to protein-coding genes (Frith et al., 2005; Li and Liu, 2019; Shabalina and Spiridonov, 2004).

As such, aberrations in the non-coding genome transcription and regulation have been implicated in metabolic dysfunction and a range of diseases including cancer (Esteller, 2011).

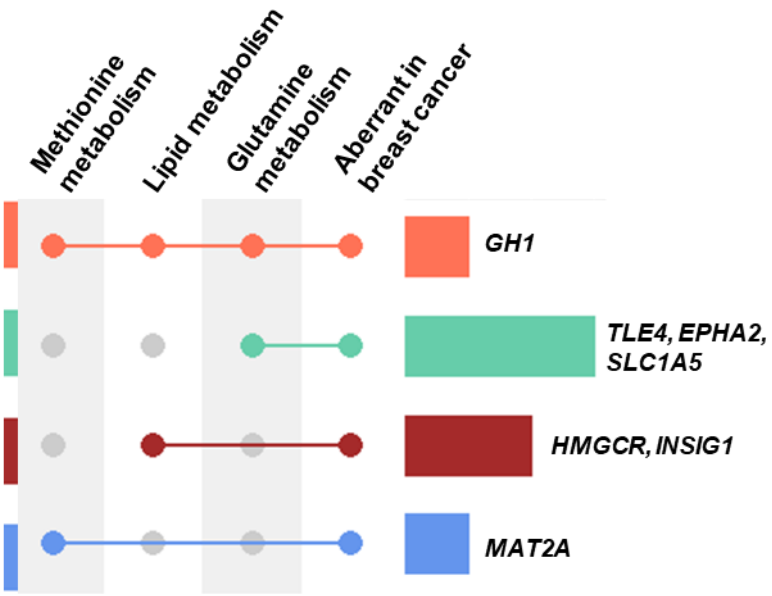

Figure 5.7. Functional representation of genes in mRNA-miRNA-lncRNA cluster
UpSet plot showing overlapping genes from the triple network cluster and the pathways in which they are involved. The genes involved in each pathway are represented as horizontal bars. The matrix represents the intersection between these sets, with intersections colour-coded (left) and the corresponding genes (right).

A cluster of correlated transcripts was identified, which contained six mRNA (*MAT2A*, *SLC1A5*, *TLE4*, *EPHA2*, *HMGCR* and *INSIG1*), one miRNA (*MIR193B*) and one lncRNA (*AGAP2-AS1*). All the genes were positively correlated with *MIR193B* whereas there was a negative correlation between *AGAP2-AS1* and *MIR193B*. Expression analysis of these RNAs in MCF-10A cells at the different time points showed that following a small spike in expression at 30 min, GH treatment leads to downregulation of *AGAP2-AS1* expression at 90 min, followed by a gradual increase in expression through 180 and 360 min. This is in contrast to the expression pattern of *MIR193B*, which decreases at 30 min after GH treatment, increases at 90 min, and decreases again through 180 and 360 min. All of the mRNAs follow a similar trend as that shown by *MIR193B* in response to GH treatment. It should be noted that the logFC values of the significant differentially expressed transcripts were not large in absolute terms but were significant in the context of the time-course experiment. These fold change values, however, are fairly consistent with those reported in previous studies. For example, a treatment time-course analysis of MCF-7 (breast cancer) cell-line with exogenous

GH by Xu *et al* showed comparable fold change differences as a result of this treatment (Xu et al., 2005). logFC is often considered to be arbitrary, as it can depend on a range of different factors, including signal intensity in microarray experiments and the number of replicates.

Previous observations of the functions of these genes revealed involvement in different facets of metabolic pathways that are all related to GH. Additionally, previous studies have demonstrated the effects of autocrine GH on the oncogenic transformation of MCF-10A cells. This is in accordance with our hypothesis that GH potentially mediates some of its functions by fine-tuning the coding/non-coding RNA regulatory circuit, aberrations in which could result in an oncogenic state.

AGAP2-AS1 has been reported to epigenetically silence gene expression by binding to enhancer of zeste homolog 2 (EZH2), which is a histone methyl-transferase and forms one of the important components of the polycomb repressive complex 2 (PRC2) (Li et al., 2016; Luo et al., 2019). PRC2 is a crucial complex involved in modulation of histone modifications. EZH2 has been shown to bind to the promoter of several miRNAs, resulting in epigenetic repression by increased methylation (Liu et al., 2013). Studies conducted on different cancer types have reported that *MIR193B* is regulated by increased methylation, causing epigenetic silencing (Lü et al., 2016; Mazzu et al., 2019a; Rauhala et al., 2010). This gives rise to the possibility that *AGAP2-AS1* could potentially be regulating the expression of the *MIR193B* through interaction with EZH2, thus altering the transcription of *MIR193B*. It has also been shown that *AGAP2-AS1* acts as a molecular sponge/decoy for miRNAs, which results in increased expression of the target genes of these miRNAs (Liu et al., 2019; Zheng et al., 2019). Therefore, it is possible that *AGAP2-AS1* exerts a similar effect on *MIR193B*, though, considering the complexity of the ceRNA process, this may not always be reflected on the *MIR193B* expression level. Although all the identified genes are targets of *MIR19B*, increased *MIR193B* did not result in decreased mRNA of these genes. However, miRNA may function through inhibited translation, and therefore altered mRNA levels often do not become evident on the transcript level.

HMGCR and *INSIG1* are both involved in lipid biosynthesis. At low cellular cholesterol levels, sterol regulatory element–binding proteins (SREBPs) enter the nucleus and bind to the *HMGCR* gene promoter, thus initiating its expression. This nuclear transport of SPREBPs is regulated by INSIG1. Under high cholesterol conditions, INSIG1 is involved in proteasomal degradation of HMGCR (Tsai et al., 2012). Genes associated with lipid metabolism,

including *HMGCR*, *INSIG1* and *INSIG2*, are downregulated by *MIR193B* upon treatment with the insulin sensitiser metformin, leading to activation of PARP cleavage and apoptosis, thus resulting in the death of triple-negative breast cancer cells (Wahdan-Alaswad et al., 2014). This study also reported that increased *MIR193B* leads to cell death only in breast cancer cell-lines, and not in the normal mammary epithelial cell-line MCF-10A. Notably, the *MIR193B* locus is controlled by STAT5 in mammary epithelial cells, which further links GH activity to these results (Yoo et al., 2014). On the other hand, *AGAP2-AS1*, an oncogenic lncRNA, promotes cellular proliferation and inhibits apoptosis in breast cancer cells by promoting activation of the NF-κβ pathway (Dong et al., 2018). *AGAP2-AS1* was found to be negatively correlated with *MIR193B* in the MCF-10A cell-line. This suggests that an oncogenic state, often characterised by increased levels of GH, may result in upregulation of *AGAP2-AS1*, which in turn, may cause a decrease in *MIR193B*, resulting in uncontrolled cell proliferation.

Expression of *MAT2A* is found to be aberrant in breast cancer and could arise due to the promoter region of *MAT2A* possessing a binding site for NF-κβ. Such an increase could result in increased anti-apoptotic signalling, which further enhances *MAT2A* expression (Thi et al., 2015). As previously mentioned, *AGAP2-AS1* functions as an activator of NF-κβ signalling (Dong et al., 2018). EPHA2, SLC1A5 and TLE4 are all involved in glutamine metabolism in breast cancer cells. EPHA2 promotes glutamine metabolism by activating the transcription factors YAP1 and TAZ. These transcription factors increase the expression of SLC1A5 (Edwards et al., 2018). TLE4 recruits the oncogenic factor, c-myc, which promotes glutamine uptake by directly trans-activating SLC1A5 expression (Bott et al., 2015; Jennings and Ish-Horowicz, 2008). *MIR193B* is reported to target YAP1 and decrease its expression (Mazzu et al., 2019b). There is also a double-negative regulatory loop between *MIR193B* and c-myc (Gao et al., 2020; Wahdan-Alaswad et al., 2014). GH has been shown to stimulate oncogenic transformation of MCF-10A cells through increased expression of c-myc (Zhu et al., 2005). *AGAP2-AS1* has been reported to promote upregulation of WNT signalling and cell proliferation, of which c-myc is a major target (Taciak et al., 2018; Zheng et al., 2019). Elevated WNT signalling is also associated with oncogenic conversion of human mammary epithelial cells (Ayyanan et al., 2006). These results suggest that coordinated regulation of GH, *MIR193B*, and *AGAP2-AS1* may contribute towards preventing the transformation of MCF-10A cells to an oncogenic state.

Several studies have demonstrated the effects of GH on regulation of expression of multiple non-coding RNAs in normal and diseased states (Chang et al., 2016; Melia et al., 2015; Perry et al., 2013; Zhang et al., 2015b). Zhang *et al.* reported that the expression of the miRNA cluster, miR-96-182-183, is stimulated by GH in MCF-7 cells and results in suppression of *BRMS1L* (breast cancer metastasis suppressor 1-like) mRNA, thus promoting epithelial-mesenchymal transition (EMT) and invasion in breast cancer (Zhang et al., 2015b). Melia *et al.* characterised a subset of GH-regulated hepatic lncRNAs that correlated strongly with the sex-specific effects of GH. They also identified chromatin state patterns associated with GH-mediated regulation that are distinct in male-specific and female-specific lncRNAs (Melia et al., 2015). Moreover, several miRNAs and lncRNAs have also been documented to modulate the transcription of GH and the GHR. For example, Zhang *et al.* showed that two of the miRNAs in the GH-stimulated miR-96-182-183 cluster can, in turn, target the GHR, which could be responsible for mediating a negative feedback loop of the GH/GHR signalling pathway (Zhang et al., 2015b). Another study demonstrated that four miRNAs, miR-129–5p, miR-142–3p, miR-202, and miR-16, can repress the expression of the *GHR* gene and the GHR protein in both normal (HEK293) and cancer (MCF-7 and LNCaP) cells (Elzein and Goodyer, 2014). In addition, some studies have reported the effects of non-coding RNAs on GH secretion. For example, the *MIR205HG* lncRNA plays a role in regulation of GH production in the anterior pituitary (Du et al., 2019). An expression study by Chang *et al.* identified a set of correlated mRNA, miRNA and lncRNA that are irregularly expressed in GHR knockout (GHR-KO) mice and are implicated in the steroid biosynthesis pathway (Chang et al., 2016). This highlights the importance of studying the interplay between the coding and the non-coding expression profiles in the context of GH functions. Our study extends these observation and adds further insight by constructing and analysing the mRNA-miRNA-lncRNA triple network.

Knockdown of single miRNA or lncRNA has been shown to yield particularly subtle effects, even on its direct targets. It is now proposed that most non-coding RNAs function as micromanagers of gene translation and serve as rheostats that polish the gene transcription machinery to balance phenotypes, which are associated with several different pathways and mechanisms (Bartel and Chen, 2004; Ebert and Sharp, 2012; Hornstein and Shomron, 2006). Since cancer is often reflective of dysregulated cellular machinery, this might explain why miRNA and lncRNA are associated with so many diverse types of cancer (Reddy, 2015). It is plausible that the regulation of (and the interaction between) miRNA, lncRNA, and the target

mRNA is important in the modulation of biological pathways, as opposed to more marked effects, such as the complete alteration/knockdown of a phenotype. A novel ceRNA network was constructed, which was associated with a time series of GH treatment in human mammary epithelial cell-line. It was shown that evaluating gene expression in terms of coding as well as non-coding genes can provide insights into the complexity of functions mediated by GH and could potentially function as a model to understand the oncogenic actions of GH. However, subsequent cell-based assays and knockdown studies in both normal and cancer cells are required to further validate these findings and would have been the logical next step. Unfortunately, due to time constraints, it was not possible to conduct these studies.

CHAPTER 6

General Discussion

6.1 Summary of findings

GH plays an important role in multiple facets of growth and metabolism. Compromised GH signalling has been linked to growth disorders, cancer and other diseases (Guevara-Aguirre et al., 2018). In this thesis, the intricacy of GH-induced effects on the genome and the proteome was explored, and the data generated was utilised to address the three key aims in Chapters 3, 4 and 5, respectively.

Chapter 3 investigated whether genetic variation occurring in genes associated with the *GH* locus may contribute to disease through deregulating spatial connections within the genome. Computational analysis formed the basis of that investigation. Common SNPs across the *GH* locus were analysed using the computational pipeline, CoDeS3D. This analysis identified the locus as a potential regulatory hub for multiple genes, located both proximally as well as distally (including on other chromosomes). Several genes that were identified by this analysis are involved in GH-related cellular signalling pathways, including pathways in cancer. These results suggest that regulatory regions within the *GH* locus region could potentially be involved in modulation of GH functions and GH-induced carcinogenesis.

The GHR has previously been shown to translocate to the nucleus, but the function of this nuclear localisation is unclear (Lobie et al., 1994a). In Chapter 4, the GHR was found to translocate to the nucleus of RL95-2 cells following 5 min stimulation with GH. It was shown that the GHR co-localises with HMGN1 in the nucleus after GH treatment, and that this co-localisation correlates with the differential expression of HMGN1 gene targets in this cell-line. HMGN1 is a transcription factor that alters DNA-histone interactions and potentially affects chromatin conformation, leading to changes in gene expression. This study further corroborates the effect of GH on spatial genomic interactions, which was investigated in Chapter 3, and substantiates the significance of GH-induced nuclear import of GHR as an alternative regulatory mechanism for mediation of GH function.

Classically, the effects of GH on cell growth and differentiation are effected through interaction with the GHR. GH binding to the GHR triggers a downstream signal cascade, which is critical for cell growth and survival. This results in pivotal effects of GH through altered coding and non-coding transcription profiles, by direct stimulation of transcription through activation of transcription factors including STAT5 and regulation of epigenetic alterations. In Chapter 5, it was determined that GH stimulation of MCF-10A cells can lead to changes in coding and non-coding RNA expression. These results were used to construct an mRNA-miRNA-lncRNA

triple regulatory network. This network identified novel mechanisms that could potentially be involved in the mediation of GH activity and in the impairment of oncogenic transformation of MCF-10A cells. This analysis showed that evaluating gene expression in terms of both the coding and the non-coding genome can help unravel the complexity of GH functions and may serve as a model to realise the oncogenic potential of GH.

6.2 Dysregulation of non-coding genome can alter transcriptional regulation and lead to disease

The significance of the non-coding genome in cellular homeostasis and disease is gaining momentum after being marginalised as "junk" for decades. There is now increasing evidence suggesting that diseases are often the consequence of faulty wiring of the regulatory circuit between the coding and the non-coding genome. The most-well-characterised non-coding RNAs involve miRNAs and lncRNAs. Aberrant non-coding RNA processing and transcription, a consequence of genetic and epigenetic perturbation, are implicated heavily not just in cancer but in developmental, neurological and cardiovascular disorders (Chen and Chen, 2020; Esteller, 2011; Hanly et al., 2018). In Chapter 5, it was shown that the identified coordinated regulation of miRNA/mRNA/lncRNA may have implications in the oncogenic transformation of MCF-10A cells. This is in accordance with the literature, in which multiple studies implicate non-coding RNAs (ncRNAs) in malignant progression, indicating a prospective use for ncRNAs as biomarkers of cancer (Barsyte-Lovejoy et al., 2006; Guo et al., 2014; Liang et al., 2019; Lu et al., 2017; Piao and Ma, 2012).

Therapeutic approaches such as ncRNA silencers and antagonists are being explored but these approaches are limited by the lack of knowledge regarding the different mechanisms of non-coding RNA function, and by the enormous scale of the whole cell ncRNAome. This can be addressed partially by employing computational biology methods for large-scale target discovery and drug screening, as well as by developing transgenic models, which can aid greatly in prioritising more viable targets.

6.3 Limitations

Certain limitations were identified over the course of this research. The eQTL data (used for analysis in Chapter 3) obtained from GTEx is derived from primarily old (age > 50 years, 68.5%), Caucasian (85.2%), males (65.8%) (GTEx v7) (Ardlie et al., 2015). Additionally,

placental samples were absent from this dataset. Therefore, this dataset precludes the accurate delineation of sex-specific effects on gene expression. This is of relevance as there is a clear sexually dimorphic pattern in GH secretion as shown in both in humans and experimental animal models (Geary et al., 2003; Lichanska and Waters, 2008; Liu et al., 2016b). In males, GH is secreted in a discontinuous, higher-amplitude pulsatile manner (Bonert and Melmed, 2017) whilst in females is released in a more tonic pattern with higher baseline concentrations. This difference in secretion pattern is likely a result of delayed feedback loop between GH and somatostatin. Thus, GH secretory pattern is critical to the expression of sexually dimorphic genes in the adipose tissue, muscle, and liver, leading to sex-specific effects on growth, puberty, and metabolism (Lichanska and Waters, 2008).

The age distribution of data obtained from GTEx is also of consideration as samples were procured from males of primarily old age (> 50 years, 68.5%). Since there is a decline in GH secretion with age, and since GH secretion is functionally more important during childhood and puberty, some interesting connections may be missing from this analysis (Geary et al., 2003; Hindmarsh et al., 1999). Despite this inherent limitation in the dataset, it was possible to interpret some biological meaning, given the presence of some female-tissue-specific information in the GTEx data.

In Chapter 4, combined IP/mass spectrometry (IP/MS) and microarray analysis was used to identify potential transcription factors that interact with the GHR in the nucleus. The inclusion of complementary techniques such as ChIP-seq and siRNA knockdown of identified transcription factors would have provided a greater depth of detail regarding transcriptional regulation by nuclear GHR. The choice of techniques was greatly restricted by time and cost. Additionally, conducting IP/MS at only one time-point meant that it was not possible to observe all GHR interactions, owing to the transient nature of protein-protein interactions. The 5 min GH treatment time-point was chosen based on combined preliminary experiments which included western blotting, immunoprecipitation and immunofluorescence, which demonstrated the rapid translocation of GHR into the nucleus. A comprehensive time course experiment that captures GHR-protein interactions in the nucleus and the cytoplasm would allow GHR interactions in different subcellular locations to be tracked in detail over time. This may provide a more nuanced characterisation of the kinetics of GHR nuclear translocation.

In Chapter 5, the interactions between mRNA/miRNA/lncRNA, using multiple time points of GH treatment in the MCF-10A cell-line were analysed. Due to cost limitations, this time course

analysis was carried out in only one cell-line. Comparative analysis with a cancerous cell-line would give a broader overview of how coding and non-coding transcriptome function together as a regulatory circuit and how erroneous wiring of this circuit can lead to oncogenic transformation of normal cells. Another limitation is the smaller log fold change values of differentially expressed transcripts at specific time points, which could possibly be because MCF-10A is a normal cell-line and the effects of GH treatment are likely subtle. Repeating this experiment with a different cell-line, might give better results in terms of absolute fold change values. Although, these fold change values are comparable with those shown in a GH time-course analysis performed with MCF-7 (breast cancer) cell-line (Xu et al., 2005). Additionally, fold changes as a result of GH treatment in RL952 cell-line (endometrial cancer cell-line) were also very similar (Chapter 4). Validation of this data was the intended next step in this experiment, but due to time limitations this was not able to be carried out.

6.4 Future directions

In Chapter 3, it was discerned that SNPs located in the GH locus that are associated with gene expression alterations of local and distal genes. The analysis identified regulatory hubs in the GH locus region that impact on GH functions and contribute to diseases including cancer. These regulatory regions can be functionally validated using a combination of reporter assays, genome editing techniques such as CRISPR/Cas9, and proximity ligation assays such as Hi-C. It could be interesting to implement these techniques on a diverse array of cell and animal models - both normal and diseased- to reflect cell-specific effects of GH.

Enhancer activity of identified regulatory sequences could be tested using luciferase enhancer assays. Regions harbouring disease-associated SNPs can be amplified using PCR, cloned into a luciferase reporter vector, and sequenced to confirm genotype. These regions can then be assayed for their ability to drive luciferase expression in transfected cell-lines to characterise their enhancer potential. Consequently, the identified SNPs can be edited using CRISPR/Cas9 genome editing. The impact of SNP modification on spatial connectivity and consequent gene expression could then be assessed by coupling Hi-C with RNA-seq. However, one needs to be cautious in interpreting data from studies of this nature, as genome editing processes are extremely complex and often result in additional undesired off-target effects that make it difficult to confidently assign any resultant change to our intended edit. Additionally, regulatory regions often act in conjunction with other loci, which means that modification of

isolated SNPs may only yield a very subtle effect on gene expression and phenotype, or even none at all.

Studies examining the impact of hormones on 3D chromatin conformation and spatial interactions within the nucleus of target cells are minimal and predominantly address steroid hormones (Le Dily and Beato, 2018; Le Dily et al., 2019; Zhou et al., 2019). The GH locus is of particular interest, since regulation of spatial interactions may serve as an alternative mechanism for GH-mediated actions. The key downstream effects of GH occur through altered transcription profiles of coding and non-coding RNA, by stimulation of transcription and regulation of epigenetic modifications (Álvarez-Nava and Lanes, 2017; Chia, 2014; Shafiei et al., 2008; Xu et al., 2005). In Chapter 4, it was demonstrated that GHR co-localises with the transcription factor, HMGN1, in the nucleus, and also correlates with differential expression of HMGN1 target genes. HMGN1 is a chromosomal protein that dynamically binds to chromatin despite the absence of a specific DNA binding sequence, leading to changes in epigenetic make-up and chromosomal conformation (Lim et al., 2004; Nanduri et al., 2020).

These observations led to the hypothesis that GH stimulation of GHR may lead to measurable changes in the 3D spatial organisation in the cell nucleus. Alterations in 3D genome conformation following GH stimulation could be explored using chromosome conformation capture techniques (Hi-C) and antagonism of GH. RNA-seq could be performed to study consequent changes in gene transcription. This approach may further identify novel GH-responsive regulatory regions in the genome and even novel GH functions.

A striking feature of *GH* gene evolution is that it is in contrast to the classical view of molecular evolution, in that it evolved by short bursts of accelerated changes, rather than slow, constant change over time (Wallis, 2014). One of these rapid changes during GH evolution accounts for significant differences observed between human and non-primate GH; this is why non-primate GHs are inactive in humans (Wallis and Wallis, 2006). Genetic regions that are associated with complex regulatory landscapes generally remain evolutionarily conserved between different species as part of their maintenance (Reilly and Noonan, 2016). Expression patterns of genes involved in coordinated regulation of critical biological processes remain typically stable between species (Merkin et al., 2012). However, it is heavily debated as to how these expression patterns are evolutionarily maintained by a continually differentiating array of regulatory elements because of the absence of consistent regulatory datasets across divergent species (Berthelot et al., 2018; Necsulea and Kaessmann, 2014). Analysis in Chapter 3 showed

that the *GH* locus forms a composite spatial regulatory region that coordinates transcription of both local and distal genes involved in mediation of GH-related signalling pathways. It would be interesting to identify similar regulatory regions that regulate the expression of genes in the *GH* locus. Comparative dissection and analysis of the human *GH* locus spatial regulatory hub with other primate and non-primate mammals could help elucidate the evolutionary significance of this co-regulatory phenomenon. Unlike primates, non-primates do not contain a five gene *GH* locus, but some of the regulatory regions may still be conserved.

As mentioned previously, it is apparent that there exist two distinct mechanisms for signal transduction by cell-surface receptors, namely the cytoplasmic signalling pathway and the nuclear translocation of receptors. The overarching question, however, remains regarding how these two fundamentally different signalling modes are coordinated. It would be worthwhile to explore whether one of these modes evolved exclusively to serve as a compensation mechanism for some of the functions mediated by the other, or whether both these mechanisms evolved at the same point to work synergistically towards the same functional outcome.

6.5 Conclusions

This thesis has contributed to understanding the mechanism of GH actions in human development and disease through the application of genomic, transcriptomic, and proteomic approaches. Novel putative genetic regulatory hubs in the *GH* locus were identified that potentially mediate GH-linked functions in homeostasis and disease development. In addition, a novel mRNA-miRNA-lncRNA regulatory network was identified that is modulated by GH and may be associated with the prevention of oncogenic transformation of MCF-10A breast epithelial cells. Further, it was demonstrated that GHR translocates into the nucleus of RL952 cells following GH treatment for 5 min and co-localises with transcription factors in the nucleus. Finally, it was shown that the target genes of these transcription factors are differentially expressed as a result of GH treatment, thus providing novel insights into the consequence of GHR nuclear import. Since GH plays an important physiological role in the body and is implicated in numerous diseases, such as cancer, it is imperative to understand the molecular mechanism underlying GH actions and the complexity associated with it. Overall, the data generated in this thesis highlight multiple mechanisms involved in the modulation of GH function.

Appendices

Appendix I

Appendix files for Chapter 3 are available on Figshare, doi: 10.17608/k6.auckland.12731906.v1

Supplementary Table 1-List of common SNPs (dbSNP147) across *GH* gene locus (Chr17:62080000-61920000; GRCh37/hg19) used for analysis

Supplementary Table 2-Results of analysis of common SNPs across *GH* locus by a computational pipeline, CoDeS3D. These results are sorted by SNPs located 5'- to 3'-in the locus

Supplementary Table 3-Overview of the number of SNPs and tissues which show an eQTL with the respective eGenes in cis as well as in trans

Supplementary Table 4-(a) Summary of number of trans-eGenes which have eQTLs with varying number of SNPs, (b) List of eGenes which show eQTLs with multiple SNPs

Supplementary Table 5-Number of trans-eGenes which have connections with analysed SNPs in each tissue

Supplementary Table 6-Number of cis-eGenes which have connections with analysed SNPs in each tissue

Supplementary Table 7-Summary of number of eQTLs and tissues associated with identified SNPs

Supplementary Table 8-Number of SNPs which have eQTLs in different tissues

Supplementary Table 9-Summary of number of SNPs which have eQTLs in all varying number of tissues

Supplementary Table 10-List of identified eGenes associated with *GH* locus SNPs in 48 tissues

Supplementary Table 11-Pathway analysis in g:Profiler restricted to KEGG database sorted by adjusted *P* value

Supplementary Table 12-Pathway analysis in g:Profiler restricted to WikiPathways database sorted by adjusted *P* value

Supplementary Table 13-List of eGenes enriched in mTOR and Wnt pathways

Supplementary Table 14- Overlap between SNPs which are reported to be important for regulatory activity and SNPs identified by CoDeS3D.

Supplementary Table 15-Summary of CoDeS3D connections of SNPs associated with regulatory functions (identified by van Arenbergen *et al*) in K562 cell-line

Supplementary Table 16- Reported enhancer regions encompassing GH locus region identified in GeneHancer database and SNPs which have spatial-eQTLs with genes in cis and trans.

Supplementary Table 17-SNPs associated with eGenes enriched in mTOR pathway in different tissues

Supplementary Table 18-SNPs associated with eGenes enriched in Wnt pathway in different tissues

Appendix II

Appendix files for Chapter 4 are available on Figshare, doi: 10.17608/k6.auckland.12731912.v1

Supplementary Table 1-Results of mass spectrometry analysis of differentially expressed proteins which are increased after treatment of RL95-2 cells with GH for 5min

Supplementary Table 2-Results of mass spectrometry analysis of differentially expressed proteins which are exclusively enriched after treatment of RL95-2 cells with GH for 5min

Supplementary Table 3-Results of mass spectrometry analysis of differentially expressed proteins which are decreased after treatment of RL95-2 cells with GH for 5min

Supplementary Table 4-Pathway analysis in g:Profiler restricted to REACTOME database sorted by adjusted *P* value

Supplementary Table 5-Pathway analysis in g:Profiler restricted to KEGG database sorted by adjusted *P* value

Supplementary Table 6-Pathway analysis in g:Profiler restricted to WIKIPATHWAYS database sorted by adjusted *P* value

Supplementary Table 7-Results of microarray analysis of differentially expressed genes obtained after treatment of RL95-2 cells with GH for 90min, sorted by adjusted *P* value

Supplementary Table 8-Pathway analysis in g:Profiler restricted to REACTOME database sorted by adjusted *P* value

Supplementary Table 9-Pathway analysis in g:Profiler restricted to KEGG database sorted by adjusted *P* value

Supplementary Table 10-Pathway analysis in g:Profiler restricted to WIKIPATHWAYS database sorted by adjusted *P* value

Supplementary Table 11: Intersection results between gene targets of transcription factors, HMGN1 and SUMO1, and differentially expressed genes in microarray analysis

Appendix III

Appendix files for Chapter 5 are available on Figshare, doi: 10.17608/k6.auckland.12731918.v1

Supplementary Table 1-Results of microarray analysis of differentially expressed genes (mRNA) obtained after treatment of MCF10A cells with GH for 30, 90, 180 and 360 min, sorted by adjusted P value

Supplementary Table 2-Results of microarray analysis of differentially expressed micro RNA (miRNA) obtained after treatment of MCF10A cells with GH for 30, 90, 180 and 360 min, sorted by adjusted P value

Supplementary Table 3-Results of microarray analysis of differentially expressed long non-coding RNA (lncRNA) obtained after treatment of MCF10A cells with GH for 30, 90, 180 and 360 min, sorted by adjusted P value

Supplementary Table 4-Pathway analysis in g:Profiler restricted to Wikipathways database sorted by adjusted P value

Supplementary Table 5-Pathway analysis in g:Profiler restricted to KEGG database sorted by adjusted P value

Supplementary Table 6-Pearson's correlation coefficient values between differentially expressed mRNA, miRNA and lncRNA

References

Aguilar, R.C., and Wendland, B. (2005). Endocytosis of membrane receptors: Two pathways are better than one. Proc. Natl. Acad. Sci. U. S. A. *102*, 2679–2680.

Akerman, I., Tu, Z., Beucher, A., Rolando, D.M.Y., Sauty-Colace, C., Benazra, M., Nakic, N., Yang, J., Wang, H., Pasquali, L., et al. (2017). Human Pancreatic β Cell lncRNAs Control Cell-Specific Regulatory Networks. Cell Metab. *25*, 400–411.

Alam, T., Medvedeva, Y.A., Jia, H., Brown, J.B., Lipovich, L., and Bajic, V.B. (2014). Promoter Analysis Reveals Globally Differential Regulation of Human Long Non-Coding RNA and Protein-Coding Genes. PLoS One *9*, e109443.

Aleksic, T., Chitnis, M.M., Perestenko, O. V., Gao, S., Thomas, P.H., Turner, G.D., Protheroe, A.S., Howarth, M., and Macaulay, V.M. (2010). Type 1 insulin-like growth factor receptor translocates to the nucleus of human tumor cells. Cancer Res. *70*, 6412–6419.

Aleksic, T., Gray, N., Wu, X., Rieunier, G., Osher, E., Mills, J., Verrill, C., Bryant, R.J., Han, C., Hutchinson, K., et al. (2018). Nuclear IGF1R interacts with regulatory regions of chromatin to promote RNA polymerase II recruitment and gene expression associated with advanced tumor stage. Cancer Res. *78*, 3497–3509.

Alfarano, A., Indraccolo, S., Circosta, P., Minuzzo, S., Vallario, A., Zamarchi, R., Fregonese, A., Calderazzo, F., Faldella, A., Aragno, M., et al. (1999). An alternatively spliced form of CD79b gene may account for altered B- cell receptor expression in B-chronic lymphocytic leukemia. Blood *93*, 2327–2335.

Álvarez-Nava, F., and Lanes, R. (2017). GH/IGF-1 signaling and current knowledge of epigenetics; A review and considerations on possible therapeutic options. Int. J. Mol. Sci. *18*, 1–13.

Ambros, V., Bartel, B., Bartel, D.P., Burge, C.B., Carrington, J.C., Chen, X., Dreyfuss, G., Eddy, S.R., Griffiths-Jones, S., Marshall, M., et al. (2003). A uniform system for microRNA annotation. RNA *9*, 277–279.

Ardlie, K.G., Deluca, D.S., Segre, A. V., Sullivan, T.J., Young, T.R., Gelfand, E.T.,

Trowbridge, C.A., Maller, J.B., Tukiainen, T., Lek, M., et al. (2015). The Genotype-Tissue Expression (GTEx) pilot analysis: Multitissue gene regulation in humans. Science (80-.). *348*, 648–660.

van Arensbergen, J., Pagie, L., FitzPatrick, V.D., de Haas, M., Baltissen, M.P., Comoglio, F., van der Weide, R.H., Teunissen, H., Võsa, U., Franke, L., et al. (2019). High-throughput identification of human SNPs affecting regulatory element activity. Nat. Genet. *51*, 1160–1169.

Ayyanan, A., Civenni, G., Ciarloni, L., Morel, C., Mueller, N., Lefort, K., Mandinova, A., Raffoul, W., Fiche, M., Dotto, G.P., et al. (2006). Increased Wnt signaling triggers oncogenic conversion of human breast epithelial cells by a Notch-dependent mechanism. Proc. Natl. Acad. Sci. U. S. A. *103*, 3799–3804.

Bach, L.A. (2018). 40 years of IGF1: IGF-binding proteins. J. Mol. Endocrinol. *61*, T11–T28.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. Cell Res. *21*, 381–395.

Barlow, D.P., Stöger, R., Herrmann, B.G., Saito, K., and Schweifer, N. (1991). The mouse insulin-like growth factor type-2 receptor is imprinted and closely linked to the Tme locus. Nature *349*, 84–87.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-Resolution Profiling of Histone Methylations in the Human Genome. Cell *129*, 823–837.

Barsyte-Lovejoy, D., Lau, S.K., Boutros, P.C., Khosravi, F., Jurisica, I., Andrulis, I.L., Tsao, M.S., and Penn, L.Z. (2006). The c-Myc oncogene directly induces the H19 noncoding RNA by allele-specific binding to potentiate tumorigenesis. Cancer Res. *66*, 5330–5337.

Bartel, D.P. (2004). MicroRNAs: Genomics, Biogenesis, Mechanism, and Function. Cell *116*, 281–297.

Bartel, D.P. (2009). MicroRNAs: Target Recognition and Regulatory Functions. Cell *136*, 215–233.

Bartel, D.P., and Chen, C.Z. (2004). Micromanagers of gene expression: The potentially widespread influence of metazoan microRNAs. Nat. Rev. Genet. *5*, 396–400.

Bartke, A. (2011). Growth hormone, insulin and aging: The benefits of endocrine defects. Exp. Gerontol. *46*, 108–111.

Bartolomei, M.S., Zemel, S., and Tilghman, S.M. (1991). Parental imprinting of the mouse H19 gene. Nature *351*, 153–155.

Barton, D.E., Foellmer, B.E., Woodland, W.I., and Francke, U. (1989). Chromosome mapping of the growth hormone receptor gene in man and mouse. Cytogenet. Genome Res. *50*, 137–141.

Baszczyński, J., and Goldstein, L. (1967). [Clinical picture of intraventricular septal defects in children of various age groups]. Pediatr. Pol. *42*, 1071–1075.

Batista, P.J., and Chang, H.Y. (2013). Long noncoding RNAs: Cellular address codes in development and disease. Cell *152*, 1298–1307.

Baumann, G.P. (2009). Growth hormone isoforms. Growth Horm. IGF Res. *19*, 333–340.

B Wightman, I.H.G.R., Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell *75*, 855–862.

Ben-Jonathan, N., Hugo, E.R., Brandebourg, T.D., and LaPensee, C.R. (2006). Focus on prolactin as a metabolic hormone. Trends Endocrinol. Metab. *17*, 110–116.

Ben-Jonathan, N., LaPensee, C.R., and LaPensee, E.W. (2008). What Can We Learn from Rodents about Prolactin in Humans? Endocr. Rev. *29*, 1–41.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J. R. Stat. Soc. Ser. B *57*, 289–300.

Benko, S., Fantes, J.A., Amiel, J., Kleinjan, D.J., Thomas, S., Ramsay, J., Jamshidi, N., Essafi, A., Heaney, S., Gordon, C.T., et al. (2009). Highly conserved non-coding elements on either side of SOX9 associated with Pierre Robin sequence. Nat. Genet. *41*, 359–364.

Berg, M.A., Argente, J., Chernausek, S., Gracia, R., Guevara-Aguirre, J., Hopp, M., Perez-Jurado, L., Rosenbloom, A., Toledo, S.P.A., and Francke, U. (1993). Diverse growth hormone receptor gene mutations in Laron syndrome. Am. J. Hum. Genet. *52*, 998–1005.

Berger, P., Untergasser, G., Hermann, M., Hittmair, A., Madersbacher, S., and Dirnhofer, S. (1999). The testis-specific expression pattern of the growth hormone/placental lactogen

(GH/PL) gene cluster changes with malignancy. Hum. Pathol. *30*, 1201–1206.

Bergman, D., Halje, M., Nordin, M., and Engström, W. (2013). Insulin-like growth factor 2 in development and disease: A mini-review. Gerontology *59*, 240–249.

Berthelot, C., Villar, D., Horvath, J.E., Odom, D.T., and Flicek, P. (2018). Complexity and conservation of regulatory landscapes underlie evolutionary resilience of mammalian gene expression. Nat. Ecol. Evol. *2*, 152–163.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigó, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E., et al. (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799–816.

De Boer, H., Blok, G.J., and Van Der Veen, E.A. (1995). Clinical aspects of growth hormone deficiency in adults. Endocr. Rev. *16*, 63–86.

Boguszewski, C.L. (2003). Molecular heterogeneity of human GH: From basic research to clinical implications. J. Endocrinol. Invest. *26*, 274–288.

Bole-Feysot, C., Goffin, V., Edery, M., Binart, N., and Kelly, P.A. (1998). Prolactin (PRL) and its receptor: Actions, signal transduction pathways and phenotypes observed in PRL receptor knockout mice. Endocr. Rev. *19*, 225–268.

Bonert, V.S., and Melmed, S. (2017). Growth Hormone. In The Pituitary, (Elsevier), pp. 85–127.

Bonev, B., and Cavalli, G. (2016). Organization and function of the 3D genome. Nat. Rev. Genet. *17*, 661–678.

Borsani, G., Tonlorenzi, R., Simmler, M.C., Dandolo, L., Arnaud, D., Capra, V., Grompe, M., Pizzuti, A., Muzny, D., Lawrence, C., et al. (1991). Characterization of a murine gene expressed from the inactive X chromosome. Nature *351*, 325–329.

Bott, A.J., Peng, I.C., Fan, Y., Faubert, B., Zhao, L., Li, J., Neidler, S., Sun, Y., Jaber, N., Krokowski, D., et al. (2015). Oncogenic Myc Induces Expression of Glutamine Synthetase through Promoter Demethylation. Cell Metab. *22*, 1068–1077.

Bougen, N.M., Steiner, M., Pertziger, M., Banerjee, A., Brunet-Dunand, S.E., Zhu, T., Lobie, P.E., and Perry, J.K. (2012). Autocrine human GH promotes radioresistance in mammary and

endometrial carcinoma cells. Endocr. Relat. Cancer *19*, 625–644.

Branco, M.R., and Pombo, A. (2006). Intermingling of chromosome territories in interphase suggests role in translocations and transcription-dependent associations. PLoS Biol. *4*, 780–788.

Brand, T.M., Iida, M., Li, C., and Wheeler, D.L. (2011). The nuclear epidermal growth factor receptor signaling network and its role in cancer. Discov. Med. *12*, 419–432.

Brandebourg, T., Hugo, E., and Ben-Jonathan, N. (2007). Adipocyte prolactin: regulation of release and putative functions. Diabetes, Obes. Metab. *9*, 464–476.

Brannan, C.I., Dees, E.C., Ingram, R.S., and Tilghman, S.M. (1990). The product of the H19 gene may function as an RNA. Mol. Cell. Biol. *10*, 28–36.

Braun, J.E., Truffault, V., Boland, A., Huntzinger, E., Chang, C. Te, Haas, G., Weichenrieder, O., Coles, M., and Izaurralde, E. (2012). A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5′ exonucleolytic degradation. Nat. Struct. Mol. Biol. *19*, 1324–1331.

Brittain, A.L., Basu, R., Qian, Y., and Kopchick, J.J. (2017). Growth hormone and the epithelial-To-mesenchymal transition. J. Clin. Endocrinol. Metab. *102*, 362–3673.

Broderick, J.A., and Zamore, P.D. (2014). Competitive Endogenous RNAs Cannot Alter MicroRNA Function InVivo. Mol. Cell *54*, 711–713.

Brouwer-Visser, J., and Huang, G.S. (2015). IGF2 signaling and regulation in cancer. Cytokine Growth Factor Rev. *26*, 371–377.

Brown, C.J., Ballabio, A., Rupert, J.L., Lafreniere, R.G., Grompe, M., Tonlorenzi, R., and Willard, H.F. (1991). A gene from the region of the human X inactivation centre is expressed exclusively from the inactive X chromosome. Nature *349*, 38–44.

Bryant, D.M., and Stow, J.L. (2005). Nuclear translocation of cell-surface receptors: Lessons from fibroblast growth factor. Traffic *6*, 947–954.

Buniello, A., Macarthur, J.A.L., Cerezo, M., Harris, L.W., Hayhurst, J., Malangone, C., McMahon, A., Morales, J., Mountjoy, E., Sollis, E., et al. (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res. *47*, D1005–D1012.

Cairns, B.R. (2009). The logic of chromatin architecture and remodelling at promoters. Nature *461*, 193–198.

Carpenter, G., and Liao, H.-J. (2013). Receptor Tyrosine Kinases in the Nucleus. Cold Spring Harb. Perspect. Biol. *5*, a008979–a008979.

Carré, N., and Binart, N. (2014). Prolactin and adipose tissue. Biochimie *97*, 16–21.

Carter-Su, C., Schwartz, J., and Argetsinger, L.S. (2016). Growth hormone signaling pathways. Growth Horm. IGF Res. *28*, 11–15.

Catto, J.W.F., Alcaraz, A., Bjartell, A.S., De Vere White, R., Evans, C.P., Fussel, S., Hamdy, F.C., Kallioniemi, O., Mengual, L., Schlomm, T., et al. (2011). MicroRNA in prostate, bladder, and kidney cancer: A systematic review. Eur. Urol. *59*, 671–681.

Chang, L., Qi, H., Xiao, Y., Li, C., Wang, Y., Guo, T., Liu, Z., and Liu, Q. (2016). Integrated analysis of noncoding RNAs and mRNAs reveals their potential roles in the biological activities of the growth hormone receptor. Growth Horm. IGF Res. *29*, 11–20.

Chang, L., Zhou, G., Soufan, O., and Xia, J. (2020). miRNet 2.0: network-based visual analytics for miRNA functional analysis and systems biology. Nucleic Acids Res. *48*, W244–W251.

Chaulk, S. (2011). Role of pri-miRNA tertiary structure in miR-17~92 miRNA biogenesis. RNA Biol. *8*, 1105–1114.

Chen, K.-W., and Chen, J.-A. (2020). Functional Roles of Long Non-coding RNAs in Motor Neuron Development and Disease. J. Biomed. Sci. *27*, 38.

Chen, Y., and Chen, A. (2019). Unveiling the gene regulatory landscape in diseases through the identification of dnase I-hypersensitive sites (Review). Biomed. Reports *11*, 87–97.

Chen, E.Y.E.Y., Liao, Y.C.Y.-C., Smith, D.H.D.H., Barrera-Saldana, H.A., Gelinas, R.E.R.E., Seeburg, P.P.H., Barrera-Saldaña, H.A., Gelinas, R.E.R.E., Seeburg, P.P.H., Barrera-Saldana, H.A., et al. (1989). The human growth hormone locus: nucleotide sequence. Biol. Evol. Genomics *4*, 479–497.

Chen, H., Levo, M., Barinov, L., Fujioka, M., Jaynes, J.B., and Gregor, T. (2018). Dynamic interplay between enhancer–promoter topology and gene activity. Nat. Genet. *50*, 1296–1303.

Chen, J., Xu, J., Li, Y., Zhang, J., Chen, H., Lu, J., Wang, Z., Zhao, X., Xu, K., Li, Y., et al. (2017). Competing endogenous RNA network analysis identifies critical genes among the different breast cancer subtypes. Oncotarget *8*, 10171–10184.

Chesnokova, V., Zonis, S., Zhou, C., Recouvreux, M.V., Ben-Shlomo, A., Araki, T., Barrett, R., Workman, M., Wawrowsky, K., Ljubimov, V.A., et al. (2016). Growth hormone is permissive for neoplastic colon growth. Proc. Natl. Acad. Sci. U. S. A. *113*, E3250–E3259.

Chhabra, Y., Waters, M.J., and Brooks, A.J. (2011). Role of the growth hormone-IGF-1 axis in cancer. Expert Rev. Endocrinol. Metab. *6*, 71–84.

Chhabra, Y., Wong, H.Y., Nikolajsen, L.F., Steinocher, H., Papadopulos, A., Tunny, K.A., Meunier, F.A., Smith, A.G., Kragelund, B.B., Brooks, A.J., et al. (2018). A growth hormone receptor SNP promotes lung cancer by impairment of SOCS2-mediated degradation. Oncogene *37*, 489–501.

Chi, S.W., Zang, J.B., Mele, A., and Darnell, R.B. (2009). Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature *460*, 479–486.

Chia, D.J. (2014). Mechanisms of growth hormone- mediated gene regulation. Mol. Endocrinol. *28*, 1012–1025.

Choi, M., Chang, C.-Y., Clough, T., Broudy, D., Killeen, T., Maclean, B., and Vitek, O. (2014). MSstats: an R package for statistical analysis of quantitative mass spectrometry-based proteomic experiments. *30*, 2524–2526.

Chou, R.H., Wang, Y.N., Hsieh, Y.H., Li, L.Y., Xia, W., Chang, W.C., Chang, L.C., Cheng, C.C., Lai, C.C., Hsu, J.L., et al. (2014). EGFR Modulates DNA Synthesis and Repair through Tyr Phosphorylation of Histone H4. Dev. Cell *30*, 224–237.

Ciabrelli, F., and Cavalli, G. (2015). Chromatin-Driven Behavior of Topologically Associating Domains. J. Mol. Biol. *427*, 608–625.

Clasen, B.F., Poulsen, M.M., Escande, C., Pedersen, S.B., Møller, N., Chini, E.N., Jessen, N., and Jørgensen, J.O.L. (2014). Growth hormone signaling in muscle and adipose tissue of obese human subjects: Associations with measures of body composition and interaction with resveratrol treatment. J. Clin. Endocrinol. Metab. *99*, E2565–E2573.

Clayton, P.E., Banerjee, I., Murray, P.G., and Renehan, A.G. (2011). Growth hormone, the insulin-like growth factor axis, insulin and cancer risk. Nat. Rev. Endocrinol. *7*, 11–24.

Clevenger, C. V., Plank, T.L., CV, C., and TL, P. (1997). Prolactin as an autocrine/paracrine factor in breast tissue. J. Mammary Gland Biol. Neoplasia *2*, 59–68.

Clevenger, C. V., Furth, P.A., Hankinson, S.E., and Schuler, L.A. (2003). The Role of Prolactin in Mammary Carcinoma. Endocr. Rev. *24*, 1–27.

Conte, F., Salles, J.P., Raynal, P., Fernandez, L., Molinas, C., Tauber, M., and Bieth, E. (2002). Identification of a region critical for proteolysis of the human growth hormone receptor. Biochem. Biophys. Res. Commun. *290*, 851–857.

Conway-Campbell, B.L., Jong, W.W., Brooks, A.J., Gordon, D., Brown, R.J., Lichanska, A.M., Hong, S.C., Barton, C.L., Boyle, G.M., Parsons, P.G., et al. (2007). Nuclear targeting of the growth hormone receptor results in dysregulation of cell proliferation and tumorigenesis. Proc. Natl. Acad. Sci. U. S. A. *104*, 13331–13336.

Conway-Campbell, B.L., Brooks, A.J., Robinson, P.J., Perani, M., and Waters, M.J. (2008). The Extracellular Domain of the Growth Hormone Receptor Interacts with Coactivator Activator to Promote Cell Proliferation. Mol. Endocrinol. *22*, 2190–2202.

Cremer, T., Kreth, G., Koester, H., Fink, R.H.A., Heintzmann, R., Cremer, M., Solovei, I., Zink, D., and Cremer, C. (2000). Chromosome territories, interchromatin domain compartment, and nuclear matrix: An integrated view of the functional nuclear architecture. Crit. Rev. Eukaryot. Gene Expr. *10*, 179–212.

Creyghton, M.P., Cheng, A.W., Welstead, G.G., Kooistra, T., Carey, B.W., Steine, E.J., Hanna, J., Lodato, M.A., Frampton, G.M., Sharp, P.A., et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc. Natl. Acad. Sci. U. S. A. *107*, 21931–21936.

Croft, J.A., Bridger, J.M., Boyle, S., Perry, P., Teague, P., and Bickmore, W.A. (1999). Differences in the localization and morphology of chromosomes in the human nucleus. J. Cell Biol. *145*, 1119–1131.

Crowley, W.R. (2015). Neuroendocrine regulation of lactation and milk production. Compr. Physiol. *5*, 255–291.

Daughaday, W.H., Rotwein, P., and Rotwein, P. (1989). Insulin-like growth factors I and II. Peptide, messenger ribonucleic acid and gene structures, serum, and tissue concentrations. Endocr. Rev. *10*, 68–91.

Davidson, E.H. (2010). Emerging properties of animal gene regulatory networks. Nature *468*, 911–920.

Dehkhoda, F., Lee, C.M.M., Medina, J., and Brooks, A.J. (2018). The Growth Hormone Receptor: Mechanism of Receptor Activation, Cell Signaling, and Physiological Aspects. Front. Endocrinol. (Lausanne). *9*, 35.

Dekker, J., and Heard, E. (2015). Structural and functional diversity of Topologically Associating Domains. FEBS Lett. *589*, 2877–2884.

Dekker, J., and Mirny, L. (2016). The 3D Genome as Moderator of Chromosomal Communication. Cell *164*, 1110–1121.

Dekker, J., Marti-Renom, M.A., and Mirny, L.A. (2013). Exploring the three-dimensional organization of genomes: Interpreting chromatin interaction data. Nat. Rev. Genet. *14*, 390–403.

Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., Guernec, G., Martin, D., Merkel, A., Knowles, D.G., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: Analysis of their gene structure, evolution, and expression. Genome Res. *22*, 1775–1789.

Devesa, J., Almengló, C., and Devesa, P. (2016). Multiple effects of growth hormone in the body: Is it really the hormone for growth? Clin. Med. Insights Endocrinol. Diabetes *9*, 47–71.

Deveson, I.W., Hardwick, S.A., Mercer, T.R., and Mattick, J.S. (2017). The Dimensions, Dynamics, and Relevance of the Mammalian Noncoding Transcriptome. Trends Genet. *33*, 464–478.

Le Dily, F., and Beato, M. (2018). Signaling by steroid hormones in the 3D nuclear space. Int. J. Mol. Sci. *19*, 306.

Le Dily, F., Vidal, E., Cuartero, Y., Quilez, J., Nacht, A.S., Vicent, G.P., Carbonell-Caballero, J., Sharma, P., Villanueva-Cañas, J.L., Ferrari, R., et al. (2019). Hormone-control regions mediate steroid receptor–dependent genome organization. Genome Res. *29*, 29–39.

Divisova, J., Kuiatse, I., Lazard, Z.W., Weiss, H., Vreeland, F., Hadsell, D.L., Schiff, R., Osborne, C.K., and Lee, A. V. (2006). The growth hormone receptor antagonist pegvisomant blocks both mammary gland development and MCF-7 breast cancer xenograft growth. Breast Cancer Res. Treat. *98*, 315–327.

Dixon, A.L., Liang, L., Moffatt, M.F., Chen, W., Heath, S., Wong, K.C.C., Taylor, J., Burnett, E., Gut, I., Farrall, M., et al. (2007). A genome-wide association study of global gene expression. Nat. Genet. *39*, 1202–1207.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376–380.

Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W., et al. (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331–336.

Dixon, J.R., Gorkin, D.U., and Ren, B. (2016). Chromatin Domains: The Unit of Chromosome Organization. Mol. Cell *62*, 668–680.

Djebali, S., Davis, C.A., Merkel, A., Dobin, A., Lassmann, T., Mortazavi, A., Tanzer, A., Lagarde, J., Lin, W., Schlesinger, F., et al. (2012). Landscape of transcription in human cells. Nature *489*, 101–108.

Dong, H., Wang, W., Mo, S., Chen, R., Zou, K., Han, J., Zhang, F., and Hu, J. (2018). SP1-induced lncRNA AGAP2-AS1 expression promotes chemoresistance of breast cancer by epigenetic regulation of MyD88. J. Exp. Clin. Cancer Res. *37*, 202.

Du, Q., Hoover, A.R., Dozmorov, I., Raj, P., Khan, S., Molina, E., Chang, T.C., de la Morena, M.T., Cleaver, O.B., Mendell, J.T., et al. (2019). MIR205HG Is a Long Noncoding RNA that Regulates Growth Hormone and Prolactin Production in the Anterior Pituitary. Dev. Cell *49*, 618–631.e5.

Duan, K., Ezzat, S., Asa, S.L., and Mete, O. (2015). Pancreatic Neuroendocrine Tumors Producing GHRH, GH, Ghrelin, PTH, or PTHrP. In Pancreatic Neuroendocrine Neoplasms, (Cham: Springer International Publishing), pp. 125–139.

Dunham, I., Kundaje, A., Aldred, S.F., Collins, P.J., Davis, C.A., Doyle, F., Epstein, C.B., Frietze, S., Harrow, J., Kaul, R., et al. (2012). An integrated encyclopedia of DNA elements in the human genome. Nature *489*, 57–74.

Ea, V., Baudement, M.O., Lesne, A., and Forné, T. (2015). Contribution of topological domains and loop formation to 3D chromatin organization. Genes (Basel). *6*, 734–750.

Ebert, M.S., and Sharp, P.A. (2012). Roles for MicroRNAs in conferring robustness to

biological processes. Cell *149*, 515–524.

Edwards, D.N., Ngwa, V.M., Wang, S., Shiuan, E., Brantley-Sieders, D., Kim, L.C., Reynolds, A.B., and Chen, J. (2018). Regulation of cancer glutamine metabolism by EphA2 RTK-dependent activation of transcriptional co-activators YAP and TAZ HHS Public Access.

van den Eijnden, M.J.M., Lahaye, L.L., and Strous, G.J. (2006). Disulfide bonds determine growth hormone receptor folding, dimerisation and ligand binding. J. Cell Sci. *119*, 3078–3086.

Eijsbouts, C.Q., Burren, O.S., Newcombe, P.J., and Wallace, C. (2019). Fine mapping chromatin contacts in capture Hi-C data. BMC Genomics *20*, 77.

Eilebrecht, S., Smet-Nocca, C., Wieruszeski, J.-M., and Benecke, A. (2010). SUMO-1 possesses DNA binding activity. BMC Res. Notes *3*, 146.

Elzein, S., and Goodyer, C.G. (2014). Regulation of Human Growth Hormone Receptor Expression by MicroRNAs. Mol. Endocrinol. *28*, 1448–1459.

Emilsson, V., Thorleifsson, G., Zhang, B., Leonardson, A.S., Zink, F., Zhu, J., Carlson, S., Helgason, A., Walters, G.B., Gunnarsdottir, S., et al. (2008). Genetics of gene expression and its effect on disease. Nature *452*, 423–428.

Esteller, M. (2011). Non-coding RNAs in human disease. Nat. Rev. Genet. *12*, 861–874.

Evans, A., Jamieson, S.M.F., Liu, D.X., Wilson, W.R., and Perry, J.K. (2016). Growth hormone receptor antagonism suppresses tumour regrowth after radiotherapy in an endometrial cancer xenograft model. Cancer Lett. *379*, 117–123.

Fadason, T., Ekblad, C., Ingram, J.R., Schierding, W.S., and O'Sullivan, J.M. (2017). Physical Interactions and Expression Quantitative Traits Loci Identify Regulatory Connections for Obesity and Type 2 Diabetes Associated SNPs. Front. Genet. *8*, 150.

Feng, W., and Michaels, S.D. (2015). Accessing the Inaccessible: The Organization, Transcription, Replication, and Repair of Heterochromatin in Plants. Annu. Rev. Genet. *49*, 439–459.

Figueiredo, M.A., Boyle, R.T., Sandrini, J.Z., Varela, A.S., and Marins, L.F. (2016). High level of GHR nuclear translocation in skeletal muscle of a hyperplasic transgenic zebrafish. J.

Mol. Endocrinol. *56*, 47–54.

Fiorillo, A.A., Medler, T.R., Feeney, Y.B., Liu, Y., Tommerdahl, K.L., and Clevenger, C. V. (2011). HMGN2 inducibly binds a novel transactivation domain in nuclear PRLr to coordinate stat5a-mediated transcription. Mol. Endocrinol. *25*, 1550–1564.

Firth, S.M., and Baxter, R.C. (2002). Cellular actions of the insulin-like growth factor binding proteins. Endocr. Rev. *23*, 824–854.

Fishilevich, S., Nudel, R., Rappaport, N., Hadar, R., Plaschkes, I., Iny Stein, T., Rosen, N., Kohn, A., Twik, M., Safran, M., et al. (2017). GeneHancer: genome-wide integration of enhancers and target genes in GeneCards. Database *2017*.

Flynn, R.A., and Chang, H.Y. (2014). Long noncoding RNAs in cell-fate programming and reprogramming. Cell Stem Cell *14*, 752–761.

Franz, M., Lopes, C.T., Huck, G., Dong, Y., Sumer, O., and Bader, G.D. (2016). Cytoscape.js: a graph theory library for visualisation and analysis. Bioinformatics *32*, 309–311.

Frick, M., Bettstetter, M., Bertz, S., Schwarz-Furlan, S., Hartmann, A., Richter, T., Lenze, D., Hummel, M., Dreyling, M., Lenz, G., et al. (2018). Mutational frequencies of CD79B and MYD88 vary greatly between primary testicular DLBCL and gastrointestinal DLBCL. Leuk. Lymphoma *59*, 1260–1263.

Frietsch, J.J., Grunewald, T.G.P., Jasper, S., Kammerer, U., Herterich, S., Kapp, M., Honig, A., and Butt, E. (2010). Nuclear localisation of LASP-1 correlates with poor long-term survival in female breast cancer. Br. J. Cancer *102*, 1645–1653.

Frith, M.C., Pheasant, M., and Mattick, J.S. (2005). The amazing complexity of the human transcriptome. Eur. J. Hum. Genet. *13*, 894–897.

Fukami, M., Tsuchiya, T., Takada, S., Kanbara, A., Asahara, H., Igarashi, A., Kamiyama, Y., Nishimura, G., and Ogata, T. (2012). Complex genomic rearrangement in the SOX9 5′ region in a patient with Pierre Robin sequence and hypoplastic left scapula. Am. J. Med. Genet. Part A *158 A*, 1529–1534.

Fukao, A., Mishima, Y., Takizawa, N., Oka, S., Imataka, H., Pelletier, J., Sonenberg, N., Thoma, C., and Fujiwara, T. (2014). MicroRNAs trigger dissociation of eIF4AI and eIF4AII from target mRNAs in humans. Mol. Cell *56*, 79–89.

Furlong, E.E.M., and Levine, M. (2018). Developmental enhancers and chromosome topology. Science (80-.). *361*, 1341–1345.

Gadelha, M.R., Kasuki, L., Lim, D.S.T., and Fleseriu, M. (2018). Systemic complications of acromegaly and the impact of the current treatment landscape: An update. Endocr. Rev. *40*, 268–332.

Galvez, T., Gilleron, J., Zerial, M., and O'Sullivan, G.A. (2012). SnapShot: Mammalian Rab Proteins in Endocytic Trafficking. Cell *151*, 234–234.e2.

Gan, Y., Buckels, A., Liu, Y., Zhang, Y., Paterson, A.J., Jiang, J., Zinn, K.R., and Frank, S.J. (2014). Human GH Receptor-IGF-1 Receptor Interaction: Implications for GH Signaling. Mol. Endocrinol. *28*, 1841–1854.

Ganguly, E., Bock, M.E., and Cattini, P.A. (2015). Expression of Placental Members of the Human Growth Hormone Gene Family Is Increased in Response to Sequential Inhibition of DNA Methylation and Histone Deacetylation. Biores. Open Access *4*, 446–456.

Gao, J., Ma, S., Yang, F., Chen, X., Wang, W., Zhang, J., Li, Y., Wang, T., and Shan, L. (2020). miR-193b exhibits mutual interaction with MYC, and suppresses growth and metastasis of osteosarcoma. Oncol. Rep. *44*.

Ge, G., Fernández, C.A., Moses, M.A., and Greenspan, D.S. (2007). Bone morphogenetic protein 1 processes prolactin to a 17-kDa antiangiogenic factor. Proc. Natl. Acad. Sci. U. S. A. *104*, 10010–10015.

Geary, M.P.P., Pringle, P.J., Rodeck, C.H., Kingdom, J.C.P., and Hindmarsh, P.C. (2003). Sexual Dimorphism in the Growth Hormone and Insulin-Like Growth Factor Axis at Birth. J. Clin. Endocrinol. Metab. *88*, 3708–3714.

Gendrel, A.-V., and Heard, E. (2014). Noncoding RNAs and Epigenetic Mechanisms During X-Chromosome Inactivation. Annu. Rev. Cell Dev. Biol. *30*, 561–580.

Gesing, A., Wiesenborn, D., Do, A., Menon, V., Schneider, A., Victoria, B., Stout, M.B., Kopchick, J.J., Bartke, A., and Masternak, M.M. (2017). A Long-lived Mouse Lacking Both Growth Hormone and Growth Hormone Receptor: A New Animal Model for Aging Studies. Journals Gerontol. - Ser. A Biol. Sci. Med. Sci. *72*, 1054–1061.

Gibcus, J.H., and Dekker, J. (2013). The Hierarchy of the 3D Genome. Mol. Cell *49*, 773–782.

Giustina, A., Mazziotti, G., and Canalis, E. (2008). Growth hormone, insulin-like growth factors, and the skeleton. Endocr. Rev. *29*, 535–559.

Godowski, P.J., Leung, D.W., Meacham, L.R., Galgani, J.P., Hellmiss, R., Keret, R., Rotwein, P.S., Parks, J.S., Laron, Z., and Wood, W.I. (1989). Characterization of the human growth hormone receptor gene and demonstration of a partial gene deletion in two patients with Laron-type dwarfism. Proc. Natl. Acad. Sci. U. S. A. *86*, 8083–8087.

Goffin, V., Shiverick, K.T., Kelly, P.A., and Martial, J.A. (1996). Sequence-Function Relationships Within the Expanding Family of Prolactin, Growth Hormone, Placental Lactogen, and Related Proteins in Mammals*. Endocr. Rev. *17*, 385–410.

Goffin, V., Bouchard, B., Ormandy, C.J., Weimann, E., Ferrag, F., Touraine, P., Bole-Feysot, C., Maaskant, R.A., Clement-Lacroix, P., Edery, M., et al. (1998). Prolactin: A hormone at the crossroads of neuroimmunoendocrinology. In Annals of the New York Academy of Sciences, (Blackwell Publishing Inc.), pp. 498–509.

Goffin, V., Hoang, D.T., Bogorad, R.L., and Nevalainen, M.T. (2011). Prolactin regulation of the prostate gland: A female player in a male game. Nat. Rev. Urol. *8*, 597–607.

Gokuladhas, S., Schierding, W., Cameron-Smith, D., Wake, M., Scotter, E.L., and O'Sullivan, J. (2020). Shared Regulatory Pathways Reveal Novel Genetic Correlations Between Grip Strength and Neuromuscular Disorders. Front. Genet. *11*, 393.

Gómez-Rubio, V. (2017). ggplot2 - Elegant Graphics for Data Analysis (2nd Edition). J. Stat. Softw. *77*, 3–5.

González Alvarez, R., Revol de Mendoza, A., Esquivel Escobedo, D., Corrales Félix, G., Rodríguez Sánchez, I., González, V., Dávila, G., Cao, Q., de Jong, P., Fu, Y.-X., et al. (2006). Growth hormone locus expands and diverges after the separation of New and Old World Monkeys. Gene *380*, 38–45.

Grattan, D.R. (2015). The hypothalamo-prolactin axis. J. Endocrinol. *226*, T101–T122.

Greenhalgh, C.J., Bertolino, P., Asa, S.L., Metcalf, D., Corbin, J.E., Adams, T.E., Davey, H.W., Nicola, N.A., Hilton, D.J., and Alexander, W.S. (2002). Growth enhancement in suppressor of cytokine signaling 2 (SOCS-2)-deficient mice is dependent on signal transducer and activator of transcription 5b (STAT5b). Mol. Endocrinol. *16*, 1394–1406.

Greenhalgh, C.J., Rico-Bautista, E., Lorentzon, M., Thaus, A.L., Morgan, P.O., Willson,

T.A., Zervoudakis, P., Metcalf, D., Street, I., Nicola, N.A., et al. (2005). SOCS2 negatively regulates growth hormone action in vitro and in vivo. J. Clin. Invest. *115*, 397–406.

Grimm, M.O.W., Lauer, A.A., Grösgen, S., Thiel, A., Lehmann, J., Winkler, J., Janitschke, D., Herr, C., Beisswenger, C., Bals, R., et al. (2019). Profiling of Alzheimer's disease related genes in mild to moderate vitamin D hypovitaminosis. J. Nutr. Biochem. *67*, 123–137.

Gross, D.S., and Garrard, W.T. (1988). Nuclease Hypersensitive Sites in Chromatin. Annu. Rev. Biochem. *57*, 159–197.

Grundberg, E., Small, K.S., Hedman, Å.K., Nica, A.C., Buil, A., Keildson, S., Bell, J.T., Yang, T.P., Meduri, E., Barrett, A., et al. (2012). Mapping cis-and trans-regulatory effects across multiple tissues in twins. Nat. Genet. *44*, 1084–1089.

Gu, Z., Gu, L., Eils, R., Schlesner, M., and Brors, B. (2014). Circlize implements and enhances circular visualization in R. Bioinformatics *30*, 2811–2812.

Gu, Z., Eils, R., and Schlesner, M. (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics *32*, 2847–2849.

Guevara-Aguirre, J., Balasubramanian, P., Guevara-Aguirre, M., Wei, M., Madia, F., Cheng, C.W., Hwang, D., Martin-Montalvo, A., Saavedra, J., Ingles, S., et al. (2011). Growth hormone receptor deficiency is associated with a major reduction in pro-aging signaling, cancer, and diabetes in humans. Sci. Transl. Med. *3*, 70ra13-70ra13.

Guevara-Aguirre, J., Guevara, A., Palacios, I., Pérez, M., Prócel, P., and Terán, E. (2018). GH and GHR signaling in human disease. Growth Horm. IGF Res. *38*, 34–38.

Guo, G., Kang, Q., Zhu, X., Chen, Q., Wang, X., Chen, Y., Ouyang, J., Zhang, L., Tan, H., Chen, R., et al. (2014). A long noncoding RNA critically regulates Bcr-Abl-mediated cellular transformation by acting as a competitive endogenous RNA. Oncogene *34*, 1768–1779.

Guo, H., Ingolia, N.T., Weissman, J.S., and Bartel, D.P. (2010). Mammalian microRNAs predominantly act to decrease target mRNA levels. Nature *466*, 835–840.

Gupta, R.A., Shah, N., Wang, K.C., Kim, J., Horlings, H.M., Wong, D.J., Tsai, M.C., Hung, T., Argani, P., Rinn, J.L., et al. (2010). Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. Nature *464*, 1071–1076.

Haddad, N., Jost, D., and Vaillant, C. (2017). Perspectives: using polymer modeling to

understand the formation and function of nuclear compartments. Chromosom. Res. *25*, 35–50.

Hahn, M., Dambacher, S., and Schotta, G. (2010). Heterochromatin dysregulation in human diseases. J. Appl. Physiol. *109*, 232–242.

Hakuno, F., and Takahashi, S.I. (2018). 40 years of IGF1: IGF1 receptor signaling pathways. J. Mol. Endocrinol. *61*, T69–T86.

Hanly, D.J., Esteller, M., and Berdasco, M. (2018). Interplay between long non-coding RNAs and epigenetic machinery: Emerging targets in cancer? Philos. Trans. R. Soc. B Biol. Sci. *373*.

Hannon, A.M., Thompson, C.J., and Sherlock, M. (2017). Diabetes in Patients With Acromegaly. Curr. Diab. Rep. *17*, 8.

Harrow, J., Frankish, A., Gonzalez, J.M., Tapanari, E., Diekhans, M., Kokocinski, F., Aken, B.L., Barrell, D., Zadissa, A., Searle, S., et al. (2012). GENCODE: The reference human genome annotation for the ENCODE project. Genome Res. *22*, 1760–1774.

Harvey, P.W., Everett, D.J., and Springall, C.J. (2008). Adverse effects of prolactin in rodents and humans: Breast and prostate cancer. J. Psychopharmacol. *22*, 20–27.

Harvey, S., Martínez-Moreno, C.G., Luna, M., and Arámburo, C. (2015). Autocrine/paracrine roles of extrapituitary growth hormone and prolactin in health and disease: An overview. Gen. Comp. Endocrinol. *220*, 103–111.

Hayashi, A.A., and Proud, C.G. (2007). The rapid activation of protein synthesis by growth hormone requires signaling through mTOR. Am. J. Physiol. - Endocrinol. Metab. *292*, E1647–E1655.

He, X., Duque, T.S.P.C., and Sinha, S. (2012). Evolutionary origins of transcription factor binding site clusters. Mol. Biol. Evol. *29*, 1059–1070.

Hermann, M., and Berger, P. (2001). Hormonal changes in aging men: A therapeutic indication? Exp. Gerontol. *36*, 1075–1082.

Hindmarsh, P.C., Dennison, E., Pincus, S.M., Cooper, C., Fall, C.H.D., Matthews, D.R., Pringle, P.J., and Brook, C.G.D. (1999). A Sexually Dimorphic Pattern of Growth Hormone Secretion in the Elderly. J. Clin. Endocrinol. Metab. *84*, 2679–2685.

Hirschhorn, J.N., and Gajdos, Z.K.Z. (2011). Genome-Wide Association Studies: Results from the First Few Years and Potential Implications for Clinical Medicine. Annu. Rev. Med. *62*, 11–24.

Hirst, J., Edgar, J.R., Borner, G.H.H., Li, S., Sahlender, D.A., Antrobus, R., and Robinson, M.S. (2015). Contributions of epsinR and gadkin to clathrin-mediated intracellular trafficking. Mol. Biol. Cell *26*, 3085–3103.

Ho, Y., Elefant, F., Cooke, N., and Liebhaber, S. (2002). A defined locus control region determinant links chromatin domain acetylation with long-range gene activation. Mol. Cell *9*, 291–302.

Ho, Y., Liebhaber, S.A., and Cooke, N.E. (2004). Activation of the human GH gene cluster: Roles for targeted chromatin modification. Trends Endocrinol. Metab. *15*, 40–45.

Ho, Y., Tadevosyan, A., Liebhaber, S.A., and Cooke, N.E. (2008). The juxtaposition of a promoter with a locus control region transcriptional domain activates gene expression. EMBO Rep. *9*, 891–898.

Hornstein, E., and Shomron, N. (2006). Canalization of development by micrornas. Nat. Genet. *38*, S20.

Housman, G., and Ulitsky, I. (2016). Methods for distinguishing between protein-coding and long noncoding RNAs and the elusive biological purpose of translation of long noncoding RNAs. Biochim. Biophys. Acta - Gene Regul. Mech. *1859*, 31–40.

Hu, Z., Wang, X., Cui, Y., Li, C., and Wang, S. (2018). LASP1 in tumor and tumor microenvironment. Curr. Mol. Med. *18*.

Huang, Y., Kim, S.-O., Yang, N., Jiang, J., and Frank, S.J. (2004). Physical and Functional Interaction of Growth Hormone and Insulin-Like Growth Factor-I Signaling Elements. Mol. Endocrinol. *18*, 1471–1485.

Hung, L.-Y., Tseng, J.T., Lee, Y.-C., Xia, W., Wang, Y.-N., Wu, M.-L., Chuang, Y.-H., Lai, C.-H., and Chang, W.-C. (2008). Nuclear epidermal growth factor receptor (EGFR) interacts with signal transducer and activator of transcription 5 (STAT5) in activating Aurora-A gene expression. Nucleic Acids Res. *36*, 4337–4351.

Irizarry, R.A., Hobbs, B., Collin, F., Beazer-Barclay, Y.D., Antonellis, K.J., Scherf, U., and Speed, T.P. (2003). Exploration, normalization, and summaries of high density

oligonucleotide array probe level data. Biostatistics *4*, 249–264.

Jennings, B.H., and Ish-Horowicz, D. (2008). The Groucho/TLE/Grg family of transcriptional co-repressors. Genome Biol. *9*, 205.

Ji, X., Dadon, D.B., Powell, B.E., Fan, Z.P., Borges-Rivera, D., Shachar, S., Weintraub, A.S., Hnisz, D., Pegoraro, G., Lee, T.I., et al. (2016). 3D Chromosome Regulatory Landscape of Human Pluripotent Cells. Cell Stem Cell *18*, 262–275.

Jin, F., Li, Y., Dixon, J.R., Selvaraj, S., Ye, Z., Lee, A.Y., Yen, C.A., Schmitt, A.D., Espinoza, C.A., and Ren, B. (2013). A high-resolution map of the three-dimensional chromatin interactome in human cells. Nature *503*, 290–294.

Jin, J., Ravindran, P., Di Meo, D., and Püschel, A.W. (2019). Igf1R/InsR function is required for axon extension and corpus callosum formation. PLoS One *14*, e0219362.

Jonas, S., and Izaurralde, E. (2015). Towards a molecular understanding of microRNA-mediated gene silencing. Nat. Rev. Genet. *16*, 421–433.

Kai, M. (2016). Roles of RNA-Binding Proteins in DNA Damage Response. Int. J. Mol. Sci. *17*, 310.

Kannan, S., and Kennedy, L. (2013). Diagnosis of acromegaly: State of the art. Expert Opin. Med. Diagn. *7*, 443–453.

Karreth, F.A., and Pandolfi, P.P. (2013). CeRNA cross-talk in cancer: When ce-bling rivalries go awry. Cancer Discov. *3*, 1113–1121.

Ke, C., Zhu, K., Sun, Y., Ni, Y., Zhang, Z., and Li, X. (2019). SUMO1 promotes the proliferation and invasion of non-small cell lung cancer cells by regulating NF-κB. Thorac. Cancer *10*, 33–40.

Kehl, T., Schneider, L., Schmidt, F., Stöckel, D., Gerstner, N., Backes, C., Meese, E., Keller, A., Schulz, M.H., and Lenhof, H.-P. (2017). RegulatorTrail: a web service for the identification of key transcriptional regulators. Nucleic Acids Res. *45*, W146–W153.

Kehrer-Sawatzki, H., Mautner, V.F., and Cooper, D.N. (2017). Emerging genotype–phenotype relationships in patients with large NF1 deletions. Hum. Genet. *136*, 349–376.

Kellis, M., Wold, B., Snyder, M.P., Bernstein, B.E., Kundaje, A., Marinov, G.K., Ward, L.D., Birney, E., Crawford, G.E., Dekker, J., et al. (2014). Defining functional DNA elements in

the human genome. Proc. Natl. Acad. Sci. U. S. A. *111*, 6131–6138.

Van Kerkhof, P., Govers, R., Alves Dos Santos, C.M., and Strous, G.J. (2000). Endocytosis and degradation of the growth hormone receptor are proteasome-dependent. J. Biol. Chem. *275*, 1575–1580.

Kikuchi, M., Hara, N., Hasegawa, M., Miyashita, A., Kuwano, R., Ikeuchi, T., and Nakaya, A. (2019). Enhancer variants associated with Alzheimer's disease affect gene expression via chromatin looping. BMC Med. Genomics *12*.

Kilpinen, H., Waszak, S.M., Gschwind, A.R., Raghav, S.K., Witwicki, R.M., Orioli, A., Migliavacca, E., Wiederkehr, M., Gutierrez-Arcelus, M., Panousis, N.I., et al. (2013). Coordinated effects of sequence variation on DNA binding, chromatin structure, and transcription. Science (80-.). *342*, 744–747.

Kim, S.J. (2003). Placental site trophoblastic tumour. Bailliere's Best Pract. Res. Clin. Obstet. Gynaecol. *17*, 969–984.

Kim, Y.D., Li, T., Ahn, S.W., Kim, D.K., Lee, J.M., Hwang, S.L., Kim, Y.H., Lee, C.H., Lee, I.K., Chiang, J.Y.L., et al. (2012). Orphan nuclear receptor small heterodimer partner negatively regulates growth hormone-mediated induction of hepatic gluconeogenesis through inhibition of signal transducer and activator of transcription 5 (STAT5) transactivation. J. Biol. Chem. *287*, 37098–37108.

Kimura, A.P., Sizova, D., Handwerger, S., Cooke, N.E., and Liebhaber, S.A. (2007). Epigenetic Activation of the Human Growth Hormone Gene Cluster during Placental Cytotrophoblast Differentiation. Mol. Cell. Biol. *27*, 6555–6568.

Kino, T., Hurt, D.E., Ichijo, T., Nader, N., and Chrousos, G.P. (2010). Noncoding RNA Gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. Sci. Signal. *3*.

Koch, C.M., Andrews, R.M., Flicek, P., Dillon, S.C., Karaöz, U., Clelland, G.K., Wilcox, S., Beare, D.M., Fowler, J.C., Couttet, P., et al. (2007). The landscape of histone modifications across 1% of the human genome in five human cell lines. Genome Res. *17*, 691–707.

Kong, S., and Zhang, Y. (2019). Deciphering Hi-C: from 3D genome to function. Cell Biol. Toxicol. *35*, 15–32.

Kong, X., Wu, W., Yuan, Y., Pandey, V., Wu, Z., Lu, X., Zhang, W., Chen, Y., Wu, M.,

Zhang, M., et al. (2016). Human growth hormone and human prolactin function as autocrine/paracrine promoters of progression of hepatocellular carcinoma. Oncotarget *7*, 29465–29479.

Kostopoulou, E., Rojas-Gil, A.P., Karvela, A., and Spiliotis, B.E. (2017). Epidermal growth factor receptor (EGFR) involvement in successful growth hormone (GH) signaling in GH transduction defect. J. Pediatr. Endocrinol. Metab. *30*, 221–230.

Krijger, P.H.L., and De Laat, W. (2016). Regulation of disease-associated gene expression in the 3D genome. Nat. Rev. Mol. Cell Biol. *17*, 771–782.

Křížková, K., Chrudinová, M., Povalová, A., Selicharová, I., Collinsová, M., Vaněk, V., Brzozowski, A.M., Jiráček, J., and Zakova, L. (2016). Insulin-Insulin-like Growth Factors Hybrids as Molecular Probes of Hormone:Receptor Binding Specificity. Biochemistry *55*, 2903–2913.

Lambert, S.A., Jolma, A., Campitelli, L.F., Das, P.K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T.R., and Weirauch, M.T. (2018). The Human Transcription Factors. Cell *172*, 650–665.

Lan, H., Li, W., Li, R., Zheng, X., and Luo, G. (2019). Endocytosis and Degradation of Pegvisomant and a Potential New Mechanism That Inhibits the Nuclear Translocation of GHR. J. Clin. Endocrinol. Metab. *104*, 1887–1899.

Lan, H.N., Hong, P., Li, R.N., Shan, A.S., and Zheng, X. (2017). Growth hormone-specific induction of the nuclear localization of porcine growth hormone receptor in porcine hepatocytes. Domest. Anim. Endocrinol. *61*, 39–47.

Lanctôt, C., Cheutin, T., Cremer, M., Cavalli, G., and Cremer, T. (2007). Dynamic genome architecture in the nuclear space: Regulation of gene expression in three dimensions. Nat. Rev. Genet. *8*, 104–115.

Laron, Z., and Kauli, R. (2016). Fifty seven years of follow-up of the Israeli cohort of Laron Syndrome patients-From discovery to treatment. Growth Horm. IGF Res. *28*, 53–56.

Laskey, R.A., Honda, B.M., Mills, A.D., and Finch, J.T. (1978). Nucleosomes are assembled by an acidic protein which binds histones and transfers them to DNA. Nature *275*, 416–420.

Lee, H., Han, S., Kwon, C.S., and Lee, D. (2016). Biogenesis and regulation of the let-7 miRNAs and their functional implications. Protein Cell *7*, 100–113.

Lee, R.C., Feinbaum, R.L., Ambros, V., and RC Lee, R.F.V.A. (1993). The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell *75*, 843–854.

Lemon, B., and Tjian, R. (2000). Orchestrated response: A symphony of transcription factors for gene control. Genes Dev. *14*, 2551–2569.

LeRoith, D., and Yakar, S. (2007). Mechanisms of disease: Metabolic effects of growth hormone and insulin-like growth factor 1. Nat. Clin. Pract. Endocrinol. Metab. *3*, 302–310.

Lettice, L.A., Heaney, S.J.H., Purdie, L.A., Li, L., De Beer, P., Oostra, B.A., Goode, D., Elgar, G., Hill, R.E., and De Graaff, E. (2003). A long-range Shh enhancer regulates expression in the developing limb and fin and is associated with preaxial polydactyly. Hum. Mol. Genet. *12*, 1725–1735.

Li, J., and Liu, C. (2019). Coding or noncoding, the converging concepts of RNAs. Front. Genet. *10*, 496.

Li, C., Iida, M., Dunn, E.F., Ghia, A.J., and Wheeler, D.L. (2009). Nuclear EGFR contributes to acquired resistance to cetuximab. Oncogene *28*, 3801–3813.

Li, W., Sun, M., Zang, C., Ma, P., He, J., Zhang, M., Huang, Z., Ding, Y., and Shu, Y. (2016). Upregulated long non-coding RNA AGAP2-AS1 represses LATS2 and KLF2 expression through interacting with EZH2 and LSD1 in non-small-cell lung cancer cells. Cell Death Dis. *7*, e2225–e2225.

Li, X., Huang, Y., Jiang, J., and Frank, S.J. (2008). ERK-dependent threonine phosphorylation of EGF receptor modulates receptor downregulation and signaling. Cell. Signal. *20*, 2145–2155.

Liang, W.Q., De Zeng, Chen, C.F., Sun, S.M., Lu, X.F., Peng, C.Y., and Lin, H.Y. (2019). Long noncoding RNA H19 is a critical oncogenic driver and contributes to epithelial-mesenchymal transition in papillary thyroid carcinoma. Cancer Manag. Res. *11*, 2059–2072.

Liao, S., Vickers, M.H., Stanley, J.L., Baker, P.N., and Perry, J.K. (2018). Human Placental Growth Hormone Variant in Pathological Pregnancies. Endocrinology *159*, 2186–2198.

Liccardi, G., Hartley, J.A., and Hochhauser, D. (2011). EGFR nuclear translocation modulates DNA repair following cisplatin and ionizing radiation treatment. Cancer Res. *71*, 1103–1114.

Lichanska, A.M., and Waters, M.J. (2008). How growth hormone controls growth, obesity and sexual dimorphism. Trends Genet. *24*, 41–47.

Lieberman-Aiden, E., Van Berkum, N.L., Williams, L., Imakaev, M., Ragoczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., et al. (2009). Comprehensive mapping of long-range interactions reveals folding principles of the human genome. Science (80-.). *326*, 289–293.

Lim, J.H., Catez, F., Birger, Y., West, K.L., Prymakowska-Bosak, M., Postnikov, Y. V., and Bustin, M. (2004). Chromosomal protein HMGN1 modulates histone H3 phosphorylation. Mol. Cell *15*, 573–584.

Lin, P.Y., Yu, S.L., and Yang, P.C. (2010). MicroRNA in lung cancer. Br. J. Cancer *103*, 1144–1148.

Lin, S.Y., Makino, K., Xia, W., Matin, A., Wen, Y., Kwong, K.Y., Bourguignon, L., and Hung, M.C. (2001). Nuclear localization of EGF receptor and its potential new role as a transcription factor. Nat. Cell Biol. *3*, 802–808.

Lin, Y., Liu, H., Waraky, A., Haglund, F., Agarwal, P., Jernberg-Wiklund, H., Warsito, D., and Larsson, O. (2017). SUMO-modified insulin-like growth factor 1 receptor (IGF-1R) increases cell cycle progression and cell proliferation. J. Cell. Physiol. *232*, 2722–2730.

Lincoln, D.T., Sinowatz, F., Temmim-Baker, L., Baker, H.I., Kölle, S., and Waters, M.J. (1998). Growth hormone receptor expression in the nucleus and cytoplasm of normal and neoplastic cells. Histochem. Cell Biol. *109*, 141–159.

Liu, G.Y., and Sabatini, D.M. (2020). mTOR at the nexus of nutrition, growth, ageing and disease. Nat. Rev. Mol. Cell Biol.

Liu, C., Zhang, N., Yu, H., Chen, Y., Liang, Y., Deng, H., and Zhang, Z. (2011). Proteomic analysis of human serum for Finding pathogenic factors and potential biomarkers in preeclampsia. Placenta *32*, 168–174.

Liu, X., Chen, X., Yu, X., Tao, Y., Bode, A.M., Dong, Z., and Cao, Y. (2013). Regulation of microRNAs by epigenetics and their interplay involved in cancer. J. Exp. Clin. Cancer Res. *32*, 1–8.

Liu, Y., Zhang, Y., Jiang, J., Lobie, P.E., Paulmurugan, R., Langenheim, J.F., Chen, W.Y., Zinn, K.R., and Frank, S.J. (2016a). GHR/PRLR heteromultimer is composed of GHR

homodimers and PRLR homodimers. Mol. Endocrinol. *30*, 504–517.

Liu, Z., Mohan, S., and Yakar, S. (2016b). Does the GH/IGF-1 axis contribute to skeletal sexual dimorphism? Evidence from mouse studies. Growth Horm. IGF Res. *27*, 7–17.

Liu, Z., Wang, Y., Wang, L., Yao, B., Sun, L., Liu, R., Chen, T., Niu, Y., Tu, K., and Liu, Q. (2019). Long non-coding RNA AGAP2-AS1, functioning as a competitive endogenous RNA, upregulates ANXA11 expression by sponging miR-16-5p and promotes proliferation and metastasis in hepatocellular carcinoma. J. Exp. Clin. Cancer Res. *38*, 1–15.

Lizio, M., Harshbarger, J., Shimoji, H., Severin, J., Kasukawa, T., Sahin, S., Abugessaisa, I., Fukuda, S., Hori, F., Ishikawa-Kato, S., et al. (2015). Gateways to the FANTOM5 promoter level mammalian expression atlas. Genome Biol. *16*, 22.

Lo, H.W., Hsu, S.C., Ali-Seyed, M., Gunduz, M., Xia, W., Wei, Y., Bartholomeusz, G., Shih, J.Y., and Hung, M.C. (2005). Nuclear interaction of EGFR and STAT3 in the activation of the iNOS/NO pathway. Cancer Cell *7*, 575–589.

Lobie, P.E., García-Aragón, J., Wang, B.S., Baumbach, W.R., and Waters, M.J. (1992). Cellular localization of the growth hormone binding protein in the rat. Endocrinology *130*, 3057–3065.

Lobie, P.E., Wood, T.J.J., Chen, C.M., Waters, M.J., and Norstedt, G. (1994a). Nuclear translocation and anchorage of the growth hormone receptor. J. Biol. Chem. *269*, 31735–31746.

Lobie, P.E., Mertani, H., Morel, G., Morales-Bustos, O., Norstedt, G., and Waters, M.J. (1994b). Receptor-mediated nuclear translocation of growth hormone. J. Biol. Chem. *269*, 21330–21339.

Lobie, P.E., Sadir, R., Graichen, R., Mertani, H.C., and Morel, G. (1999). Caveolar internalization of growth hormone. Exp. Cell Res. *246*, 47–55.

Loda, A., and Heard, E. (2019). Xist RNA in action: Past, present, and future. PLoS Genet. *15*.

Lonfat, N., and Duboule, D. (2015). Structure, function and evolution of topologically associating domains (TADs) at HOX loci. FEBS Lett. *589*, 2869–2876.

Lu, M., Flanagan, J.U., Langley, R.J., Hay, M.P., and Perry, J.K. (2019). Targeting growth

hormone function: strategies and therapeutic applications. Signal Transduct. Target. Ther. *4*, 3.

Lu, W., Zhang, H., Niu, Y., Wu, Y., Sun, W., Li, H., Kong, J., Ding, K., Shen, H.M., Wu, H., et al. (2017). Long non-coding RNA linc00673 regulated non-small cell lung cancer proliferation, migration, invasion and epithelial mesenchymal transition by sponging miR-150-5p. Mol. Cancer *16*, 118.

Lü, L., Liu, T., Gao, J., Zeng, H., Chen, J., Gu, X., and Mei, Z. (2016). Aberrant methylation of microRNA-193b in human Barrett's esophagus and esophageal adenocarcinoma. Mol. Med. Rep. *14*, 283–288.

Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. Nature *389*, 251–260.

Lund, E., Güttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear export of microRNA precursors. Science *303*, 95–98.

Luo, W., Li, X., Song, Z., Zhu, X., and Zhao, S. (2019). Long non-coding RNA AGAP2-AS1 exerts oncogenic properties in glioblastoma by epigenetically silencing TFPI2 through EZH2 and LSD1. Aging (Albany. NY). *11*, 3811–3823.

Ma, W., Ay, F., Lee, C., Gulsoy, G., Deng, X., Cook, S., Hesson, J., Cavanaugh, C., Ware, C.B., Krumm, A., et al. (2014). Fine-scale chromatin interaction maps reveal the cis-regulatory landscape of human lincRNA genes. Nat. Methods *12*, 71–78.

Maas, S. (2012). Posttranscriptional recoding by RNA editing. In Advances in Protein Chemistry and Structural Biology, (Academic Press Inc.), pp. 193–224.

Maher, P.A. (1996). Nuclear translocation of fibroblast growth factor (FGF) receptors in response to FGF-2. J. Cell Biol. *134*, 529–536.

Männik, J., Vaas, P., Rull, K., Teesalu, P., Rebane, T., and Laan, M. (2010). Differential expression profile of Growth Hormone/Chorionic Somatomammotropin genes in placenta of small- and large-for-gestational-age newborns. J. Clin. Endocrinol. Metab. *95*, 2433–2442.

Manolio, T.A. (2010). Genomewide association studies and assessment of the risk of disease. N. Engl. J. Med. *363*, 166–176.

Marano, R.J., and Ben-Jonathan, N. (2014). Extrapituitary Prolactin: An Update on the

Distribution, Regulation, and Functions. Mol. Endocrinol. *28*, 622–633.

Marsman, J., and Horsfield, J.A. (2012). Long distance relationships: Enhancer-promoter communication and dynamic gene transcription. Biochim. Biophys. Acta - Gene Regul. Mech. *1819*, 1217–1227.

Marti, U., Ruchti, C., Kämpf, J., Thomas, G.A., Williams, E.D., Peter, H.J., Gerber, H., and Bürgi, U. (2001). Nuclear localization of epidermal growth factor and epidermal growth factor receptors in human thyroid tissues. Thyroid *11*, 137–145.

Martirosyan, A., De Martino, A., Pagnani, A., and Marinari, E. (2017a). ceRNA crosstalk stabilizes protein expression and affects the correlation pattern of interacting proteins. Sci. Rep. *7*, 43673.

Martirosyan, A., Marsili, M., and De Martino, A. (2017b). Translating ceRNA Susceptibilities into Correlation Functions. Biophys. J. *113*, 206–213.

Massone, S., Ciarlo, E., Vella, S., Nizzari, M., Florio, T., Russo, C., Cancedda, R., and Pagano, A. (2012). NDM29, a RNA polymerase III-dependent non coding RNA, promotes amyloidogenic processing of APP and amyloid β secretion. *1823*, 1170–1177.

Mazzu, Y.Z., Yoshikawa, Y., Nandakumar, S., Chakraborty, G., Armenia, J., Jehane, L.E., Lee, G.M., and Kantoff, P.W. (2019a). Methylation-associated miR-193b silencing activates master drivers of aggressive prostate cancer. Mol. Oncol. *13*, 1944–1958.

Mazzu, Y.Z., Hu, Y., Shen, Y., Tuschl, T., and Singer, S. (2019b). miR-193b regulates tumorigenesis in liposarcoma cells via PDGFR, TGFβ, and Wnt signaling. Sci. Rep. *9*, 1–11.

McGee, S.L., and Hargreaves, M. (2011). Histone modifications and exercise adaptations. J. Appl. Physiol. *110*, 258–263.

McMahon, H.T., and Boucrot, E. (2011). Molecular mechanism and physiological functions of clathrin-mediated endocytosis. Nat. Rev. Mol. Cell Biol. *12*, 517–533.

Meireles-Filho, A.C., and Stark, A. (2009). Comparative genomics of gene regulation-conservation and divergence of cis-regulatory information. Curr. Opin. Genet. Dev. *19*, 565–570.

Melia, T., Hao, P., Yilmaz, F., and Waxman, D.J. (2015). Hepatic lincRNAs: high promoter conservation and dynamic, sex-dependent transcriptional regulation by growth hormone.

Mol. Cell. Biol. *36*, MCB.00861-15.

Melmed, S. (2006). Acromegaly. N. Engl. J. Med. *355*, 2558–2573.

Merkin, J., Russell, C., Chen, P., and Burge, C.B. (2012). Evolutionary dynamics of gene and isoform regulation in mammalian tissues. Science (80-.). *338*, 1593–1599.

Mertani, H.C., Garcia-Caballero, T., Lambert, A., Gérard, F., Palayer, C., Boutin, J.M., Vonderhaar, B.K., Waters, M.J., Lobie, P.E., and Morel, G. (1998). Cellular expression of growth hormone and prolactin receptors in human breast disorders. Int. J. Cancer *79*, 202–211.

Mestdagh, P., Boström, A.K., Impens, F., Fredlund, E., Van Peer, G., De Antonellis, P., von Stedingk, K., Ghesquière, B., Schulte, S., Dews, M., et al. (2010). The miR-17-92 MicroRNA Cluster Regulates Multiple Components of the TGF-β Pathway in Neuroblastoma. Mol. Cell *40*, 762–773.

Miotto, M., Marinari, E., and De Martino, A. (2019). Competing endogenous RNA crosstalk at system level. PLoS Comput. Biol. *15*, e1007474.

Mirabello, L., Yu, K., Berndt, S.I., Burdett, L., Wang, Z., Chowdhury, S., Teshome, K., Uzoka, A., Hutchinson, A., Grotmol, T., et al. (2011). A comprehensive candidate gene approach identifies genetic variation associated with osteosarcoma. BMC Cancer *11*, 209.

Møller, N., Gormsen, L.C., Schmitz, O., Lund, S., Jørgensen, J.O.L., and Jessen, N. (2009). Free fatty acids inhibit growth hormone/signal transducer and activator of transcription-5 signaling in human muscle: A potential feedback mechanism. J. Clin. Endocrinol. Metab. *94*, 2204–2207.

Mondal, T., Subhash, S., Vaid, R., Enroth, S., Uday, S., Reinius, B., Mitra, S., Mohammed, A., James, A.R., Hoberg, E., et al. (2015). MEG3 long noncoding RNA regulates the TGF-β pathway genes through formation of RNA–DNA triplex structures. Nat. Commun. *6*, 7743.

Montgomery, D.W. (2001). Prolactin production by immune cells. Lupus *10*, 665–675.

Morán, I., Akerman, I., Van De Bunt, M., Xie, R., Benazra, M., Nammo, T., Arnes, L., Nakić, N., García-Hurtado, J., Rodríguez-Seguí, S., et al. (2012). Human β cell transcriptome analysis uncovers lncRNAs that are tissue-specific, dynamically regulated, and abnormally expressed in type 2 diabetes. Cell Metab. *16*, 435–448.

Mulrane, L., McGee, S.F., Gallagher, W.M., and O'Connor, D.P. (2013). miRNA dysregulation in breast cancer. Cancer Res. *73*, 6554–6562.

Nagano, T., and Fraser, P. (2011). No-nonsense functions for long noncoding RNAs. Cell *145*, 178–181.

Nagle, A.M., Levine, K.M., Tasdemir, N., Scott, J.A., Burlbaugh, K., Kehm, J., Katz, T.A., Boone, D.N., Jacobsen, B.M., Atkinson, J.M., et al. (2018). Loss of E-cadherin enhances IGF1-IGF1R pathway activation and sensitizes breast cancers to anti-IGF1R/InsR inhibitors. Clin. Cancer Res. *24*, 5165–5177.

Nam, J.W., Choi, S.W., and You, B.H. (2016). Incredible RNA: Dual functions of coding and noncoding. Mol. Cells *39*, 367–374.

Nanduri, R., Furusawa, T., and Bustin, M. (2020). Biological Functions of HMGN Chromosomal Proteins. Int. J. Mol. Sci. *21*, 449.

Necsulea, A., and Kaessmann, H. (2014). Evolutionary dynamics of coding and non-coding transcriptomes. Nat. Rev. Genet. *15*, 734–748.

Niall, H.D., Hogan, M.L., Sauer, R., Rosenblum, I.Y., and Greenwood, F.C. (1971). Sequences of pituitary and placental lactogenic and growth hormones: evolution from a primordial peptide by gene reduplication. Proc. Natl. Acad. Sci. U. S. A. *68*, 866–870.

Nightingale, K.P., O'Neill, L.P., and Turner, B.M. (2006). Histone modifications: Signalling receptors and potential elements of a heritable epigenetic code. Curr. Opin. Genet. Dev. *16*, 125–136.

Nilsson, A., Carlsson, B., Isgaard, J., Isaksson, O.G.P., and Rymo, L. (1990). Regulation by GH of insulin-like growth factor-I mRNA expression in rat epiphyseal growth plate as studied with in-situ hybridization. J. Endocrinol. *125*, 67–74.

Nyaga, D.M., Vickers, M.H., Jefferies, C., Perry, J.K., and O'Sullivan, J.M. (2018). Type 1 Diabetes Mellitus-Associated Genetic Variants Contribute to Overlapping Immune Regulatory Networks. Front. Genet. *9*, 535.

Ohlsson, C., Mohan, S., Sjögren, K., Tivesten, Å., Isgaard, J., Isaksson, O., Jansson, J.O., and Svensson, J. (2009). The role of liver-derived insulin-like growth factor-I. Endocr. Rev. *30*, 494–535.

Ohlsson, R., Renkawitz, R., and Lobanenkov, V. (2001). CTCF is a uniquely versatile transcription regulator linked to epigenetics and disease. Trends Genet. *17*, 520–527.

Oki, S., Ohta, T., Shioi, G., Hatanaka, H., Ogasawara, O., Okuda, Y., Kawaji, H., Nakaki, R., Sese, J., and Meno, C. (2018). ChIP -Atlas: a data-mining suite powered by full integration of public ChIP-seq data. EMBO Rep. *19*, e46255.

Orth, M.F., Cazes, A., Butt, E., and Grunewald, T.G.P. (2015). An update on the LIM and SH3 domain protein 1 (LASP1): A versatile structural, signaling, and biomarker protein. Oncotarget *6*, 26–42.

Osmundsen, A.M., Keisler, J.L., Mark Taketo, M., and Davis, S.W. (2017). Canonical WNT signaling regulates the pituitary organizer and pituitary gland formation. Endocrinology *158*, 3339–3353.

Pachnis, V., Belayew, A., and Tilghman, S.M. (1984). Locus unlinked to α-fetoprotein under the control of the murine raf and Rif genes. Proc. Natl. Acad. Sci. U. S. A. *81*, 5523–5527.

Pandey, V., Perry, J.K., Mohankumar, K.M., Kong, X.J., Liu, S.M., Wu, Z.S., Mitchell, M.D., Zhu, T., and Lobie, P.E. (2008). Autocrine human growth hormone stimulates oncogenicity of endometrial carcinoma cells. Endocrinology *149*, 3909–3919.

Pandini, G., Frasca, F., Mineo, R., Sciacca, L., Vigneri, R., and Belfiore, A. (2002). Insulin/insulin-like growth factor I hybrid receptors have different biological characteristics depending on the insulin receptor isoform involved. J. Biol. Chem. *277*, 39684–39695.

Papadopoulos, N., Lennartsson, J., and Heldin, C.H. (2018). PDG FRβ translocates to the nucleus and regulates chromatin remodeling via TATA element-modifying factor 1. J. Cell Biol. *217*, 1701–1717.

Pekic, S., Stojanovic, M., and Popovic, V. (2017). Controversies in the risk of neoplasia in GH deficiency. Best Pract. Res. Clin. Endocrinol. Metab. *31*, 35–47.

Perrini, S., Laviola, L., Carreira, M.C., Cignarelli, A., Natalicchio, A., and Giorgino, F. (2010). The GH/IGF1 axis and signaling pathways in the muscle and bone: Mechanisms underlying age-related skeletal muscle wasting and osteoporosis. J. Endocrinol. *205*, 201–210.

Perry, J.K., Emerald, B.S., Mertani, H.C., and Lobie, P.E. (2006). The oncogenic potential of growth hormone. Growth Horm. IGF Res. *16*, 277–289.

Perry, J.K., Mohankumar, K.M., Emerald, B.S., Mertani, H.C., and Lobie, P.E. (2008). The contribution of growth hormone to mammary neoplasia. J. Mammary Gland Biol. Neoplasia *13*, 131–145.

Perry, J.K., Liu, D.-X., Wu, Z.-S., Zhu, T., and Lobie, P.E. (2013). Growth hormone and cancer. Curr. Opin. Endocrinol. Diabetes Obes. *20*, 307–313.

Perry, J.K., Wu, Z.S., Mertani, H.C., Zhu, T., and Lobie, P.E. (2017). Tumour-Derived Human Growth Hormone As a Therapeutic Target in Oncology. Trends Endocrinol. Metab. *28*, 587–596.

Piao, H., and Ma, L. (2012). Non-Coding RNAs as Regulators of Mammary Development and Breast Cancer. J. Mammary Gland Biol. Neoplasia *17*, 33–42.

Podlecki, D.A., Smith, R.M., Kao, M., Tsai, P., Huecksteadt, T., Brandenburg, D., Lasher, R.S., Jarett, L., and Olefskyd, J.M. (1987). Nuclear Translocation of the Insulin Receptor: A Possible Mediator Of Insulin's Long Term Effects. J. Biol. Chem. *262*, 3362–3368.

Politz, J.C.R., Scalzo, D., and Groudine, M. (2013). Something Silent This Way Forms: The Functional Organization of the Repressive Nuclear Compartment. Annu. Rev. Cell Dev. Biol. *29*, 241–270.

Pollak, M. (2012). The insulin and insulin-like growth factor receptor family in neoplasia: An update. Nat. Rev. Cancer *12*, 159–169.

Poreba, E., and Durzynska, J. (2020). Nuclear localization and actions of the insulin-like growth factor 1 (IGF-1) system components: Transcriptional regulation and DNA damage response. Mutat. Res. Mutat. Res. *784*, 108307.

Priyanka, Kotiya, D., Rana, M., Subbarao, N., Puri, N., and Tyagi, R.K. (2016). Transcription regulation of nuclear receptor PXR: Role of SUMO-1 modification and NDSM in receptor function. Mol. Cell. Endocrinol. *420*, 194–207.

Psyrri, A., Yu, Z., Weinberger, P.M., Sasaki, C., Haffty, B., Camp, R., Rimm, D., and Burtness, B.A. (2005). Quantitative determination of nuclear and cytoplasmic epidermal growth factor receptor expression in oropharyngeal squamous cell cancer by using automated quantitative analysis. Clin. Cancer Res. *11*, 5856–5862.

Qi, H.-L., Li, C.-S., Qian, C.-W., Xiao, Y.-S., Yuan, Y.-F., Liu, Q., and Liu, Z.-S. (2016). The long noncoding RNA, EGFR-AS1, a target of GHR, increases the expression of EGFR in

hepatocellular carcinoma. Tumor Biol. *37*, 1079–1089.

Qi, X., Zhang, D.H., Wu, N., Xiao, J.H., Wang, X., and Ma, W. (2015). ceRNA in cancer: Possible functions and clinical implications. J. Med. Genet. *52*, 710–718.

Quinn, J.J., and Chang, H.Y. (2016). Unique features of long non-coding RNA biogenesis and function. Nat. Rev. Genet. *17*, 47–62.

Rachagani, S., Kumar, S., and Batra, S.K. (2010). MicroRNA in pancreatic cancer: Pathological, diagnostic and therapeutic implications. Cancer Lett. *292*, 8–16.

Rakowicz-Szulczynska, E.M., Rodeck, U., Herlyn, M., and Koprowski, H. (1986). Chromatin binding of epidermal growth factor, nerve growth factor and platelet-derived growth factor in cells bearing the appropriate surface receptor. Proc. Natl. Acad. Sci. U. S. A. *83*, 3728–3732.

Rao, S.S.P., Huntley, M.H., Durand, N.C., Stamenova, E.K., Bochkov, I.D., Robinson, J.T., Sanborn, A.L., Machol, I., Omer, A.D., Lander, E.S., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. Cell *159*, 1665–1680.

Rao, Y.P., Buckley, D.J., and Buckley, A.R. (1995). The nuclear prolactin receptor: a 62-kDa chromatin-associated protein in rat Nb2 lymphoma cells. Arch. Biochem. Biophys. *322*, 506–515.

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., and Vilo, J. (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Res. *47*, W191–W198.

Rauhala, H.E., Jalava, S.E., Isotalo, J., Bracken, H., Lehmusvaara, S., Tammela, T.L.J., Oja, H., and Visakorpi, T. (2010). miR-193b is an epigenetically regulated putative tumor suppressor in prostate cancer. Int. J. Cancer *127*, 1363–1372.

Reddy, K.B. (2015). MicroRNA (miRNA) in cancer. Cancer Cell Int. *15*, 38.

Reilly, S.K., and Noonan, J.P. (2016). Evolution of Gene Regulation in Humans. Annu. Rev. Genomics Hum. Genet. *17*, 45–67.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. *43*, e47–e47.

Roelfsema, V., and Clark, R.G. (2001). The growth hormone and insulin-like growth factor

axis: Its manipulation for the benefit of growth disorders in renal failure. J. Am. Soc. Nephrol. *12*, 1297–1306.

Le Roith, D., Bondy, C., Yakar, S., Liu, J.L., and Butler, A. (2001). The somatomedin hypothesis: 2001. Endocr. Rev. *22*, 53–74.

Rother, K.I., and Accili, D. (2000). Role of insulin receptors and IGF receptors in growth and development. Pediatr. Nephrol. *14*, 558–561.

Rotwein, P., and Chia, D.J. (2010). Gene regulation by growth hormone. Pediatr. Nephrol. *25*, 651–658.

RW Carthew, E.S. (2009). Origins and mechanisms of miRNAs and siRNAs. Cell *136*, 642–655.

Sabin, L.R., Delás, M.J., and Hannon, G.J. (2013). Dogma Derailed: The Many Influences of RNA on the Genome. Mol. Cell *49*, 783–794.

Sachse, M., van Kerkhof, P., Strous, G.J., and Klumperman, J. (2001). The ubiquitin-dependent endocytosis motif is required for efficient incorporation of growth hormone receptor in clathrin-coated pits, but not clathrin-coated lattices. J. Cell Sci. *114*, 3943–3952.

Salmena, L., Poliseno, L., Tay, Y., Kats, L., and Pandolfi, P.P. (2011). A ceRNA hypothesis: The rosetta stone of a hidden RNA language? Cell *146*, 353–358.

Samani, A.A., Yakar, S., LeRoith, D., and Brodt, P. (2007). The role of the IGF system in cancer growth and metastasis: Overview and recent insights. Endocr. Rev. *28*, 20–47.

Sanborn, A.L., Rao, S.S.P., Huang, S.C., Durand, N.C., Huntley, M.H., Jewett, A.I., Bochkov, I.D., Chinnappan, D., Cutkosky, A., Li, J., et al. (2015). Chromatin extrusion explains key features of loop and domain formation in wild-type and engineered genomes. Proc. Natl. Acad. Sci. U. S. A. *112*, E6456–E6465.

Sanyal, A., Lajoie, B.R., Jain, G., and Dekker, J. (2012). The long-range interaction landscape of gene promoters. Nature *489*, 109–113.

Sarfstein, R., Pasmanik-Chor, M., Yeheskel, A., Edry, L., Shomron, N., Warman, N., Wertheimer, E., Maor, S., Shochat, L., and Werner, H. (2012). Insulin-like Growth Factor-I Receptor (IGF-IR) translocates to nucleus and autoregulates IGF-IR gene expression in breast cancer cells. J. Biol. Chem. *287*, 2766–2776.

Savinko, T., Guenther, C., Uotila, L.M., Llort Asens, M., Yao, S., Tojkander, S., and Fagerholm, S.C. (2018). Filamin A Is Required for Optimal T Cell Integrin-Mediated Force Transmission, Flow Adhesion, and T Cell Trafficking. J. Immunol. *200*, 3109–3116.

Saxton, R.A., and Sabatini, D.M. (2017). mTOR Signaling in Growth, Metabolism, and Disease. Cell *168*, 960–976.

Schierding, W., and O'Sullivan, J.M. (2015). Connecting SNPs in Diabetes: A Spatial Analysis of Meta-GWAS Loci. Front. Endocrinol. (Lausanne). *6*, 102.

Schierding, W., Cutfield, W.S., and O'Sullivan, J.M. (2014). The missing story behind Genome Wide Association Studies: Single nucleotide polymorphisms in gene deserts have a story to tell. Front. Genet. *5*.

Schierding, W., Antony, J., Cutfield, W.S., Horsfield, J.A., and O'Sullivan, J.M. (2016). Intergenic GWAS SNPs are key components of the spatial and regulatory network for human growth. Hum. Mol. Genet. *25*, 3372–3382.

Schierding, W., Farrow, S., Fadason, T., Graham, O.E.E., Pitcher, T.L., Qubisi, S., Davidson, A.J., Perry, J.K., Anderson, T.J., Kennedy, M.A., et al. (2020). Common Variants Coregulate Expression of <scp> *GBA* </scp> and Modifier Genes to Delay Parkinson's Disease Onset. Mov. Disord. mds.28144.

Schrader, A.M.R., Jansen, P.M., Willemze, R., Vermeer, M.H., Cleton-Jansen, A.M., Somers, S.F., Veelken, H., Van Eijk, R., Kraan, W., Kersten, M.J., et al. (2018). High prevalence of MYD88 and CD79B mutations in intravascular large B-cell lymphoma. Blood *131*, 2086–2089.

Sehat, B., Tofigh, A., Lin, Y., Trocme, E., Liljedahl, U., Lagergren, J., and Larsson, O. (2010). SUMOylation Mediates the Nuclear Translocation and Signaling of the IGF-1 Receptor. Sci. Signal. *3*, ra10-ra10.

Selbach, M., Schwanhäusser, B., Thierfelder, N., Fang, Z., Khanin, R., and Rajewsky, N. (2008). Widespread changes in protein synthesis induced by microRNAs. Nature *455*, 58–63.

Seymour, S.L., and Hunter, C. (2015). ProteinPilot TM Software Overview High Quality , In-Depth Protein Identification and Protein Expression Analysis.

Shabalina, S.A., and Spiridonov, N.A. (2004). The mammalian transcriptome and the function of non-coding DNA sequences. Genome Biol. *5*, 105.

Shafiei, F., Rahnama, F., Pawella, L., Mitchell, M.D., Gluckman, P.D., and Lobie, P.E. (2008). DNMT3A and DNMT3B mediate autocrine hGH repression of plakoglobin gene transcription and consequent phenotypic conversion of mammary carcinoma cells. Oncogene *27*, 2602–2612.

Shah, Chaumet, Royle, and Bard (2019). The NAE Pathway: Autobahn to the Nucleus for Cell Surface Receptors. Cells *8*, 915.

Shay-Salit, A., Shushy, M., Wolfovitz, E., Yahav, H., Breviario, F., Dejana, E., and Resnick, N. (2002). VEGF receptor 2 and the adherens junction as a mechanical transducer in vascular endothelial cells. Proc. Natl. Acad. Sci. U. S. A. *99*, 9462–9467.

Shibata, H., Kanadome, T., Sugiura, H., Yokoyama, T., Yamamuro, M., Moss, S.E., and Maki, M. (2015). A new role for annexin A11 in the early secretory pathway via stabilizing Sec31A protein at the endoplasmic reticulum exit sites (ERES). J. Biol. Chem. *290*, 4981–4993.

Shrestha, S., Hsu, S.-D., Huang, W.-Y., Huang, H.-Y., Chen, W., Weng, S.-L., and Huang, H.-D. (2014). A systematic review of microRNA expression profiling studies in human gastric cancer. Cancer Med. *3*, 878–888.

Simpson, A., Petnga, W., Macaulay, V.M., Weyer-Czernilofsky, U., and Bogenrieder, T. (2017). Insulin-Like Growth Factor (IGF) Pathway Targeting in Cancer: Role of the IGF Axis and Opportunities for Future Combination Studies. Target. Oncol. *12*, 571–597.

Smemo, S., Tena, J.J., Kim, K.H., Gamazon, E.R., Sakabe, N.J., Gómez-Marín, C., Aneas, I., Credidio, F.L., Sobreira, D.R., Wasserman, N.F., et al. (2014). Obesity-associated variants within FTO form long-range functional connections with IRX3. Nature *507*, 371–375.

Smet-Nocca, C., Wieruszeski, J.-M., Léger, H., Eilebrecht, S., and Benecke, A. (2011). SUMO-1 regulates the conformational dynamics of thymine-DNA Glycosylase regulatory domain and competes with its DNA binding activity. BMC Biochem. *12*, 4.

Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N., and De Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the β-globin locus. Genes Dev. *20*, 2349–2354.

St.Laurent, G., Wahlestedt, C., and Kapranov, P. (2015). The Landscape of long noncoding RNA classification. Trends Genet. *31*, 239–251.

Stadhouders, R., Vidal, E., Serra, F., Di Stefano, B., Le Dily, F., Quilez, J., Gomez, A., Collombet, S., Berenguer, C., Cuartero, Y., et al. (2018). Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. Nat. Genet. *50*, 238–249.

van Steensel, B., and Belmont, A.S. (2017). Lamina-Associated Domains: Links with Chromosome Architecture, Heterochromatin, and Gene Repression. Cell *169*, 780–791.

Steuerman, R., Shevah, O., and Laron, Z. (2011). Congenital IGF1 deficiency tends to confer protection against post-natal development of malignancies. Eur. J. Endocrinol. *164*, 485–489.

Storey, J.D., Madeoy, J., Strout, J.L., Wurfel, M., Ronald, J., and Akey, J.M. (2007). Gene-expression variation within and among human populations. Am. J. Hum. Genet. *80*, 502–509.

Stranger, B.E., Montgomery, S.B., Dimas, A.S., Parts, L., Stegle, O., Ingle, C.E., Sekowska, M., Smith, G.D., Evans, D., Gutierrez-Arcelus, M., et al. (2012). Patterns of Cis regulatory variation in diverse human populations. PLoS Genet. *8*, e1002639.

Strous, G.J., and Van Kerkhof, P. (2002). The ubiquitin-proteasome pathway and the regulation of growth hormone receptor availability. Mol. Cell. Endocrinol. *197*, 143–151.

Su, Y., Liebhaber, S.A., and Cooke, N.E. (1997). The human growth hormone locus control region supports pituitary- and placental-specific patterns of gene expression in transgenic mice. FASEB J. *11*, 7902–7909.

Sui, Y., Yang, Z., Xiong, S., Zhang, L., Blanchard, K.L., Peiper, S.C., Dynan, W.S., Tuan, D., and Ko, L. (2007). Gene amplification and associated loss of 5′ regulatory sequences of CoAA in human cancers. Oncogene *26*, 822–835.

Symmons, O., Uslu, V.V., Tsujimura, T., Ruf, S., Nassari, S., Schwarzer, W., Ettwiller, L., and Spitz, F. (2014). Functional and topological characteristics of mammalian regulatory domains. Genome Res. *24*, 390–400.

Szcześniak, M.W., and Makałowska, I. (2016). lncRNA-RNA Interactions across the Human Transcriptome. PLoS One *11*, e0150353.

Szklarczyk, D., Gable, A.L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N.T., Morris, J.H., Jensen, L.J., et al. (2018). STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. *47*, 607–613.

Taciak, B., Pruszynska, I., Kiraga, L., Bialasek, M., and Krol, M. (2018). Wnt signaling pathway in development and cancer. J. Physiol. Pharmacol. *69*, 185–196.

T Conrad, A.M.M.G.U.O. (2014). Microprocessor activity controls differential miRNA biogenesis In Vivo. Cell Rep. *9*, 542–554.

Tak, Y.G., and Farnham, P.J. (2015). Making sense of GWAS: using epigenomics and genome engineering to understand the functional relevance of SNPs in non-coding regions of the human genome. Epigenetics Chromatin *8*, 57.

Tan, C.L., Plotkin, J.L., Venø, M.T., Von Schimmelmann, M., Feinberg, P., Mann, S., Handler, A., Kjems, J., Surmeier, D.J., O'Carroll, D., et al. (2013). MicroRNA-128 governs neuronal excitability and motor behavior in mice. Science (80-.). *342*, 1254–1258.

Tang, Z., Luo, O.J., Li, X., Zheng, M., Zhu, J.J., Szalaj, P., Trzaskoma, P., Magalska, A., Wlodarczyk, J., Ruszczycki, B., et al. (2015). CTCF-Mediated Human 3D Genome Architecture Reveals Chromatin Topology for Transcription. Cell *163*, 1611–1627.

Tarnawski, A.S., Ahluwalia, A., Gergely, H.M., and Jones, M.K. (2015). 800 Expression and Co-Localization of IGF-1, Its Receptor and Survivin in Esophageal Progenitor Cells: Implications for Esophageal Mucosal Renewal, Protection and Healing. Gastroenterology *148*, S-157.

Tay, Y., Rinn, J., and Pandolfi, P.P. (2014). The multilayered complexity of ceRNA crosstalk and competition. Nature *505*, 344–352.

Thankamony, G.A., Dunger, D.B., and Acerini, C.L. (2009). Pegvisomant: Current and potential novel therapeutic applications. Expert Opin. Biol. Ther. *9*, 1553–1563.

Therizols, P., Illingworth, R.S., Courilleau, C., Boyle, S., Wood, A.J., and Bickmore, W.A. (2014). Chromatin decondensation is sufficient to alter nuclear organization in embryonic stem cells. Science (80-.). *346*, 1238–1242.

Thi, N., Phuong, T., Kim, S.K., Im, J.H., Yang, J.W., Choi, M.C., Lim, S.C., Lee, K.Y., Kim, Y.-M., Yoon, J.H., et al. (2015). Induction of methionine adenosyltransferase 2A in tamoxifen-resistant breast cancer cells.

Thomson, D.W., and Dinger, M.E. (2016). Endogenous microRNA sponges: Evidence and controversy. Nat. Rev. Genet. *17*, 272–283.

Thurman, R.E., Rynes, E., Humbert, R., Vierstra, J., Maurano, M.T., Haugen, E., Sheffield, N.C., Stergachis, A.B., Wang, H., Vernot, B., et al. (2012). The accessible chromatin landscape of the human genome. Nature *489*, 75–82.

Du Toit, A. (2015). Endocytosis: A new gateway into cells. Nat. Rev. Mol. Cell Biol. *16*, 68.

Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and De Laat, W. (2002). Looping and interaction between hypersensitive sites in the active β-globin locus. Mol. Cell *10*, 1453–1465.

Traynor, A.M., Weigel, T.L., Oettel, K.R., Yang, D.T., Zhang, C., Kim, K.M., Salgia, R., Iida, M., Brand, T.M., Hoang, T., et al. (2013). Nuclear EGFR protein expression predicts poor survival in early stage non-small cell lung cancer. Lung Cancer *81*, 138–141.

Treiber, T., Treiber, N., and Meister, G. (2019). Regulation of microRNA biogenesis and its crosstalk with other cellular pathways. Nat. Rev. Mol. Cell Biol. *20*, 5–20.

Trott, J.F., Schennink, A., Petrie, W.K., Manjarin, R., VanKlompenberg, M.K., and Hovey, R.C. (2012). Prolactin: The multifaceted potentiator of mammary growth and function1,2. J. Anim. Sci. *90*, 1674–1686.

Tsai, M.C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J.K., Lan, F., Shi, Y., Segal, E., and Chang, H.Y. (2010). Long noncoding RNA as modular scaffold of histone modification complexes. Science (80-.). *329*, 689–693.

Tsai, Y.C., Leichner, G.S., Pearce, M.M.P., Wilson, G.L., Wojcikiewicz, R.J.H., Roitelman, J., and Weissman, A.M. (2012). Differential regulation of HMG-CoA reductase and Insig-1 by enzymes of the ubiquitin-proteasome system. Mol. Biol. Cell *23*, 4484–4494.

Tsai, Y.C., Cooke, N.E., and Liebhaber, S.A. (2016). Long-range looping of a locus control region drives tissue-specific chromatin packing within a multigene cluster. Nucleic Acids Res. *44*, 4651–4664.

Uhlmann, S., Mannsperger, H., Zhang, J.D., Horvat, E., Schmidt, C., Küblbeck, M., Henjes, F., Ward, A., Tschulena, U., Zweig, K., et al. (2012). Global microRNA level regulation of EGFR-driven cell-cycle protein network in breast cancer. Mol. Syst. Biol. *8*, 570.

Ulianov, S. V., Khrameeva, E.E., Gavrilov, A.A., Flyamer, I.M., Kos, P., Mikhaleva, E.A., Penin, A.A., Logacheva, M.D., Imakaev, M. V., Chertovich, A., et al. (2016). Active chromatin and transcription play a key role in chromosome partitioning into topologically

associating domains. Genome Res. *26*, 70–84.

Ulitsky, I., and Bartel, D.P. (2013). XLincRNAs: Genomics, evolution, and mechanisms. Cell *154*, 26.

Vaquerizas, J.M., Kummerfeld, S.K., Teichmann, S.A., and Luscombe, N.M. (2009). A census of human transcription factors: Function, expression and evolution. Nat. Rev. Genet. *10*, 252–263.

Vesterlund, M., Zadjali, F., Persson, T., Nielsen, M.L., Kessler, B.M., Norstedt, G., and Flores-Morales, A. (2011). The SOCS2 Ubiquitin Ligase Complex Regulates Growth Hormone Receptor Levels. PLoS One *6*, e25358.

Vouyovitch, C.M., Perry, J.K., Liu, D.X., Bezin, L., Vilain, E., Diaz, J.-J., Lobie, P.E., and Mertani, H.C. (2016). WNT4 mediates the autocrine effects of growth hormone in mammary carcinoma cells. Endocr. Relat. Cancer *23*, 571–585.

Wahdan-Alaswad, R.S., Cochrane, D.R., Spoelstra, N.S., Howe, E.N., Edgerton, S.M., Anderson, S.M., Thor, A.D., and Richer, J.K. (2014). Metformin-Induced Killing of Triple-Negative Breast Cancer Cells Is Mediated by Reduction in Fatty Acid Synthase via miRNA-193b. Horm. Cancer *5*, 374–389.

Wallis, M. (2014). Molecular evolution of growth hormone. Biochem. (Lond). *36*, 4–8.

Wallis, O.C., and Wallis, M. (2006). Evolution of Growth Hormone in Primates: The GH Gene Clusters of the New World Monkeys Marmoset (Callithrix jacchus) and White-Fronted Capuchin (Cebus albifrons). J. Mol. Evol. *63*, 591–601.

Wang, S.C., and Hung, M.C. (2009). Nuclear translocation of the epidermal growth factor receptor family membrane tyrosine kinase receptors. Clin. Cancer Res. *15*, 6484–6489.

Wang, J.-J., Chong, Q.-Y., Sun, X.-B., You, M.-L., Pandey, V., Chen, Y.-J., Zhuang, Q.-S., Liu, D.-X., Ma, L., Wu, Z.-S., et al. (2017). Autocrine hGH stimulates oncogenicity, epithelial-mesenchymal transition and cancer stem cell-like behavior in human colorectal carcinoma. Oncotarget *8*, 103900–103918.

Wang, S.C., Nakajima, Y., Yu, Y.L., Xia, W., Chen, C. Te, Yang, C.C., McIntush, E.W., Li, L.Y., Hawke, D.H., Kobayashi, R., et al. (2006). Tyrosine phosphorylation controls PCNA function through protein stability. Nat. Cell Biol. *8*, 1359–1368.

Warnes, G., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T., Maechler, M., Magnusson, A., Moeller, S., et al. (2020). Gplots: Various R Programming Tools for Plotting Data, 3.0.4.

Warsito, D., Sjöström, S., Andersson, S., Larsson, O., and Sehat, B. (2012). Nuclear IGF1R is a transcriptional co-activator of LEF1/TCF. EMBO Rep. *13*, 244–250.

Waters, M.J. (2016). The growth hormone receptor. Growth Horm. IGF Res. *28*, 6–10.

Van De Werken, H.J.G., Landan, G., Holwerda, S.J.B., Hoichman, M., Klous, P., Chachik, R., Splinter, E., Valdes-Quezada, C., Öz, Y., Bouwman, B.A.M., et al. (2012). Robust 4C-seq data analysis to screen for regulatory DNA interactions. Nat. Methods *9*, 969–972.

Weroha, S.J., and Haluska, P. (2012). The Insulin-Like Growth Factor System in Cancer. Endocrinol. Metab. Clin. North Am. *41*, 335–350.

Wijchers, P.J., Krijger, P.H.L., Geeven, G., Zhu, Y., Denker, A., Verstegen, M.J.A.M., Valdes-Quezada, C., Vermeulen, C., Janssen, M., Teunissen, H., et al. (2016). Cause and Consequence of Tethering a SubTAD to Different Nuclear Compartments. Mol. Cell *61*, 461–473.

Willier, S., Butt, E., Richter, G.H.S., Burdach, S., and Grunewald, T.G.P. (2011). Defining the role of TRIP6 in cell physiology and cancer. Biol. Cell *103*, 573–591.

Xia, W., Wei, Y., Du, Y., Liu, J., Chang, B., Yu, N.L., Huo, L.F., Miller, S., and Hung, M.C. (2009). Nuclear expression of epidermal growth factor receptor is a novel prognostic value in patients with ovarian cancer. Mol. Carcinog. *48*, 610–617.

Xu, W., Gorman, P.A., Rider, S.H., Hedge, P.J., Moore, G., Prichard, C., Sheer, D., and Solomon, E. (1988). Construction of a genetic map of human chromosome 17 by use of chromosome-mediated gene transfer. Proc. Natl. Acad. Sci. U. S. A. *85*, 8563–8567.

Xu, X.Q., Emerald, B.S., Goh, E.L.K., Kannan, N., Miller, L.D., Gluckman, P.D., Liut, E.T., and Lobie, P.E. (2005). Gene expression profiling to identify oncogenic determinants of autocrine human growth hormone in human mammary carcinoma. J. Biol. Chem. *280*, 23987–24003.

Yang, X., and Friedl, A. (2015). A positive feedback loop between prolactin and stat5 promotes angiogenesis. Adv. Exp. Med. Biol. *846*, 265–280.

Yang, N., Huang, Y., Jiang, J., and Frank, S.J. (2004). Caveolar and lipid raft localization of the growth hormone receptor and its signaling elements: Impact on growth hormone signaling. J. Biol. Chem. *279*, 20898–20905.

Yoo, K.H., Kang, K., Feuermann, Y., Jang, S.J., Robinson, G.W., and Hennighausen, L. (2014). The STAT5-regulated miR-193b locus restrains mammary stem and progenitor cell activity and alveolar differentiation. Dev. Biol. *395*, 245–254.

Yu, M., and Ren, B. (2017). The Three-Dimensional Organization of Mammalian Genomes. Annu. Rev. Cell Dev. Biol. *33*, 265–289.

Zerbino, D.R., Achuthan, P., Akanni, W., Amode, M.R., Barrell, D., Bhai, J., Billis, K., Cummins, C., Gall, A., Girón, C.G., et al. (2018). Ensembl 2018. Nucleic Acids Res. *46*, D754–D761.

Zhang, Y., and Reinberg, D. (2001). Transcription regulation by histone methylation: Interplay between different covalent modifications of the core histone tails. Genes Dev. *15*, 2343–2360.

Zhang, J., Huang, F.F., Wu, D.S., Li, W.J., Zhan, H.E., Peng, M.Y., Fang, P., Cao, P.F., Zhang, M.M., Zeng, H., et al. (2015a). SUMOylation of insulin-like growth factor 1 receptor, promotes proliferation in acute myeloid leukemia. Cancer Lett. *357*, 297–306.

Zhang, W., Qian, P., Zhang, X., Zhang, M., Wang, H., Wu, M., Kong, X., Tan, S., Ding, K., Perry, J.K., et al. (2015b). Autocrine/Paracrine Human Growth Hormone-stimulated MicroRNA 96-182-183 Cluster Promotes Epithelial-Mesenchymal Transition and Invasion in Breast Cancer. J. Biol. Chem. *290*, 13812–13829.

Zheng, Y., Lu, S., Xu, Y., and Zheng, J. (2019). Long non-coding RNA AGAP2-AS1 promotes the proliferation of glioma cells by sponging miR-15a/b-5p to upregulate the expression of HDGF and activating Wnt/β-catenin signaling pathway. Int. J. Biol. Macromol. *128*, 521–530.

Zhou, Y., Gerrard, D.L., Wang, J., Li, T., Yang, Y., Fritz, A.J., Rajendran, M., Fu, X., Schiff, R., Lin, S., et al. (2019). Temporal dynamic reorganization of 3D chromatin architecture in hormone-induced breast cancer and endocrine resistance. Nat. Commun. *10*, 1–14.

Zhu, T., Emerald, B.S., Zhang, X., Lee, K.-O.O., Gluckman, P.D., Mertani, H.C., Lobie, P.E., Starling-Emerald, B., Zhang, X., Lee, K.-O.O., et al. (2005). Oncogenic transformation of

human mammary epithelial cells by autocrine human growth hormone. Cancer Res. *65*, 317–324.