Genetic Variations in the Human *Growth Hormone Receptor* Gene: functional relevance and associations with short stature and obesity.

by

Christel Dias

Division of Experimental Medicine

Department of Medicine

Faculty of Medicine

McGill University

Montreal, Quebec, Canada

November 2018

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

aa	amino acid
ACAN	Aggrecan gene
ACTH	Adrenocorticotrophic Hormone
AI	Allelic Imbalance
Akt	serine/threonine kinase (also known as PKB)
ALS	Acid Labile Subunit
ANOVA	Analysis of Variance
B2M	Beta 2 Microglobulin
BAD	BCL-2 Associated Death Promoter
BAI	Body Adiposity Index
BAX	BCL2 Associated X protein
BCL2	B Cell Lymphoma 2
BF%	Body Fat %
BMD	Bone Mineral Density
BMI	Body Mass Index
BRET	Bioluminescence resonance energy transfer
BRG1	Brahma-related gene 1
cAMP/PKA	cyclic AMP/protein kinase A
C/EBP	CCAAT enhancer binding protein
CEPH	Centre d'Etude du Polymorphisme Humain

Cfp1	CXXC finger protein 1
CGHD	Compound GH Deficiency
ChIP	Chromatin Immunoprecipitation
CHOP	C/EBP Homologous Protein
CI	Confidence Interval
CIS	Cytokine-inducible SH2-domain protein
CS	Chorionic Somatomammotropin
CTCF	CCCTC-binding Factor
C-terminus	carboxyl-terminus
DBP	D-binding Protein
ddPCR	drop digital Polymerase Chain Reaction
DNA	Deoxyribonucleic acid
DNAse	Deoxyribonuclease
DNMT	DNA methyltransferase
DTT	Dithiotreitol
EBV	Epstein-Barr Virus
ECD	Extracellular domain
EMS(S)A	Electrophoretic Mobility Shift (Super)Assay
EPO	Erythropoietin
eQTL	Expression Quantitative Trait Locus
ER	Estrogen receptor

ERE	Estrogen response element
ERK	Extracellular signal-regulated kinase
eSTR	Expression Short Tandem Repeats
FERM	N-terminal 4.1, Ezrin, Radixin, Moesin
FFAs	Free fatty acids
FFM	Fat Free Mass
FFMI	Fat Free Mass Index
FIPA	Familial Isolated Pituitary Adenoma
FM	Fat Mass
FMI	Fat Mass Index
FRET	Fluorescence resonance energy transfer
FSH	Follicle Stimulating Hormone
FTO	Fat Mass and Obesity-Associated
Gα	G protein alpha subunit
GDF	Growth Differentiation Factor
Gfi-1/1b	Growth Factor Independence-1/1b
GH	Growth Hormone
GHBP	Growth Hormone Binding Protein
GHD	Growth Hormone Deficient
GHI	Growth Hormone Insensitive
GHR	Growth Hormone Receptor

GHRH	Growth Hormone Releasing Hormone
GHRHR	Growth Hormone Releasing Hormone Receptor
GHRfl	Full-length Growth Hormone Receptor
GHRtr	Truncated Growth hormone Receptor
GHS	GH secretagogue
GHSR	GH secretagogue receptor
GIANT	Genetic Investigation of Anthropometric Traits
GM-CSF	Granulocyte-Macrophage Colony-Stimulation Factor
Grb2	Growth factor receptor-bound protein 2
GSK-3	Glycogen synthase kinase 3
GWAS	Genome Wide Association Study
HC	Hip Circumference
HDAC	Histone Deacetylase
HIF	Hypoxia Inducible Factor
HMGA2	High Mobility Group AT-Hook 2
HNF	Hepatic Nuclear Factor
HPRT	Hypoxanthine-guanine phosphoribosyltransferase
HS	Hypersensitive sites
HSL	Hormone Sensitive Lipase
ICD	Intracellular domain
IFN	Interferon

IGF-1	Insulin-like Growth Factor 1
IGF-1R	Insulin-like Growth Factor 1 Receptor
IGF-2	Insulin-like Growth Factor 2
IGF-2R	Insulin Growth Factor 2 Receptor
IGFBP	Insulin-like Growth Factor Binding Protein
IGHD	Isolated GH Deficiency
IL	Interleukin
IRS	Insulin Receptor Substrate
ISS	Idiopathic Short Stature
IUGR	Intrauterine Growth Retardation
JAK	Janus Kinase
kDa	kiloDalton
KIR	Kinase Inhibitory Region
L	Long
LCL	Lymphoblastoid Cell Lines
LD	Linkage Disequilibrium
LINES	Long Interspersed Elements
LPL	Lipoprotein Lipase
LTR	Long Terminal Repeat
M	Medium
MAF	Minor Allele Frequency

MAPK	Mitogen activated protein kinase
MBP	Methyl Binding Protein
MEK	Mitogen-activated protein kinase kinase
MEN	Multiple Endocrine Neoplasia
meQTL	Methylated Expression Quantitative Trait Loci
miRISCs	miRNA-induced silencing complex
miRNA	Micro RNAs
mRNA	messenger ribonucleic acid
mTOR	mammalian Target of Rapamycin
NESTEGG	Network of European studies of genes in growth
NFκB	Nuclear Factor kappa B
NRP2	Natriuretic peptide receptor 2 gene
N-terminal	amino-terminal
PIAS	Protein Inhibitor of Activated STATS
PI3K	Phosphoinositol 3' kinase
PK	Protein kinase
POU1F1/Pit-1	Pou domain, class1, transcription factor 1
PPAR	Peroxisome proliferator activated receptor
PRL	Prolactin
PROP-1	Prophet of Pit-1
PTP	Protein Tyrosine Phosphatases

qPCR	Quantitative PCR or quantitative reverse transcription PCR (qRT-PCR)
RACE	Rapid Amplification of cDNA Ends
RNApolII	RNA Polymerase II
RT-PCR	Reverse Transcription – Polymerase Chain Reaction
S	Short
SDS	Standard Deviation Score
SFK	Src family kinases
SGA	Small for Gestational Age
SH2	Src homology 2 domains
SHOX	Short stature homeobox gene
SINES	Short Interspersed Elements
SNP	Single nucleotide polymorphism
SOCS	Suppressor of cytokine signaling
SOS	Son of Sevenless
SS	Short stature
SST	Somatostatin
SSTR	Somatostatin receptor
STAT	Signal Transducers and Activators of Transcription
STR	Short Tandem Repeat
SWI/SNF	SWItch/Sucrose Non-Fermentable
TBP	TATA Binding Proteins

TMD	Transmembrane domain
TNFα	Tumor Necrosis Factor alpha
TPO	Thrombopoietin
TSH	Thyroid Stimulating Hormone
TYK2	Tyrosine kinase 2
UbE	Ubiquitin motif
UTR	Untranslated region
VLDL	Very low density lipoproteins
WAT	White Adipose Tissue
WC	Waist Circumference
WHO	World Health Organization
WHR	Waist/Hip Ratio
ZBP	Z-DNA binding protein

ABSTRACT

The GH/IGF-1 axis plays an important role in the acquisition of longitudinal height but also in regulating fat metabolism. GH exerts its effects through specific binding to its receptor, the GH receptor (GHR), on target cells and activating multiple signaling pathways, leading to changes in cell gene expression and function. Any dysregulation in this axis can lead to pathophysiological consequences. For example, it is well known that a dysfunctional GHR protein leads to the Laron syndrome, characterized by proportional short stature and increased abdominal fat.

Idiopathic short stature (ISS), a less drastic phenotype, occurs in 1-2% of the human population and is characterized by an average height >2 SD below the mean but a normal endocrine profile. The etiology of this phenotype is still unknown; however, the fact that a high percentage of these individuals have slightly elevated levels of serum GH and lower IGF-1 suggests decreased GH sensitivity. Recent GWAS studies have shown an association of the *GHR* gene as well as other members of the GH/IGF-1 axis with variation in height in the Caucasian population but these studies have rarely included the extreme tails of height distribution observed in ISS.

The *GHR* gene structure is characterized by a unique 5' promoter organization. It comprises thirteen different first exons that produce transcripts that are either ubiquitously expressed or with a tissue- and developmental-specific expression pattern; however, all of the mRNAs code for the same protein. Previous studies in our lab found transcription factor-related functional mechanisms that regulate the expression of several of these transcripts. In addition, a GT microsatellite, located in the promoter of one major *GHR* transcript was shown to be polymorphic in the human population but its functional significance had not yet been characterized.

For my PhD project, I hypothesized that different genetic variants within the *GHR* gene - a set of SNPs and the GT microsatellite – would be associated with ISS and SS phenotypes and that these associations would be influenced by adiposity indices, such as body mass index (BMI). I also hypothesized that the variants would have a regulatory effect on *GHR* transcription.

Through a series of logistic regression analyses, I found sex-related associations of specific GT repeat genotypes and one SNP in intron 2 with Caucasian ISS pediatric cohorts (Chapter II). Moreover, by including BMI and other adiposity indices as covariates in my analysis of an adult control and SS cohort, I showed that they had confounding effects on the associations with short stature; a set of six SNPs exclusively located in the promoter region of the *GHR* gene and a specific GT genotype were strongly associated with female SS (Chapter III). Using three different experimental approaches (luciferase reporter constructs carrying different lengths of the GT repeat, allele-specific expression assays of *GHR* and quantitative PCR of GH-IGF1 axis genes), I was able to show that the GT microsatellite acts, in a context- and sex-specific manner, as a *cis* regulator of *GHR* expression and potentially exerts *trans* effects on *IGF1* and *BCL2* (Chapter IV).

These studies have shown the relevance of examining different types of genetic variants in the *GHR* gene for sex-specific associations with severe short stature and obesity and shed new light on the mechanisms regulating *GHR* expression.

RÉSUMÉ

L'axe GH/IGF-1 joue un rôle important dans l'acquisition de la taille longitudinale mais aussi dans la régulation du métabolisme lipidique. La GH exerce ses effets par la liaison spécifique à son récepteur, le récepteur a l'hormone de croissance (GHR), sur les cellules cibles et par l'activation de multiples voies de signalisation, aboutissant à des changements au niveau de l'expression et de la fonction des gènes de la cellule. Par exemple, il est bien connu que la protéine GHR dysfonctionnelle entraîne le syndrome de Laron, caractérisé par une petite taille proportionnelle ainsi qu'une augmentation de l'adiposité abdominale.

La petite taille idiopathique (ISS), un phénotype moins extrême, se retrouve dans 1-2% de la population humaine et est caractérisée par une taille moyenne >2DS au-dessous de la moyenne présentant néanmoins un profil endocrinien normal. L'étiologie de ce phénotype est encore inconnue; cependant, la présence d'un fort pourcentage d'individus présentant des taux sériques de GH légèrement élevés et d'IGF-1 légèrement bas suggèrent une diminution de sensibilité à l'hormone de croissance. De récentes études GWAS ont montré une association du gène *GHR* ainsi que d'autres membres de l'axe GH/IGF-1 avec la variation de taille dans la population caucasienne mais ces études n'incluent que rarement les extrêmes de la distribution staturale tel qu'observé dans ISS.

La structure du gène *GHR* est caractérisée par une organisation unique du promoteur en 5'. Il comprend 13 différents premiers exons qui produisent des transcrits ubiquitaires ou présentant un profil d'expression spécifique au tissu ou encore au stade de développement; cependant, tous les ARNm codent pour la même protéine. De précédentes études dans notre laboratoire ont montré que des mécanismes fonctionnels mettant en jeu des facteurs de transcription régulent l'expression de plusieurs de ces transcrits. De plus, un microsatellite GT, situé dans le promoteur d'un transcrit majeur de *GHR*, a été montré comme étant polymorphique dans la population humaine mais sa fonction n'a pas été encore élucidée.

Pour mon projet doctoral, j'ai émis l'hypothèse que différentes variations génétiques au sein du gène *GHR*- comprenant un ensemble de SNPs et un microsatellite GT- pourraient être associées aux phénotypes ISS et SS et que ces associations pourraient être influencées par des

indices d'adiposité comme l'indice de masse corporelle (IMC). En utilisant des modèles de régression logistiques, j'ai trouvé des associations spécifiques et liées au sexe, de certains génotypes de la répétition GT ainsi que d'un SNP situé dans l'intron 2 avec les cohortes pédiatriques caucasiennes ISS (Chapitre II). De plus, en incluant l'IMC et d'autres indices d'adiposité comme covariables dans mon analyse d'une cohorte d'adultes contrôle et de petite taille, j'ai démontré que ces indices avaient des effets confondants sur les associations avec la petite taille; un ensemble de 6 SNPs situé exclusivement dans le promoteur du gène *GHR* ainsi qu'un génotype spécifique de la répétition GT ont montré une forte association chez la femme de petite taille (Chapitre III). En utilisant trois approches expérimentales différentes (vecteurs comprenant le gène rapporteur Luc ainsi que des longueurs variables de la répétition GT, tests d'expression allélique de *GHR* et analyse de l'expression des gènes de l'axe GH/IGF-1 par PCR quantitative), j'ai pu montrer que le microsatellite GT agit de façon spécifique selon le contexte et le sexe et constitue un régulateur en *cis* de l'expression de *GHR* tout en exerçant potentiellement des effets en *trans* sur les gènes *IGF-1* et *BCL-2* (Chapitre IV).

Ces études ont montré l'importance d'examiner les différents types de variations génétiques dans le gène *GHR* et leurs associations avec la petite taille sévère et l'obésité et ouvrent de nouvelles perspectives sur les mécanismes régulant l'expression de *GHR*.

ACKNOWLEDGEMENTS

I would first like to express all my gratitude to my supervisor, Dr. Cindy Goodyer, for the opportunity to be a part of her team. Throughout the years, she has been a true mentor and her passion has been a real inspiration. Her patience and her incredible dedication while editing too many versions and always challenging me to push further. Her constant encouragement, trust and wisdom made that journey particularly rich for my scientific and professional growth.

Next, I would like to say how grateful I am to have had such a dedicated and collaborative committee, composed of Dr. Michael Witcher, Dr. Rob Sladek, Dr. Thomas Stroh and Dr. Maxime Bouchard for their constructive critiques and insightful comments. A special thank-you to Dr. Rob Sladek for his support and for granting me access to the CEPH cell lines as well as for giving me great advice regarding my experiments. I also want to thank Dr. Raman Agnihotram for his help and teachings for my statistical analyses.

My sincere thanks to past members of the CGG lab, Chris, Ayse and Alona, for the good times, support and the sharing of ideas and protocols as well as people around our lab who shared my day to day life during all these years giving me good advice and support.

I will always be grateful to my dear friend and colleague, Samar Elzein, with whom I had the pleasure to work these past years. I am grateful for her incredible help, not to mention the laughs shared during our time in the lab together; and of course, outside the lab, in some memorable trips to yurts and cabanes perchées!

Thank you to all of my friends, who supported me throughout this process, for their wonderful and cheerful company in the lab and outside. A special call out to Maha, Rasha, Maira Christine and Lucie for the listening and good times, and Annie for her positive energy and wisdom.

None of this would have been possible without the unwavering support and love of my family who believed in me from the start to the end of this project. I would like to thank, with all my

heart, my parents, Lucien and Monique Dias, my sister, Séverine, and of course my niece, Eva, who have been the guiding stars in my life.

Finally I will end with my boyfriend, Nathaniel, who has walked with me throughout and gave such wonderful support by listening about all the experiments and troubleshootings. In helping me to understand Linux and its secrets, he enabled me to perform a significant part of the statistical analysis. Thank you for everything with all my love. And thank you for having insisted on getting a dog, Cora, who helped me so much in keeping a healthy mind by always reminding me to go for walk breaks during the long hours of writing!

CONTRIBUTIONS TO ORIGINAL KNOWLEDGE

Chapter II: Here I contributed to the identification of a new set of potential markers in the *GHR* gene that could help to better define subgroups within the ISS heterogeneous population. I identified two GT genotypes, the L/S and the S/M, that were associated with pediatric idiopathic short stature with sex-specificity. I also showed, in the Montreal ISS cohort, that the L allele carriers were the ones showing the least catch-up growth compared to the non-L carriers. Moreover, a set of other potential markers (rs4292454 located in intron 2 and a risk haplotype at the 3' end of the *GHR* gene) were associated with an increased risk for the ISS phenotype. Conversely, a set of SNPs and a risk haplotype located in the promoter region of the *GHR* gene were associated with the adult short stature phenotype.

Chapter III: In this series of case-control analyses, I demonstrated that short stature Canadian women had a significantly higher prevalence of obesity compared to normal height women while no difference was observed in men. This sex-specific association translated at the *GHR* level to a specific cluster of SNPs in the regulatory regions of *GHR* when adiposity indices effects were taken into account. The L/M GT genotype showed a consistently stronger association with short stature women for multiple adiposity indices. I also found that genetic variants located within the *GHR* gene promoter participate in the etiology of both short stature and obesity.

Chapter IV: For the first time I provided a functional analysis of the *GHR* GT microsatellite and showed evidence of a role as a fine-tuning modulator of *GHR* transcriptional activity in a context- and sex-specific manner. I also developed a new assay using droplet digital PCR technology to measure the level of allelic imbalance expression and showed a high level of differential allelic expression in *GHR* mainly in the males. Finally, I demonstrated a *continuum* of low transcriptional activity in *IGF-1* and *BCL2* in individuals carrying the GT L/M genotype in a sex-specific manner.

CONTRIBUTIONS OF AUTHORS

Chapter II: "Genetic Variations at the human *growth hormone receptor* (*GHR*) gene locus are associated with idiopathic short stature" by Christel Dias, Mara Giordano, Rosalie Frechette, Simonetta Bellone, Constantin Polychronakos, Laurent Legault, Cheri L Deal and Cynthia Gates Goodyer (*J Cell Mol Med* 21(11); 2985-2999, 2017).

CD was responsible for experimental design, data collection and analysis and writing of the manuscript; RF contributed the genotype analyses; MG and SB provided the Novara ISS and control cohorts while CP, LL and CLD provided the Montreal ISS children and control adults; CGG was the principal investigator and coordinator of the projects, reviewer of the data, analyses and manuscript drafts; all authors reviewed and accepted the final manuscript.

Chapter III: "Sex-specific increase in the risk of obesity with short stature: insights from a Canadian cohort" by Christel Dias, Agnihotram V Ramanakumar and Cynthia Gates Goodyer (Submitted to *Am J Physiol Endo Metab* in August 2018).

CD was responsible for the experimental design, data collection and statistical analyses and writing of the manuscripts; AVR provided critical advice for the statistical analyses; CGG was the principal investigator and coordinator of the project, reviewer of the data, analyses and manuscript drafts; all authors reviewed and accepted the final manuscript.

Chapter IV: "A GT microsatellite polymorphism in the human *growth hormone receptor* (*GHR*) gene affects its transcriptional activity" by Christel Dias, Samar Elzein, Robert Sladek and Cynthia Gates Goodyer (to be submitted in September 2018).

CD was responsible for the experimental design, data collection and statistical analyses and writing of the manuscripts; SE contributed to the quantitative PCR assays; RS provided the majority of the lymphoblastoid cell lines examined as well as critical advice for the analyses and served on CD's thesis committee; CGG was the principal investigator and coordinator of the project, reviewer of the data, analyses and manuscript drafts; all authors reviewed and accepted the final manuscript.

CD received doctoral studentship awards from the Montreal Children's Hospital Research Institute and the Fonds de recherche du Québec – Santé (FRQ-S). The three projects were funded by the Canadian Institutes for Health Research (CIHR) MOP-119313.

CHAPTER I

GENERAL INTRODUCTION AND LITERATURE REVIEW

1. The GH/IGF-1 axis: physiological and pathophysiological implications

A. From the anterior pituitary to the target organs

a) Human GH: synthesis, secretion and regulation

(i) GH is a heterogeneous protein: GH is a peptide hormone also called somatotropin that is found mainly in two isoforms in humans: the major bioactive form of 22kDa (191aa) that represents ~90-95% of the circulating hormone and a minor form of 20kDa (176 aa) representing ~5-10% (Baumann 2009). The human GH gene produces a family of proteins: its single genetic locus contains a cluster of five highly related genes (~94% of sequence homology) on chromosome 17 that originated from gene duplication events during primate evolution (Baumann 1991; Baumann 2009; Chen et al. 1989). The first gene, GH-N (GH1), is expressed primarily by somatotroph cells in the anterior pituitary and encodes both the 20 and 22kDa GH forms by alternative splicing. GH-V (the GH-variant gene or GH2) codes for the placental 22kDa and glycosylated 25kDa GH isoforms (Newbern and Freemark 2011). Two other genes, CS-A and CS-B are chorionic somatomammotropins or placental lactogens (PLs). The fifth gene, CS-L (CS-like gene) is a pseudogene. The four placental GH/CS genes are strictly produced and released by the placental villus syncytiotrophoblast cells and are, thus, expressed solely during pregnancy (Hoshina et al. 1982). GH-V is found exclusively in the maternal serum from the sixth week; by mid-gestation it has totally replaced the maternal pituitary GH. CS-A and CS-B are secreted into both maternal and fetal circulations (Baumann 2009; Frankenne et al. 1987; Newbern and Freemark 2011). GH-V and the CS hormones have differential affinities for the GH receptor (GHR): GH-V has high affinity for GHR while CS-A and CS-B have low affinity for GHR and high affinity for the prolactin receptor (PRL-R) (Newbern and Freemark 2011). Using mouse lines that carry a transgene containing the human GH locus, it has been shown that the human GH gene cluster is under the control of a complex locus control region (LCR) located 14-32kb upstream of GH-N

(Vakili et al. 2011). The LCR is comprised of five DNAse hypersensitive sites (HS), with HSI and HSII being pituitary-specific while HSIII-V are involved in placental-specific regulation (Jones et al. 1995). Three pituitary-specific POU homeodomain factor Pit-1 binding elements co-localizing in HSI were shown to be necessary for *in vivo* GH-N activation (Ho et al. 2004; Shewchuk et al. 2002). Combinations of local and remote regulatory regions at the GH/CS locus have been shown to control tissue specific expression through epigenetic changes and long range DNA interactions (Cattini et al. 2006; Ho et al. 2004; Jin et al. 2018).

(ii) A pulsatile and sex-specific secretory pattern for GH-N: In humans, pituitary GH shows a pulsatile secretion with peaks occurring between 7 to 10 times in a 24h period, with a major nocturnal surge shortly after sleep onset during the REM phase (Sassin et al. 1969; Takahashi et al. 1968; Van Cauter et al. 2004). GH has a short half-life in serum (~20min) and 50% is bound to the GH binding protein (GHBP). The highest levels of serum GH occur during puberty: the growth spurt in both genders results from an interaction of the gonadal axis and the GH/IGF-1 axis whereby increased levels of sex steroids stimulate GH production (Christoforidis et al. 2005; Rogol 2010). In young adult men (<40y), the GH night pulse is very large and accounts for most of the GH secreted during the 24h period (70%). In pre-menopausal women, the amounts of total GH secreted in a day is similar to the men but the daily pulses are more frequent with higher basal GH levels than in males; moreover, the GH pulse at sleep onset accounts for less of the total GH than in men (Ho et al. 1987; Jaffe et al. 1998; Van Cauter et al. 2004). GH secretion decreases with age (somatopause) in both sexes, gradually for males after the third decade and more drastically for females at the onset of menopause due to decreased estrogen levels (Veldhuis 2008). The sex specificity of the GH secretory pattern is more marked in rodents, with a pulsatile secretion in males with high amplitude peaks every ~3 hours and undetectable GH levels during trough periods vs. a more continuous secretion pattern marked by lower amplitude pulses in females (Lichanska and Waters 2008a; MacLeod et al. 1991; Steyn et al. 2011; Tannenbaum and Martin 1976). The secretory pattern of GH can by influenced by numerous factors, including other hormones (e.g. thyroid hormones, glucocorticoids), body composition (glucose levels, fat mass), nutrition, exercise, sleep and stress (see **Figure I-1**) (Steyn et al. 2016).

(iii) Regulation of GH release at the pituitary level: The expression and rhythmic release of GH are regulated primarily by two antagonistic peptide hormones: Growth Hormone-Releasing Hormone (GHRH) and Somatostatin (SST); both are secreted by hypothalamic

neurons into the median eminence portal vascular system supplying the anterior pituitary (Lichanska and Waters 2008a). These two hormones show a synchronous and alternate pulse of release that likely constitutes the driving force of pulsatile GH release (Cataldi et al. 1994).

GHRH is a 44 aa peptide that was first isolated from a human pancreatic tumor causing acromegaly (Guillemin et al. 1982); it is the product of a 108 aa pro-hormone and GHRH 1-29 is the most bioactive fragment. GHRH is expressed primarily in the arcuate nucleus and the ventromedial nucleus in the hypothalamus. It stimulates the production and release of GH from the pituitary somatotroph cells by binding to a seven transmembrane domain G-protein ($G\alpha_s$) coupled receptor, the GHRH-R, at the surface of the cells. The $G\alpha_s$ subunit activates the adenylate cyclase/PKA pathway which has a direct effect on *GH* gene transactivation and the L-type Ca^{2+} channels; the increased intracellular calcium concentration results in GH release. Both the pulsatile pattern of GH in serum as well as the amplitude of the episodic peaks are directly linked to the GHRH pulsatile stimulation.

Inactivating mutations in the GHRH-R cause GH deficiency leading to short stature and metabolic disorders while activating mutations in the α subunit of the G-protein ($G\alpha_s$) cause excess of GH synthesis and secretion leading to gigantism in children and acromegaly in adults. The $G\alpha_s$ activating mutations can also cause tumors of the anterior pituitary (Gadelha et al. 2017).

Somatostatin, also known as somatotropin release-inhibiting factor (SRIF), was originally isolated from the ovine hypothalamus (Brazeau et al. 1973; Schally et al. 1976). It is widely distributed throughout the hypothalamus and the rest of the brain, as well as peripheral nervous systems where it influences gastro-intestinal tissues and pancreatic islets. Two peptides are produced from a 116aa prohormone: SST-14 and SST-28. SST-14 is the most abundantly distributed active form in the brain while SST-28 is mostly found in the digestive tract. It is an antagonist of GHRH (and Ghrelin) by directly inhibiting hypothalamic GHRH release into the portal system and by inhibiting GH production and release from the somatotroph cells. It also inhibits the synthesis and the release of other pituitary hormones (TSH and PRL). Like GHRH, it binds a G-protein coupled type of receptor but with different G-coupled proteins, $G\alpha_i$ (to inhibit adenylate cyclase) and $G\alpha_o$ (to inhibit Ca^{2+} channels), that results in suppression of GH release. There are 5 different genes encoding the different subtypes of the somatostatin receptor (SSTR1-5). SSTR2 and 5 are predominant on the

pituitary somatotroph and selective agonists (e.g. octreotide) for these receptors are used to treat GH overexpression disorders such as acromegaly.

There is a third major player in the regulation of GH secretion. Ghrelin, a 28aa peptide octanoylated on serine 3, was cloned in 1999, 3 years after its receptor, the growth hormone secretagogue (GHS) receptor (Goldenberg and Barkan 2007; Howard et al. 1996). The GHSR is a seven transmembrane G-protein coupled receptor ($G\alpha_{11}$) linked to phospholipase C/PKC pathways and is widely expressed in GHRH neurons and pituitary somatotrophs (Goldenberg and Barkan 2007; Howard et al. 1996). Ghrelin is a peripheral hormone primarily expressed in the stomach but it is also produced in neurons within the arcuate nucleus; it stimulates GH secretion directly from the somatotroph or indirectly by increasing hypothalamic GHRH release (Ariyasu et al. 2001; Goldenberg and Barkan 2007). Ghrelin also stimulates food intake, appetite and fat deposition in humans and rodents and has been viewed as a promising anti-obesity target although its physiological roles remain a debate in the human (Chanoine et al. 2009; Murphy and Bloom 2006; Nakazato et al. 2001; Sun et al. 2003). However, mice with a knockout of GHSR showed low IGF-1 and body weight and a resistance to diet-induced obesity (Zigman et al. 2005). Moreover, missense mutations in the GHSR have been associated with familial short stature (Pantel et al. 2006) while polymorphisms in the GHSR have also been associated with variation in height (Lango Allen et al. 2010; Lanktree et al. 2011), constitutional growth delay (Pugliese-Pires et al. 2011) and obesity (Wang and Tao 2016).

(iv) Feedback regulation of GH release **(Figure I-1):** GH secretion is regulated at multiple levels. The pituitary somatotrophs are regulated by both GHRH stimulation and somatostatin inhibition. There is an 'ultrashort' feedback loop involving somatostatin and GHRH neurons at the hypothalamic level through axonal projections from the periventricular to the arcuate nucleus. Once GH is released into the general blood circulation, it exerts a 'short' excitatory feedback on somatostatin-producing neurons, resulting in inhibition of the somatotrophs. A 'long' negative feedback effect on GH and GHRH is exerted by peripheral IGF-1 produced through GH binding to and stimulation of its target organs (e.g. liver). Numerous peripheral regulatory factors also exert stimulatory or inhibitory effects on somatotrophs or the neuroendocrine neurons producing GHRH and somatostatin.

b) GH actions on target organs

Since its isolation in 1944, GH has been known to have a central role in musculoskeletal development in the child but also important regulatory effects on protein, carbohydrate and lipid metabolism at all stages of life (Lichanska and Waters 2008a; Lichanska and Waters 2008b; Veldhuis et al. 2005) (**Figure I-2**). By binding to its specific receptor, GH triggers a cascade of signaling pathways conveying its biological actions within different cells and tissues (Lichanska and Waters 2008b). GHR is widely expressed in the human body and is found in significant levels in the liver, adipose tissue, kidney, heart, skeletal muscle, lymphocytes, fibroblasts and prostate (Ballesteros et al. 2000a). GH can act on target cells directly or indirectly by insulin-like growth factor 1 (IGF-1), the main mediator of GH actions in the body. GH peak amplitude has been shown to correlate with serum IGF-1 concentration and growth rates while there was no correlation observed with basal (interpulse) GH release (Butler and Le Roith 2001; Maiter et al. 1988).

(i) IGF-1 as the GH messenger and the IGF-1 system: Most of the serum IGF-1 concentration is produced by the liver through GH stimulation (Liu and LeRoith 1999; Roberts et al. 1987). In the 1950s, hepatic IGF-1 was thought to be the only endocrine mediator of postnatal longitudinal growth and that concept was termed "the somatomedin hypothesis"(Salmon and Daughaday 1957). In the 1980s, the discovery of 'extra-pituitary' production of GH and extra-hepatic production of IGF-1 challenged this hypothesis and led to the dual effector theory in which GH exerts its effects through both direct and indirect effects (mediated by IGF-1) (Green et al. 1985). Studies showed direct effects of GH on chondrocytes (Isaksson et al. 1982; Schlechter et al. 1986) and on differentiation of prechondrocytes (Ohlsson et al. 1992) leading to longitudinal growth at the epiphyseal growth plate (Lindsey and Mohan 2016). Later, a conditional deletion of hepatic IGF-1 in mice revealed an identical growth phenotype of the transgenic and wild type mice despite a drastic reduction in serum IGF-1, providing evidence of a significant local autocrine/paracrine effect of IGF-1 on growth (Le Roith et al. 2001; Yakar et al. 1999).

IGF-1 and IGF-2 are growth factors widely expressed throughout fetal and postnatal development (Han et al. 1988; Roberts et al. 1987); their genes as well as the *insulin* gene are derived from a common ancestral 'insulin-like' gene in vertebrates. Both IGFs bind to their respective insulin-like receptors (IGF-1R and IGF-2R) but IGF-1R is the major receptor mediating both IGF-1 and IGF-2 actions. IGF-1R is a transmembrane receptor tyrosine kinase

arranged in a α2β2 configuration which shares similarity with the insulin receptor. Upon IGF-1 binding to its receptor, IGF-1R autophosphorylates and then recruits adaptor proteins, mainly insulin-receptor substrate 1 (IRS-1). Both IGFs act as endocrine (with insulin-like effects), paracrine and autocrine factors in the regulation of growth, development and metabolism (Kadakia and Josefson 2016; Murray and Clayton 2013). The circulating IGFs are coupled to high affinity IGF binding proteins (IGFBP1-6) that prolong IGF half-lives, by protecting them from proteolytic degradation, and that do not bind insulin (Allard and Duan 2018; Bach 2018). IGFs are largely found in the circulation in a form of a high molecular ternary complex (~150kDa) composed of IGFBP3 (less often IGFBP5) and a glycoprotein called acid labile subunit (ALS); the added ALS prevents the crossing of the capillary barrier, thus maintaining a reservoir of circulating IGFs and modifying tissue IGF-1 availability (Silha et al. 2001). IGFBP3 is the predominant binding protein for IGF-1 and binds 75% of all the IGFs. The IGFBPs are expressed by most of the GH target tissues and have important regulatory roles as they bind to the IGFs with equal or more affinity than their receptors (Allard and Duan 2018; Mohseni-Zadeh and Binoux 1997). Recently these BPs have also been shown to exert IGF-independent actions, including modulation of other growth factor pathways, nuclear translocation and transcription regulation (Allard and Duan 2018; Bach 2018). Hepatocytes express high levels of the *ALS*, *IGF-1* and GHR transcripts while IGFBP3 is expressed by the endothelial cells of the hepatic sinusoids; all are stimulated by GH (Butler and Le Roith 2001; Chin et al. 1994).

Although GH has effects on many tissues, in this overview I will concentrate on bone and adipose tissues as they are the major focus of Chapters 2 and 3.

(ii) GH action on skeletal formation in growth: GH, IGF-1 and IGF-2 are key factors of growth during development. IGF-2 is especially critical during fetal development while GH and IGF-1 are the main regulators of skeletal growth and maintenance during postnatal life. IGF-1 null mice show severe growth retardation (~30% of wild type size) with decreased chondrocyte proliferation and differentiation and decreased bone mineral density (BMD) (Yakar et al. 2018).

In the healthy child, skeletal linear growth involves both chondrogenesis (cartilage formation) and osteogenesis (bone formation); these processes are regulated by genetic, nutritional and hormonal factors, with essential roles for GH, IGF-1 and thyroid hormones (**Figure I-3**) (Ohlsson et al. 1992; Yakar and Isaksson 2016). During puberty, increases in GH and

estrogen in both sexes result in a rise in serum IGF-1 and IGFBP3 levels as well as local tissue IGF-1 levels, leading to enhanced chondrogenesis in the epiphyseal growth plate (long bone growth) and bone mineralization (Christoforidis et al. 2005; Grumbach 2000). Skeletal maturation involves the progressive ossification of the epiphyseal growth plate, leading to epiphyseal fusion that ends the growth spurt and results in final adult height.

GH and IGF-1 also contribute to the acquisition of bone mass during childhood and puberty and to the peak bone mass occurring by the average age of 16 years in girls and 17 years in boys (Kasukawa et al. 2004; Lindsey and Mohan 2016). Once the skeleton has reached maturity, bone is still actively maintained by constant remodeling; this involves a balance between the formation of new bone by osteoblasts and removal of bone (resorption) by osteoclasts, leading to complete replacement of the adult skeleton every 10 years (Manolagas 2000). An imbalance between these two cell type activities can lead to osteoporosis or osteosclerosis. IGF-1 (from endocrine and local sources) has been shown to be an important factor for coupling the two activities responsible for normal bone remodeling (Yakar and Isaksson 2016).

***(iii) GH actions on adipose tissue*:** Effects of GH on lipid metabolism have been extensively studied in humans and rodents and there is a general consensus that GH levels negatively correlate with adiposity (Chaves et al. 2013). In addition humans (Laron syndrome) or transgenic mice with *GHR* deletions are severely decreased in height/length but also have significantly higher percent body fat (%BF) throughout their life, marked by increased central obesity and elevated cholesterol levels, decreased lean mass and smaller organs (Laron 2004b; Troike et al. 2017)

Adipose tissue is complex and recognized as a dynamic endocrine organ as it secretes and responds to a multitude of hormones and cytokines, including GH. It is found in discrete locations in the human body in the form of depots. The white adipose tissue (WAT) is associated with energy storage (triglycerides) and can be subcutaneous (gluteal, femoral, subcutaneous superficial or deep) or intra-abdominal (mesenteric or omental, mostly lining the internal organs) (Troike et al. 2017). There are also depot-specific differences in cell morphology, composition, receptor abundance and secretory profile (Troike et al. 2017).

Mature adipocytes are not the major cell type in normal WAT depots; the non-adipocyte cells include fibroblasts, preadipocytes, immune cells, neural cells and endothelial cells. Adipose growth occurs through increase in size (hypertrophy) or in number (hyperplasia) of adipocytes

following maturation of pre-adipocytes (adipogenesis) (Troike et al. 2017); adipogenesis occurs *via* stimulation of peroxisome proliferator activated receptor gamma (PPARγ), CCAAT/enhancer binding proteins (C/EBPs) (Lefterova et al. 2008) and hormones, including GH. Our lab has reported an increase in *GHR* mRNA expression during pre-adipocyte differentiation (Wei et al. 2006). In addition, Erman et al showed that lean women had higher *GHR* mRNA levels in omental fat compared to subcutaneous. This depot-specific difference was lost in obesity: total *GHR* expression for both depots was decreased compared to levels in lean women (Erman et al. 2011a).

Through binding to GHR at the surface of the mature adipocytes, GH can modulate their function by promoting lipolytic and anti-lipogenic effects. These effects lead to an acute rise in serum free fatty acids (FFAs) and glycerol available for energy expenditure. Contrary to a more indirect effect of GH through IGF-1 in bone, GH primarily exerts direct effects on the catalytic activity of several enzymes in the adipose tissue. GH mediates its anti-lipogenic effects by suppressing lipoprotein lipase (LPL) activity, which hydrolyses the triglycerides from circulating very low density lipoproteins (VLDL) and chylomicrons into FFAs, thus reducing their uptake and further storage in adipocytes (Moller and Jorgensen 2009; Ottosson et al. 1995; Richelsen et al. 2000). GH exerts lipolytic effects through increasing hormone sensitive lipase (HSL) activity, which hydrolyzes the stored triglycerides into FFAs and glycerol. This stimulates FFA transport from adipose tissue to the liver and muscle where they can be used as fuel. The glycerol is then taken up by the liver or the kidney and rejoins the glycolysis or gluconeogenesis pathways, decreasing insulin sensitivity. The lipolytic actions of GH in adipocytes are the result of an upregulation of β3-adrenergic receptor expression; this is a G-coupled protein receptor stimulated by catecholamines that will activate the HSL through the cAMP/PKA pathway (Dietz and Schwartz 1991; LeRoith and Yakar 2007). While IGF-1 has been shown to play an important role in pre-adipocyte differentiation, IGF-1R expression decreases during this process although mature adipocytes still express IGF-1 (Scavo et al. 2004; Zizola et al. 2002).

c) Dysregulation of the GH/IGF-1 axis: human pathophysiology

***(i) GH Deficiency and Insensitivity*:** In humans, genetic defects affecting members of the GH/IGF-1 axis can lead to a spectrum of growth disorders ranging from severe growth retardation (‘dwarfism’) to “short normal” stature. Before investigating the GH/IGF-1 axis, pediatric endocrinologists have to rule out thyroid deficiency, excess glucocorticoids,

malnutrition, social deprivation, specific genetic syndromes, chronic diseases and intrauterine growth retardation (IUGR) or other syndromes affecting growth. Short stature can be caused by GH deficiency (GHD) or a resistance to the actions of GH, termed GH insensitivity (GHI). The different types of defects involving genes in the GH/IGF-1 axis and disrupting linear growth are presented in **Table I-1** (Savage et al. 2011).

GHD occurs in ~1/4000-10 000 children and accounts for ~1% of children with short stature; it can be a congenital or acquired condition. Between 3 to 30% of the cases have a genetic origin and variants can affect the production, release and/or the functional activity of GH, leading to isolated growth hormone deficiency (IGHD) or combined growth hormone deficiency (CGHD) that is associated with other pituitary hormone deficiencies. CGHD involves mutations in pituitary transcription factors, including (a) POU1F1 (Pit-1) that causes deficiencies in GH, PRL and TSH, (b) PROP-1 (Prophet of Pit-1) that is associated with deficiencies in GH, PRL, TSH, FSH, +/- ACTH, or (c) factors involved in embryonic development of the anterior pituitary (e.g. HESX1, LHX3 and LHX4) (reviewed in (Wit et al. 2016). The majority of the mutations causing IGHD are related to *GH-N* and *GHRH receptor* gene defects; these account for ~10% of the GHD individuals. No mutations have been found in the *GHRH* gene to date. The children with congenital GHD have stunted growth, small head circumference, delayed bone age and puberty, truncal obesity and are at higher risk for hyperlipidemia. Acquired GHD usually results from hypopituitarism due to pituitary adenomas or after pituitary surgery or radiation (Ayuk and Sheppard 2006; Smuel et al. 2015).

GHI is a heterogeneous disorder: the cases present with variable phenotypes resulting mostly from defects in the GHR but also from defects in GHR signaling. These defects occur more rarely than in GHD and affect only several hundred individuals worldwide (Savage et al. 2011). The first example of GHI, and the most extreme phenotype, was reported in 1966 by Laron et al who described three siblings with symptoms resembling hypopituitarism but with high levels of GH and the absence of binding of iodinated (^{125}I)hGH to GHRs prepared from liver membranes (Eshet et al. 1984; Laron et al. 1966). This 'Laron syndrome' is caused by homozygous mutations or deletions in the *GHR* gene resulting in loss of GHR function (Godowski et al. 1989). Patients are diagnosed soon after birth because of metabolic instability and a rapid loss in growth rate, have extreme short stature, very low serum IGF-1, IGFBP3, ALS and GHBP levels, increased BMI, musculoskeletal abnormalities and hypoglycemia (Laron et al. 1966). Untreated patients with Laron syndrome have extremely

short stature, between 5-12 SD below the mean normal height, and can present with different degrees of obesity associated with a decreased lean mass (Laron 2004b; Savage et al. 2006).

Around 70 mutations, homozygous or compound heterozygous, have been described in GHR causing different degrees of GHI, including point mutations, missense, nonsense and splice mutations, with mostly autosomal recessive inheritance (Savage et al. 2006; Savage et al. 2011). Mutations in the GHR extracellular domain can affect GH binding and can lead to absent or extremely low levels of serum GHBP; mutations in the transmembrane domain can modulate receptor dimerization and membrane anchorage; and mutations in the intracellular domain may disrupt signal transduction (Savage et al. 2006; Savage et al. 2011). GHR heterozygous mutations have been identified that result in truncated GHRs, due to aberrant alternative splicing; these can have a dominant negative effect and will present with abnormally high levels of GHBP (Ayling et al. 1997; Iida et al. 1998). A point mutation causing the inclusion of an intronic pseudoexon results in an additional 108nt between exons 6 and 7 and potential defects in receptor dimerization; this was associated with a milder GHI phenotype (Metherell et al. 2001).

Mutations have also been found in members of the key GH-stimulated JAK2/STAT5b (signal transducer and activator of transcription 5b)/IGF-1 intracellular pathway. Mutations in STAT5b are very rare and cause GHR signaling defects that result in severe post-natal growth retardation, complete GHI associated with severe IGF-1 deficiency, and moderate to severe immunodeficiency. The combined impairment of the immune system comes from the importance of STAT5b in the signaling of several cytokines (e.g. IL-2, IFNγ). Seven homozygous inactivating mutations have been described in 10 patients worldwide (Hwa 2016).

Lastly, defects in IGF-1 and IGF1-R have been reported but are the rarest involved in the GHI syndrome, especially IGF-1 mutations. Homozygous and heterozygous IGF-1 mutations are associated with a variable degree of pre- and post-natal growth failure, microcephaly, mental retardation and sensorineural deafness. IGF-1 resistance has been found in some patients presenting with IGF1-R mutations that are mainly heterozygous or compound heterozygous. Haploinsufficiency of the *IGF-1R* gene is associated with impaired intrauterine and postnatal growth (Domene and Fierro-Carrion 2018). Individuals with homozygous *IGFALS* (gene coding for ALS) mutations present with severe deficiencies in circulating ALS and extreme deficiency in IGF-1 and IGFBP3; this is caused by their increased clearance due to failure to

form the ternary complex with ALS. Even with extremely low IGF-1 levels, these individuals present with mild growth failure (2-3 SD below the mean) (Domene and Fierro-Carrion 2018; Savage et al. 2011).

***(ii) Partial GHI: Idiopathic Short Stature (ISS)*:** ISS is more a clinical description than a disorder because it has an undefined etiology and is based on a diagnostic by exclusion. The ISS category represents a heterogeneous group of short stature individuals, including those who are normal variants of short stature (e.g. familial short stature or constitutional growth delay) (Wit et al. 2008). The clinical features of these children include a short stature of more than 2 SD below the mean, low to low normal serum IGF-1, with a relatively high BMIbut with no evidence of GH deficiency, hypothyroidism, malnutrition, intrauterine growth retardation (IUGR), systemic disease or specific syndromes. These children are born with normal weights and have a relatively normal growth velocity but at the lowest part of the growth curve, below the 3rd percentile on the growth curve. Half of the children with ISS will reach their genetic potential for height at puberty and are re-diagnosed as having had a constitutional growth delay due to a delay in their growth spurt (Wit et al. 2008).

Serum levels of GH binding protein, the product of enzymatic cleavage of cell-surface GHR, are below the mean in 90% of the cases and 20% have levels >2 SD below, suggesting that tissue levels of GHR are chronically low in many of these children. These findings indicate a partial GHI and potential abnormalities affecting the *GHR* gene (Attie et al. 1995; Carlsson et al. 1994). In ~5% of the ISS children, heterozygous and compound heterozygous mutations in *GHR* coding exons have been identified, mostly in the extracellular domain region but there are conflicting results concerning their functional relevance (Bonioli et al. 2005; Hujeirat et al. 2006). Post-receptor mutations have been associated with some ISS cohorts: individuals presenting with *IGFALS* haploinsufficiency have lower levels of ALS, IGF-1 and IGFBP3 than normal height individuals, leading to 1 SD height loss compared to normal height, while homozygous or compound heterozygosity mutations lead to a further loss of 1.0 to 1.5 SD suggesting a gene dosage effect (Domene et al. 2013; Fofanova-Gambetti et al. 2010).

(iii) Over-stimulation of the GH/IGF-1 axis: Acromegaly and Cancer: Acromegaly is caused by chronically elevated levels of pituitary GH synthesis and secretion leading to increased levels of IGF-1. If excessive levels of GH and IGF-1 occur during childhood before the fusion of the epiphyseal growth plate then there will be excessive growth leading to gigantism if no treatment; for example, Robert Wadlow's height at 22 years old was 2.72m

(Carter-Su et al. 2016). It is a very rare disorder with 5-9 cases per 100,000 with 5% of the cases being due to an over-secretion of GHRH caused by an ectopic or hypothalamic tumor and 40% to a pituitary adenoma due to activating mutations in the somatotroph Gαs proteins (Capatina and Wass 2015). Acromegaly can also be a part of certain syndromes, like familial isolated pituitary adenoma (FIPA) or multiple endocrine neoplasia (MEN1 and MEN4) syndromes (Gadelha et al. 2017). The over-stimulation of the GH/IGF-1 pathway leads to reduced fat mass, increased lean mass, impairments in glucose homeostasis resulting in insulin resistance and type 2 diabetes, and increased risk of cancers (thyroid, breast, colon), mortality and morbidity (Troike et al. 2017; Wolinski et al. 2017).

GH and IGF-1 promote the normal growth of tissues, through pro-proliferative, anti-apoptotic and pro-angiogenic effects. These actions are critical during childhood growth, especially for the long bones, but they are also important throughout life for those tissues where there is a continual turnover of cells (e.g. intestinal villous epithelium). However, when either of these growth-promoting factors or their receptors are chronically over-expressed they promote cell hyperplasia and tumor progression, as seen in some acromegalic patients (Brooks and Waters 2010b; Lichanska and Waters 2008a; Lichanska and Waters 2008b; Perry et al. 2006; Wolinski et al. 2017). The association of GH and IGF-1 with cancer has been shown in both human and animal models. Several studies have shown that IGF-1 promotes tumor proliferation, angiogenesis and metastasis and, with its receptor, has been a major focus for developing new therapies (Bruchim et al. 2009). Elevated *GHR* transcripts and protein expression have been reported in different tumors, including colorectal, breast and prostate cancers (Bidosee et al. 2009; Gebre-Medhin et al. 2001; Wu et al. 2007). Moreover, Weiss-Messer et al. observed a decrease in the ratio of the truncated/full length *GHR* mRNAs in prostate cancers suggesting that loss of the dominant negative receptor isoforms may also play a role in tumor progression (Weiss-Messer et al. 2004). Although loss of function mutations have been described in the GHR, as mentioned earlier, no activating mutations have been identified (Brooks and Waters 2010b; Perry et al. 2006); however, polymorphisms in the *GHR* gene have been associated with increased risk for lung and prostate cancer as well as decreased breast cancer risk (McKay et al. 2007; Rudd et al. 2006; Van Dyke et al. 2009; Wagner et al. 2006).

The cross-talk of the GH/IGF-1 axis with cancer is also supported by studies relating final height with cancer incidence: a taller height (>175cm) has been associated with an increase in breast, prostate and colorectal cancer relative to shorter people (<160cm) (Brooks and Waters

2010b; Gunnell et al. 2001). The importance of the GHR in cancer has been demonstrated through follow-up analysis of cohorts of Laron patients with dysfunctional GHRs that have shown no evidence of a malignancy. Only one case of a non-lethal malignancy has been reported from the two cohorts analyzed (~350 individuals) whereas 8 to 17% incidence of malignancies was reported in their first to fourth degree relatives (Guevara-Aguirre et al. 2011; Steuerman et al. 2011). Parallel reports have shown a significant reduction in the incidence of malignancies in the homozygous *GHR* knockout mouse as well as delayed progression of non-neoplastic lesions (Ikeno et al. 2009; List et al. 2011).

(iv) Treatments of GH/IGF-1 axis pathophysiology: In the 1950s, the first treatment for severe GH deficiency was carried out using extractions of GH from pituitaries of deceased people. The limited supply, in addition to the subsequent discovery of cases of transmitted Creuzfeldt-Jacob Disease through prion contamination, led to the discontinuation of its extraction in 1985. At the same time, the first human recombinant GH (rGH) was synthesized in *E.coli* by Genentech and approved for use in GHD children throughout North America. The 22kDa GH is given primarily by subcutaneous injections every day at night (to mimic the nocturnal GH pulse) although more-long acting forms of GH are in clinical trials (Christiansen et al. 2016). GHD children respond with an increase in height velocity and a catch up of height in the normal percentiles while treatment for GHD in both children and adults results in a reduction of fat mass, particularly abdominal fat mass, and an enhanced quality of life (Chaves et al. 2013). Initially, treatment in children was usually stopped when the gain in height was <2cm/year, but today it is recommended for life to maintain the positive effects of GH on the metabolism (Reh and Geffner 2010). While GH treatments significantly reduce body fat in the GH-deficient individuals, similar treatments of morbidly obese individuals who are not GH-deficient are unsuccessful (Troike et al. 2017).

Although secondary effects of GH are limited, the long term safety of its continuous use is not known and concerns have been raised concerning the potential risk of an increase of rare cancers (Allen et al. 2016). The administration of rGH in GHD or GHI individuals requires a tight follow-up of IGF-1 serum concentrations that should be kept under the normal upper limit to minimize the risk of cancers. rIGF-1 has been synthesized since 1986 and treatment of GHI became possible but only for a limited number of patients as it is not only a potent mitogen but can also cause various secondary effects, including hypoglycemia in 40% of the patients with GHI (controlled by administration with meals), hyperplasia of lymphoid tissue and increases in fat mass and BMI (Rosenbloom 2009; Savage et al. 2006). rIGF-1 has a

positive effect on height velocity but the final height obtained in GHI patients does not reach the level obtained with rGH treatment in GHD patients (Guevara-Aguirre et al. 1997).

Treatment by GH for ISS children was approved in the US in 2003 and in Canada in 2006. The effect of rGH in height gain in ISS is highly variable and fairly minimal (3.5-7.5cm after 4-7 years); it also depends on the dose of GH used as well as the IGF-1 deficit levels and age at the onset of treatment (Savage et al. 2010; Wit et al. 2005). Few studies have reported on the metabolic outcomes of ISS children treated with rGH: levels of IGF-1 and IGFBP-3 increase, there is an increase in insulin (transient insulin resistance) and a decrease in fat mass and glucose levels (Dahlgren 2011). Ethical and social questions have been raised concerning the treatment of ISS children by rGH (Ambler et al. 2013).

The GH excess in acromegalic patients can be treated surgically (transphenoidal) or by radiation or suppressed by dopaminergic and somatostatin analogs (e.g. octreotride). The somatostatin analogs are the most efficient pharmaceutical reagent as they can normalize the GH/IGF-1 levels in 70% of the cases and cause tumor shrinkage in 40% of the patients (Dineen et al. 2017). There is also the option of using a GHR antagonist, pegvisomant (G120K): this is a 22kDa GH with 9aa changes that can bind the receptor without activation as it prevents proper GHR dimerization; the conjugation with polyethylene glycol chains increases its half-life. It has a better effect in acromegalic patients than somatostatin analogs in reducing the production of IGF-1 and has shown promising results as an anticancer agent. It is more commonly used nowadays in conjunction with somatostatin analogues (Dineen et al. 2017).

B. The growth hormone receptor: from structure to receptor activation

a) The GHR structure

The human *GHR* gene spans ~300kb on chromosome 5 in region p13.1-p12 (Barton et al. 1989a; Godowski et al. 1989). The coding region is defined as exons 2 to 10 that encode the 638 amino acid (aa) GHR peptide comprising the mature receptor of 620 aa plus an 18 aa leader sequence. Exon 2 contains the translation start site and codes for the 18 aa signal peptide as well as the first 5 aa of the extracellular domain (ECD). Exons 3-7 mostly encode the 246 aa ECD. Exon 8 encodes the final 3 aa of the ECD, a 24 aa short hydrophobic transmembrane domain (TMD) and 4 aa of the intracellular domain (ICD). Exons 9 and 10 encode the major part of the ICD of 346 aa as well as the 3'untranslated region (**Figure I-4**).

The GHR is part of the Class I cytokine/hematopoietin receptor superfamily and its extracellular domain was the first to be cloned and crystallized (in complex with the GHBP) (de Vos et al. 1992; Leung et al. 1987). The structural resolution of the GHR ECD led to the identification of related receptors based on limited amino acid homology (15-35%) in a ~210aa region. This family presently includes ~40 receptors, including those for prolactin (PRL), erythropoietin (EPO), thrombopoietin (TPO) granulocyte-macrophage colony-stimulating factor (GM-CSF), and many interleukins (**Figure I-5**). Class II cytokine receptors (e.g. interferon-γ (IFNγ), IL-10) are similar to class I but lack the conserved WSxWS sequence motif.

The ECD region in the GHR is composed of two fibronectin type III β sandwich domains connected by a short flexible linker; each domain contains a tryptophan residue crucial for GH binding to its receptor through hydrophobic interactions (Brooks and Waters 2010a). Class I cytokine receptors are characterized by the consensus WSxWS sequence in their ECD just above the TMD, except for GHR which has a YGEFS motif (Argetsinger and Carter-Su 1996). Alanine mutational analysis of the GHR motif was shown to affect binding of GH and signal transduction (Baumgartner et al. 1994); in addition, three mutations in the YGEFS have been reported to result in Laron syndrome, showing the significance of this sequence for normal GHR activity (Brooks and Waters 2010a). The ECD of cytokine receptors also contains several conserved cysteine residues; GHR is the only member to contain seven, six of which are paired by disulfide bonds. Mutational analysis showed that these residues are critical for the interaction with GH (Bass et al. 1991). Finally, the GHR ECD contains five potential N-linked glycosylation sites; changing the five sites by site-directed mutagenesis did not block GH binding but resulted in a 20-fold reduced GH binding affinity (Harding et al. 1994).

These cytokine receptors are also characterized by the absence of intrinsic tyrosine kinase activity and the need to associate with non-receptor tyrosine kinases for signal transduction. A conserved hydrophobic proline-rich domain called Box 1 has been identified in the ICD of all the cytokine receptors; this lies just below the TMD (within 20 residues) and consists of eight residues in mammals (Zhu et al. 2001). Box 1 constitutes the interaction domain for JAK2, the cellular tyrosine kinase associated with GHR activation upon GH binding (Argetsinger et al. 1993) and site-directed mutagenesis studies have shown that it is critical for GH signaling (Wang and Wood 1995). Box 2 is a less conserved motif of 15 hydrophobic residues located just below Box 1 and contains a ubiquitin motif critical for internalization of the GHR; it is

thought to be required for full activation of JAK2 by GH. The GHR ICD also contains seven tyrosine residues that have been shown to be substrates for JAK2 and docking sites for STAT5 proteins (Colosi et al. 1993; Hansen et al. 1996; Wang and Wood 1995).

b) GHR activation mechanism

Previously, it was thought that the GH binding was causing the dimerization of the receptor and its further activation. A revised model resulted from several studies showing that, in the absence of GH, the majority of the GHRs on the cell surface are constitutively dimerized and dimerization alone was not sufficient to initiate signaling (Dehkhoda et al. 2018; Ross et al. 2001; Rowlinson et al. 1998). The constitutive dimerization of the GHR is mediated predominantly through interactions in the TMD; however, other regions, such as the extracellular dimerization domain, may provide specificity to homodimer formation (Yang et al. 2007). Using co-immunoprecipitation, fluorescence resonance energy transfer (FRET), bioluminescence resonance energy transfer (BRET) and X-ray crystallography techniques, it has been shown that GH, through its asymmetrical binding at the binding sites in the preformed constitutive GHR dimers, induces a conformational change of the GHR intracellular domains, including a rotation of one GHR subunit relative to the other, and a locking together of the extracellular receptor-receptor interaction domains (Brooks and Waters 2010a; Poger and Mark 2010). As a JAK2 peptide is bound to each of the two intracellular GHR ICDs, the structural reorientation following GH binding is transmitted through the TMD and results in a repositioning of the two JAK2s; the two kinase domains are in closer proximity, enabling trans-phosphorylation of tyrosine residues in each kinase domain and the activation of the JAK2s and associated signaling pathways (Brooks and Waters 2010a).

c) Signal transduction (Figure I-6)

As stated earlier, the Class I cytokine receptors rely on associated tyrosine kinases for signal transduction. Unlike other cytokine receptors, GHR associates primarily with JAK2 and, only rarely, with the remaining three members of the JAK family of cytosolic kinases (JAK1, JAK 3 and TYK2) (Dehkhoda et al. 2018). JAK2 binds to Box 1 of each GHR receptor monomer through its FERM (N-terminal 4.1, Ezrin, Radixin, Moesin) and Src homology 2 (SH2) domains, both located at the N-terminal region of the kinase and necessary for the interaction with the GHR ICD. The FERM domain has been shown to regulate JAK2 kinase activity and recently the crystal structure of the JAK2 FERM-SH2 domain has been solved (McNally et al.

2016). Activated JAK2s phosphorylate key tyrosine residues on the GHR ICD that act as docking sites for SH2 domain-containing signaling proteins (Derr et al. 2011). The main signaling pathways activated by GH involve the STATs (Signal Transducers and Activators of Transcription), the phosphoinositol 3' kinase (PI3K) and the extracellular signal-regulated kinase/mitogen-activated protein kinases (ERK/MAPK) (Carter-Su et al. 2016). The differential activation of these pathways is dependent on the cell type, emphasizing the pleiotropic roles of GH (Dehkhoda et al. 2018).

(i) The JAK-STAT pathway: This pathway constitutes the GH canonical activation pathway. The STAT family comprises seven members (STAT1-4, 5a, 5b and 6) that all include an SH2 domain and a conserved tyrosine residue at their C-terminus. They are the best characterized JAK targets and are implicated in the signaling of several cytokines. GH has been shown to induce the phosphorylation of STAT1, 3, 5a and 5b in several cell types through JAK2. STAT5a and b are encoded by two highly homologous genes and have been shown to be critical for metabolism, body growth, sexual dimorphism and stimulation of the expression of *IGF-1*, *ALS*, *SOCS2* (*Suppressor of Cytokine Signaling 2*) and *CIS* (*Cytokine-Inducible SH2-domain protein*) genes (Chia et al. 2010; Woelfle and Rotwein 2004). STAT5b has been reported to be the main mediator of GH actions as it is responsible for the GH-dependent postnatal body growth (Carter-Su et al. 2016). STAT5b-deleted mice have severe growth retardation, decreased IGF-1 levels and increased obesity, similar to what is observed in the GHI syndrome (Udy et al. 1997). Upon GH binding, cytoplasmic inactive STAT5 monomers are recruited to specific phosphorylated tyrosine containing regions in the GHR ICD through their SH2 domain, which is required for their subsequent tyrosine phosphorylation by JAK2 (Rowland et al. 2005). Once activated STAT5 monomers dissociate from the receptor, they form homodimers *via* their SH2 domains and translocate to the nucleus to regulate transcription of several key genes through binding to STAT5 binding sites. STAT1 and STAT3 can be activated directly through JAK2 without the requirement for receptor binding and can also heterodimerize (Brooks and Waters 2010b; Carter-Su et al. 2016).

(ii) The ERK/MAPK pathway: GH has been shown to activate this pathway by JAK2 phosphorylation of the SH2 domain of the Shc (Src homology 2 domain containing transforming protein 1) adapter protein. Phosphorylated Shc then activates adaptor proteins like Grb2 (Growth factor receptor-bound protein 2) and stimulates its association with the guanine nucleotide exchange factor, Sos (Son of sevenless), resulting in activation of the Ras/Raf/MEK pathway and the extracellular signal-regulated kinases, Erk1 and 2. These

effector kinases are involved in the phosphorylation of multiple substrates (on serine and threonine residues) in all cellular compartments, including the nucleus where they translocate and activate multiple transcription factors (e.g. Elk-1, C/EBPβ, c-jun and STATs), other kinases, cytoskeletal components and phospholipases (Carter-Su et al. 2016). It has been shown that GH stimulates C/EBPβ transactivation via p44/42 MAPK (Erk1/2) phosphorylation and regulates its nuclear relocalization (Piwien Pilipuk et al. 2003).

(iii) The PI3K pathway: This pathway is involved in cell cycle, survival, metabolism, cell motility and cancer metastasis (Wan et al. 2015). PI3K and the serine/threonine protein kinase mTOR pathways are activated by JAK2 *via* phosphorylation of IRS1, 2 and 3 (Insulin Receptor Substrate 1, 2 and 3), the same effectors involved in insulin and IGF-1 signaling. The activation of the IRS proteins has been implicated in insulin regulation of lipid metabolism and glucose transport and can explain the acute insulin-like effect of GH in the transient increase of glucose transport in adipocytes (Bergan-Roller and Sheridan 2018). Activated PI3K results in the generation of phosphoinositide products that act as second messenger molecules and activate various targets, including the anti-apoptotic serine kinase Akt (also called PKB). Activated Akt regulates different pro- and anti-apoptotic proteins substrates, including Bcl2 (B Cell Lymphoma 2) family members (Bax (Bcl-2 associated X protein) and Bad (Bcl-2 associated death promoter)), GSK-3, caspase 9 and NFκB. GH prevented muscle cell apoptosis in rats with cardiac heart failure by increasing Bcl2 expression and reducing Bax and caspase levels (Dalla Libera et al. 2004).

***(iv) The SRC pathway*:** There is evidence that activation of GHR pathways can occur independent of JAK2 *via* activation of Src family kinases (SFKs); the Lyn Src kinase has been found bound to the membrane proximal part of the GHR and a deficiency in Lyn activation results in impaired activation of Erk1 and 2. The relative activation of SFKs is dependent on the cell type (Rowlinson et al. 2008a).

d) Inhibitors of the GHR-JAK2-STAT pathway

Control of GH activation of the signaling pathways through GHR is of critical importance and three main classes of negative regulators participate in maintaining this balance, as any dysregulation of these proteins can lead to pathophysiological states.

***(i) Suppressors of cytokine signaling (SOCS)*:** This family is comprised of eight members, SOCS 1-7 and CIS, and is part of a larger family of cytokine-inducible proteins known as

STAT-induced STAT inhibitory proteins (SSI) (Wojcik et al. 2018). They are the most important class in the negative regulation of cytokine receptor signaling, inhibiting the activity of the JAK-STAT pathway by direct interaction with activated JAK proteins or the cytokine receptors (Zhang et al. 1999). SOCS protein levels are constitutively low but increase rapidly upon GH stimulation *in vitro* and *in vivo* (Greenhalgh et al. 2005; Wojcik et al. 2018). Only SOCS1-3 and CIS are activated by GH; *SOCS2* expression has been shown to increase through binding of STAT5b to its response element in intron 1 (Vidal et al. 2007). SOCS4-7 are more constitutively expressed and their precise biological roles are still unclear (Linossi et al. 2013).

The activity of the SOCS proteins relies mainly on their central SH2 domain, a conserved domain that interacts with phosphotyrosine containing peptides with high affinity, and a conserved SOCS box located at their C-terminus. The latter domain can form, with other proteins, the E3 ubiquitin ligase complex, enabling the bound proteins (JAKs, STATs and GHRs) to be directed for ubiquitin-mediated proteasomal degradation (Linossi et al. 2013; Zhang et al. 1999). SOCS2 and CIS can also exert actions via their SH2 domains, through competition with STAT5 for the phosphotyrosines on the GHR ICD, thus blocking STAT5 phosphorylation. SOCS1 and 3 have very different effects: they contain a kinase inhibitory region (KIR) at their N-termini which enables them to directly bind the JAK catalytic domain, blocking its kinase activity and ability to phosphorylate its substrates (Greenhalgh et al. 2005; Kershaw et al. 2013; Sasaki et al. 1999; Wojcik et al. 2018). The key role of SOCS2 in the regulation of GHR signaling has been shown by the generation of knockout mice: only SOCS2-deficient mice present with an ~40% increase in body growth compared to wild type mice, mimicking the acromegalic/gigantism phenotype in humans (Linossi et al. 2013; Metcalf et al. 2000). Recently, its role in controlling GHR signal duration has been shown to be crucial: studies of a SNP in the *GHR* gene (found to be associated with lung cancer) showed that the resultant amino acid change (P495T) in the GHR ICD caused a structural change in a SOCS2 binding site, impairing GHR degradation (Chhabra et al. 2018).

***(ii) Protein Tyrosine Phosphatases (PTPs)*:** Four PTPs have been shown to down-regulate GH signaling by exclusively dephosphorylating phosphotyrosine residues: SHP1, SHP2, PTP-1B and PTP-H1 (Dehkhoda et al. 2018). PTP-1B and PTP-H1 have been shown to dephosphorylate GHR activated by GH (Pasquali et al. 2003). Mice lacking PTP-H1 show increased weight compared to wild type as well as higher levels of *IGF-1* mRNA in liver and IGF-1 in serum (Pilecka et al. 2007). Mice lacking SHP1 and 2 are embryonic lethal but gain

of function mutations in the PTPN11 gene (coding for SHP2) result in an excessive activity of SHP2 in 50% of Noonan syndrome cases; these individuals are characterized by dysmorphic facial features, heart disease and proportional short stature (Tartaglia et al. 2001).

(iii) Protein Inihibitor of Activated STATs (PIAS): Although these proteins (PIAS1-4) are known as negative regulators of JAK-STAT signaling, their precise role in GH signaling is still unclear. Upon binding to the STATs, they can act through several different mechanisms to inhibit their activity: preventing STAT dimerization, blocking their DNA binding domain, modulating the localisation of co-regulators, and STAT sumoylation (Linossi et al. 2013; Wojcik et al. 2018).

C. Growth Hormone Receptor isoforms: identification and roles

a) The Growth Hormone Binding Protein (GHBP)

In several species, a substantial fraction of circulating GH is carried by a high affinity GHBP, the structure of which corresponds to the GHR ECD (Baumann 1994). As the GHBP retains its ability to bind specifically its ligand, it can compete with membrane receptors, thus potentially acting as an antagonist. In addition, soluble cytokine receptors have been thought to act as carrier proteins that increase the half-life of their respective cytokine by potentially decreasing its metabolic clearance rate (Baumann 2001; Rose-John and Heinrich 1994). In rodents, the GHBP is largely derived by translation of an alternatively spliced *GHR* mRNA encoding only the receptor ECD and a hydrophilic tail allowing the protein to be secreted instead of being retained at the plasma membrane (Baumbach et al. 1989). In humans and rabbits, a member of the metalloprotease family (TACE: TNF-α converting enzyme also called ADAM-17) cleaves the GHR ECD, releasing the GHBP into the extracellular space, a process called “shedding” (Baumann 2001); thus, the GHBP has been used as an indirect measure for GHR expression at the cell surface. Low levels of GHBP have been associated with GH resistance in conditions such as malnutrition, uncontrolled diabetes, catabolic states, renal failure and hypothyroidism (Baumann 2001). In the majority of the GHI patients, it is a reliable clinical marker: for example, patients with Laron dwarfism have very low to undetectable levels of GHBP (Schilbach and Bidlingmaier 2015). However, in certain conditions, this relationship is not linear and some GHI patients present with normal to elevated GHBP levels (Amit et al. 2000). A single mutation (D152H) in the extracellular domain abolishing receptor homodimerization can result in both normal GH binding and

GHBP levels while mutations causing alternative splicing of exon 8 or 9 result in the generation of a truncated GHR and lead to abnormal elevated GHBP levels (described below). Although these clinical observations are rare, they suggest that GHBP levels should not be an isolated measure for assessment of GHR status (Amit et al. 2000; Baumann 2001; Schilbach and Bidlingmaier 2015).

b) Truncated GHRs (GHRtr): GHR1-279 and GHR1-277

Two short GHR isoforms were first identified in human liver using RT-PCR: they showed partial or complete skipping of exon 9 due to alternative splicing (Dastot et al. 1996; Ross et al. 1997). GHR1-279 utilizes an alternative 3'-acceptor cryptic splice site within exon 9, 26bp downstream of the splice acceptor used for full length GHR (GHRfl). The predicted C-terminal residues of GHR1-279 peptide are frame-shifted to end in an early stop codon within exon 9. GHR1-277 skips all of exon 9 with exon 8 splicing directly to exon 10, resulting again in a frame shift and an early stop codon in exon 10. As a result, these truncated forms are still found in the cell membrane but lack >97% of their intracellular domains, including Box 1 and Box 2 that are critical for GHR JAK2 association, and, thus, are unable to transduce GH-JAK2 mediated actions. GHR1-279 and GHR1-277 are both naturally produced isoforms of GHR but make up only a small proportion of the total GHR: GHR1-279 represents 1-10% and GHR1-277 less than 1%. Expression of mRNA for the three receptors, GHRfl, GHR1-279 and GHR1-277, is highest in the major GH target tissues (liver, fat, muscle and kidney) and the pattern of isoform expression varies widely among tissues, implicating a tissue-specific putative role for those isoforms. It was shown that GHR1-279 levels in fetal liver were comparable to those in adult liver, whereas GHRfl and GHR1-277 were only half the levels found in adult liver, suggesting that the exon 9 alternative splicing may also be regulated developmentally (Ballesteros et al. 2000b).

c) GHR truncated forms act as dominant negative regulators

***(i) In vitro studies*:** Immunoprecipitation and western blot experiments of cells co-transfected with GHR1-279 and/or GHRfl revealed that they could form heterodimers (Ayling et al. 1997). Cells co-transfected with GHRfl and increasing amounts of GHR1-279 showed a dose-dependent inhibition of GH activation of a STAT5 reporter vector (Ross et al. 1997). A 10:1 ratio of GHRfl to GHR1-279, similar to that observed *in vivo* in certain tissues, led to an

inhibition of GH actions by up to 30%. A 1: 1 ratio decreased GH signal induction by 50% while a 1:10 ratio completely blocked activity of the GHRfl.

***(ii) In vivo studies*:** The importance of these truncated forms has also been shown *in vivo*. Two research groups have described patients with severe short stature who have a heterozygous mutation at the donor splice site of intron 9 of the *GHR* gene resulting in the complete skipping of exon 9 from one allele and the production of GHR1-277 (Ayling et al. 1997; Iida et al. 1998). As the mutations of *GHR* in these cases are heterozygous, three different types of GHR dimerization can occur: homodimers of two GHRfl, heterodimers of GHRfl and GHR-277 and homodimers of GHR-277. Homodimers of GHRfl can transduce the GH-JAK2 signal. However, as GHR-277 lacks the Box1 motif, the heterodimers of GHR-277 and GHRfl and the homodimers of GHR1-277 cannot.

***(iii) GHR truncated forms show abnormal internalization*:** GHRfl has a half-life of around 1h and is continuously degraded even in the absence of GH (Gorin and Goodman 1985). Both liganded and unoccupied GHRs are endocytosed *via* clathrin-coated vesicles and subsequently transported via endosomes to lysosomes (Strous et al. 1996). Both endocytosis and transport to lysosomes require an active ubiquitin conjugation system and a 10-amino acid UbE-motif inside the conserved Box 2 region of the GHR ICD (van Kerkhof et al. 2001). Truncated forms of GHR lack the essential internalization motif in the cytoplasmic domain and, therefore, accumulate at the cell surface, available to heterodimerize with GHRfl and to compete with the full-length receptor for GH binding, enhancing their dominant negative effect.

There are two consequences of an impaired internalization. First, the truncated receptor always demonstrates a greater level of receptor expression at the cell surface compared with the full length receptor (Ross et al. 1997). Second, as the GHRtr isoform is sustained at the cell surface, it becomes more susceptible to proteolytic cleavage and, thus, generates high levels of soluble GHBP. The latter has been observed in patients with heterozygous mutations generating GHR1-277 (Iida et al. 1999; Ross et al. 1997)**.**

d) The exon 3 deleted (3-) Growth Hormone Receptor isoform

Several groups have shown that both exon 3+ and exon 3- hGHR isoforms exist in the normal population (Wickelgren et al. 1995; Zogopoulos et al. 1996b). Although it was initially thought that the 3- hGHR isoform was caused by individual-specific alternative splicing

mechanisms, Pantel et al. showed that it is more likely caused by a homologous recombination event occurring between retroviral elements that surround exon 3 in the human *GHR* gene (Pantel et al. 2000). The production of both full-length and 3- isoforms is a species-specific event: other mammalian species (e.g. rabbits, rodents) express only the full length GHR, whereas birds, bony fish, marsupials and amphibians express only a 3- GHR form. Interestingly, the human PRL receptor, closely related to the human GHR, always lacks a corresponding exon 3 (Pantel et al. 2000).

The lack of exon 3 results in a 22aa truncation of the extracellular domain near the N-terminus of the mature receptor, a loss of one glycosylation site and the substitution of a highly conserved aa leading to a change in charge, size and hydrophobicity of the receptor domain. The GHR global folding of the ECD is supposed not to be altered but the crystal structure of the 3- GHR isoform has not been modeled to date (Brooks and Waters 2010a; Dos Santos et al. 2004). Therefore, the functional consequences of this N-terminal shortening are still unknown.

The prevalence of the 3- allele in humans is quite high (approximately 25-30%) with a homozygous frequency of 9-15% (Kenth et al. 2007). What is not clear is whether the 3-GHR form has a special physiological role in regulating responsiveness to GH. There have been controversies as to whether the 3- polymorphism is associated with a better growth response to GH in short stature children. Some groups have reported significant differences in height velocity between those with the 3+/3+ genotype *vs.* those with at least one 3- allele, including patients who have been diagnosed as small for gestational age (SGA), GH deficient, ISS or Turner syndrome (Binder et al. 2006; Dos Santos et al. 2004; Jorge et al. 2006; Wassenaar et al. 2009). In contrast, other groups report no significant association between the genotypes and response to GH therapy (Blum et al. 2006; Carrascosa et al. 2006; Pilotta et al. 2006). Recent reports have concluded that the better growth response to GH in patients with at least one 3- allele is restricted only to the first 1-2 years of treatment and that it does not alter final adult height (Dorr et al. 2011; Tauber et al. 2007).

The exon 3 deletion genotype is also reported to be associated with a lower Body Mass BMI and significantly improved glucose tolerance in patients with acromegaly (Montefusco et al. 2010). The mechanism for increased hormone sensitivity is still unclear. One *in vitro* transfection experiment has shown that the transduction of GH signaling through 3- GHR

homo- or heterodimers was 30% higher than through full length GHR homodimers (Dos Santos et al. 2004).

2. The *GHR* gene

A. *GHR* Gene Organization: a complex promoter

a) The human *GHR* gene

The human *GHR* gene spans ~300kb on the short arm of chromosome 5 close to the centromere (Barton et al. 1989a; Godowski et al. 1989; Leung et al. 1987). The coding region is defined by exons 2-10 where exon 2 contains the translation start site (Leung et al. 1987). Using RT-PCR/Southern-Blot and immunohistochemical approaches, *GHR* mRNA and protein have been detected in human tissues as early as the 9th week of fetal life. Levels increase gradually during gestation in a tissue-specific manner and, by mid-gestation, there is a tissue distribution similar to that found in the adult. These changes in tissue *GHR* expression, both pre- and post-natally, are due to developmental, hormonal, nutritional and pathophysiological signals (Goodyer et al. 2001a; Hill et al. 1992; Simard et al. 1996; Wei et al. 2006; Zogopoulos et al. 1996a).

b) The human *GHR* 5' untranslated region (5'UTR), first exons and mRNA variants

In 1992, Pekhletsky et al cloned the first eight *GHR* mRNA variants from adult human liver tissue and named them V1 to V8 according to their relative abundance. Almost ten years later, the V9 transcript was identified in the adult human heart with similar levels as V2 and V3 and its first exon was mapped within the *GHR* 5' untranslated region (5'UTR) (Goodyer et al. 2001c). To date, fourteen different *GHR* mRNAs encoding the full-length *GHR* have been reported (Goodyer et al. 2001c; Orlovskii et al. 2004; Pekhletsky et al. 1992; Wei et al. 2006). They each have a unique 5'UTR, derived from different first exons, but all splice into the same site in exon 2, 11bp upstream from the ATG translation start site and, thus, code for the same protein (**Figure I-7**). The definition of the first exons through 5'RACE and chromosomal mapping experiments showed that, while the majority are expressing *GHR* mRNA ubiquitously, four show a similar tissue- and developmental-specific expression pattern.

Seven of the variant exons form two distinct clusters (Goodyer et al. 2001c). One distal cluster (Module A), located ~140kb from exon 2, contains three of the exons (V2, V3, V9) within a 1.6kb region; they transcribe ubiquitous *GHR* mRNAs from as early as the third month of fetal life in all tissues examined. A proximal cluster (Module B), located ~18kb upstream of exon 2 within a 2kb domain, contains four exons (V1, V4, V7, V8); their mRNAs are expressed only in normal hepatocytes and only beginning ~3-4 months after birth (Goodyer et al. 2001a; Wei et al. 2006; Zogopoulos et al. 1996a). The factors responsible for the "switch" between the absence of fetal liver-specific *GHR* mRNAs and their postnatal presence are not fully understood. However, they likely explain the dramatic increase in total *GHR* mRNA (4 to 6 fold) and protein (4 fold increase in ^{125}I hGH binding) levels observed in the postnatal liver (Goodyer et al. 2001a; Kenth et al. 2011). In contrast, a postnatal tissue-specific decrease in *GHR* mRNAs is observed in the lung, kidney and small intestine. The highest levels postnatally are found in liver, kidney and adipose tissues (Goodyer et al. 2001a; Kenth et al. 2011).

The other exons expressing ubiquitous *GHR* mRNAs include V_A-V_D and V3a/b/$_E$, which are located between Modules A and B, while V5 is adjacent to exon 2. V3 mRNA subvariants are the product of alternative splicing of three separate exons in the *GHR* 5'UTR and show high homology with Alu repeat elements (Goodyer et al. 2001b). V9 can also be alternatively spliced (RB Wickelgren et al. unpublished data; UCSC Genome Browser AF 230801 V9b). V6 was shown to be a 5'RACE artefact (Orlovskii et al. 2004).

c) *GHR* 5'UTRs in alternative animal species

Heterogeneity of the 5'UTR of the *GHR* gene is a common feature across different animal species. Two major promoters and mRNA expression patterns for the *GHR* gene have been identified in the human, mouse, rat, ovine and bovine: one *GHR* mRNA is always specific for postnatal liver while there is also at least one ubiquitously expressed mRNA (Goodyer et al. 2001c; Heap et al. 1995; Menon et al. 2001; Moffat et al. 2000; Moffat et al. 1999; O'Mahoney et al. 1994; Schwartzbauer and Menon 1998). The human "V2-like" exons are L2 in the mouse, GHR2/V1 in the rat and 1B in the ovine and bovine. V9 and V3 homologues in other species have also been identified in the mouse (L3-L5) and bovine (1C) (Jiang et al. 1999; Moffat et al. 2000). The human hepatic specific "V1-like" exons are L1 in the mouse, GHR1/V2 in the rat and 1A in the ovine and bovine. Although alternative variants in the mouse (L3-L5) are located in the same L2 exon-containing cluster, their very low abundance in different tissues makes it impractical to study their regulation (Moffat et al. 2000). Despite

certain structural similarities in the *GHR* 5' flanking region amongst various species, no readily accessible animal model appears to be appropriate for investigating the complex regulation of the human *GHR* (Goodyer et al. 2008; Kenth et al. 2011; Wei et al. 2009a; Wei et al. 2006).

B) Distinct regulatory mechanisms within the *GHR* 5'UTR

For the last decade, efforts have been made to delineate the physiological and cell-specific factors that regulate *GHR* expression through interactions at the proximal regulatory domains of Module A and B exons. The promoters of exons associated with ubiquitous *GHR* mRNAs are important for influencing GH responsiveness in every cell in the body and, therefore, must respond to numerous metabolic, hormonal and intracellular signals. Thus, any changes in these regions (e.g. polymorphisms) could have more global effects. In contrast, the promoters of exons linked to postnatal liver-specific *GHR* mRNAs are likely to be responsible for translating postnatal hepatocyte-specific cues and are repressed during fetal liver development as well as in all other tissues. To date, several members of our lab have conducted extensive studies to characterize the different regulatory elements, using luciferase reporter vectors with serial deletion constructs of the different promoters, site-directed mutagenesis, transient transfections/co-transfections, EMS(S)A and ChIP assays. They have demonstrated the functionality of putative response elements and shown that the *GHR* Module A vs. B exons are regulated in unique ways (Erman et al. 2011b; Goodyer et al. 2008; Kenth et al. 2011; Wei et al. 2009a).

(a) Molecular mechanisms regulating ubiquitous *GHR* expression by Module A exons

***(i) Common promoter regulation of Module A exons*:** The three Module A exon promoters have two common regulatory elements, Sp1/Sp3 binding sites and CpG islands, both of which are present in many ubiquitously expressed genes (G Kenth et al., unpublished data). Interestingly, these elements are not present in Module B or the promoter regions of the other ubiquitously expressing exons (V5, V_A-V_D), suggesting that these represent Module A-specific regulatory mechanisms.

Within the Sp family of zinc finger transcription factors, Sp1 and Sp3 are ubiquitously expressed in mammalian cells and regulate the expression of a vast number of genes implicated in almost all cellular processes (Li et al. 2004). They bind to GC-rich boxes that are commonly found in promoters with CpG islands to regulate gene expression; TATA-less promoters have been shown to be particularly regulated by Sp proteins. Like V2 in humans, the mouse L2 promoter is GC-

rich and TATA-less. Interestingly, L2 transcription can be activated differentially by Sp1 (weak activator) and Sp3 (strong activator) depending on their ratio during fetal and postnatal development that could explain increases in postnatal GHR expression in certain murine tissues (Yu et al. 1999). In addition, Sp1 and Sp3 cooperate with other factors but only Sp1 can synergistically activate promoters by forming higher order complexes (e.g. by recruiting chromatin remodelers such as the SWI/SNF family proteins or HDAC), showing the complexity of their actions (Li et al. 2004).

***(ii) Specific regulation of V2 mRNA transcription*:** The three exons of Module A are not equally active in human tissues: V2 is the most highly expressed transcript in all tissues examined, especially adipocytes (Goodyer et al. 2001c; Wei et al. 2006). Wei et al have determined that CHOP, C/EBPs and Ets1 significantly stimulate V2 expression through sites in its proximal promoter (Wei et al. 2009b). In addition, they found an important role for ~160bp of the V2 exon sequence immediately downstream of the start site in regulating basal, CHOP-stimulated and Hes1-inhibited transcriptional activity. Finally, by using qRT-PCR assays, the expression levels of these factors in adipocytes were quantified during differentiation, showing a correlation of promoter findings with adipocyte biology (Wei et al. 2009a).

***(iii) GH regulates GHR expression via GAGA boxes*:** GH is primarily known to enhance gene expression through STAT5 signaling (Lanning and Carter-Su 2006; Woelfle and Rotwein 2004). However, lab members could not detect a STAT5 response element in proximal promoter regions of the Module A or B exons, either by *in silico* analyses or by overexpressing STAT5B in transient transfection assays (S Puzhko et al, unpublished data). Instead, multiple GAGA boxes, that have been found to function as GH response elements (GHREs) in several mammalian genes (Legraverend et al. 1996; Volpi et al. 2002; Wyse et al. 2000), are present: three GAGA elements are located in the V3 promoter and exon. Because there is no known mammalian GAGA binding factor, Kenth et al used *Drosophila* GAGA binding factor (GAF-519) to show that V3 promoter activity could be stimulated through the GAGA boxes. Nuclear extracts from cells treated with GH (but not insulin or IGF-1) resulted in a new EMSA complex forming with the GAGA element probe, suggesting that the V3 GAGA boxes do indeed function as a GHRE (Kenth et al. 2011)**.**

***(iv) Mechanisms of circadian regulation of GHR expression*:** Studies have shown that GHR has a diurnal expression pattern in murine liver, bone and skeletal muscle (Itoh et al. 2004; Zvonic et al. 2007). A member of our lab found that two transcription factors implicated in

the regulation of several diurnally controlled genes, DBP (D-binding protein) and its repressor E_4BP_4, bind to and significantly modulate V9 (as well as V1) promoter activity, suggesting that these factors may be important for circadian regulation of human GHR expression (G Kenth, unpublished data).

***(v) Regulation of GHR expression by obesity-related factors*:** In 2011, Erman et al used ChIP assays, site-directed mutagenesis and luciferase reporter constructs to characterize functional response elements for HIF1-α, NFκB (downstream TNF-α signaling) and glucocorticoids in V9 and V3 exons. The findings showed differential effects of these factors on *GHR* transcription regulation in adipocytes and HEK293 cells, suggesting mechanisms for responses to hypoxic or inflammatory environments (Erman et al. 2011b).

b) Regulation of normal postnatal liver-specific *GHR* expression by Module B exons

***(i) Common regulation of Module B exons*:** V1 is the most highly expressed *GHR* mRNA in postnatal liver, likely because the promoter of exon V1 has two active TATA boxes, while the other three exons have non-consensus TATA elements within their promoters (Goodyer et al. 2008; Goodyer et al. 2001c)**.** Interestingly, the Module B promoter constructs are actively repressed in all cell lines tested. These data suggest there is coordinated inhibitory control of V1, V4, V7 and V8 that is likely to be important in all non-hepatic tissues as well as fetal liver and hepatic tumors, since none of these tissues express Module B-derived *GHR* mRNAs (Goodyer et al. 2001a; Goodyer et al. 2001c; Wei et al. 2006; Zogopoulos et al. 1996a).

***(ii) Regulation of V1*:** Adjacent GAGA and growth factor independence-1/1b (Gfi-1/1b) response elements are present within the V1 exon region between the two TATA boxes. GAF-519 stimulates V1 promoter activity while Gfi-1/1b is a strong repressor through their respective sites, suggesting that liver-specific V1 transcription is tightly controlled by these two elements. When co-transfected, Gfi-1/1b completely inhibited GAF activity, indicating that Gfi-1/1b may be responsible for the lack of V1 expression in fetal liver, non-hepatic tissues and hepatic tumors (Kenth et al. 2011).

***(iii) Regulation of V1-derived GHR mRNAs by metabolic signals*:** Several putative binding sites for liver-enriched transcription factors are present in Module B exon promoters, including multiple HNF4 sites upstream of V1. Members of our lab showed that HNF-4α regulates V1 *GHR* expression in the human hepatocyte in response to fatty acid metabolic cues (Goodyer et al. 2008).

c) Potential epigenetic mechanisms regulating *GHR* expression

***(i) Chromatin organization and histone modifications*:** In eukaryotic cells, the chromatin is formed by DNA associated with proteins, mostly histones, organized in units called nucleosomes. The predominant form of chromatin in the nucleus is the 'beads on a string' 10nm fiber form seen in electron microscopy; the highest degree of chromatin condensation is the chromosome. The spatial organization of the chromatin has been linked to transcriptional regulation and led to the definition of two types of chromatin: the transcriptionally active euchromatin (localized in distinct regions of the nucleus, less condensed, depleted in nucleosomes as well as sensitive to DNAse I digestion, more accessible to transcription factors) and the heterochromatin (highly condensed, nucleosome dense and resistant to DNAse digestion, inaccessible to transcription factors). A nucleosome is formed by an octamer of four types of core histones (H2A, H2, H3, H4) with 146 bp of DNA wrapped around it. Histone N-terminal tails are prone to posttranslational modifications, such as acetylation, methylation, phosphorylation or ubiquitylation, and have been shown to affect chromatin organization, impacting all aspects of transcriptional regulation. For example, H3K4me3 or H3K9ac are found at actively transcribed genes with H3K27me3 and H3K9m3 at silenced genes. The functional consequences of these histone marks are mediated by chromatin remodeling enzyme complexes (e.g. SWI/SNF family) that can affect chromatin in different ways, including nucleosome sliding, eviction or histone variant exchange. Defects in these complexes have been linked to human pathologies (Tyagi et al. 2016).

***(ii) Methylation: GHR Module A region is GC rich and contains two CpG islands*:** DNA methylation is the most abundant epigenetic modification. It involves the placement of a methyl group on carbon 5 of cytosine residues (5-methylcytosine) by DNA methyltransferases (DNMT1, DNMT3a and b). This occurs primarily at CpG residues at the 5' of the guanine without affecting base pairing and can alter gene regulation as it constitutes a repressive mark of transcription by recruiting methyl binding proteins (MBPs) and by promoting chromatin remodeling, silencing gene expression (Bernstein et al. 2007). Genomic regions particularly enriched in CpG residues form CpG islands (~5-10 CpGs per 100 bp) and are frequently found in promoters. 70% of annotated genes contain CpG islands in their promoter which is a common feature of housekeeping genes and genes involved in development while the majority of the tissue-specific gene promoters lack CpG islands and TATA boxes (Deaton and Bird 2011; Zhu et al. 2008). The levels of methylation have been inversely correlated with the density of CpGs: in the genome, the sparse CpGs are highly

methylated while CpG islands located within promoters are mostly maintained in a hypomethylated state. These latter regions are associated with transcriptional activity as they co-localize with RNA polymerase II and the recruitment of ubiquitous transcription factors like Sp1 which has been shown to recruit TATA binding proteins (TBPs) to initiate transcription in TATA-less promoter (Deaton and Bird 2011). This transcriptionally permissive chromatin state is characterized by a key chromatin signature, the histone mark H3K4me3 (for active chromatin state), that is still present even when the gene is inactive. The influence on local chromatin organization is done through recruitment of CXXC finger protein 1 (Cfp1) that specifically associates with non-methylated CpG sites and with the H3K4 methyltransferase Setd1 (Thomson et al. 2010).

Differential methylation levels at CpG islands have been shown to be tissue specific, especially in genes involved in developmental processes, and is frequent in certain cancers in which tumor suppressor genes show aberrant promoter hypermethylation (Illingworth and Bird 2009; Schilling and Rehli 2007). Methylation at CpGs is critical during embryonic development and is dynamic, with usually symmetrical methylation between both alleles. However, allelic asymmetry or allele-specific methylation occurs during X chromosome inactivation in females, leading to monoallelically expressed genes, as well as in imprinted genes with parent of origin (~100 genes in the human genome). Allele-specific methylation has been shown to be frequent in the human genome and associated with cis regulatory SNPs (meQTLs) (Tycko 2010). As mentioned earlier, *GHR* Module A is GC rich and contains CpG islands; their potential regulatory mechanisms have not yet been explored.

***(iii) MicroRNAs regulate GHR expression*:** MiRNA*s* are a class of small (19-22 nucleotides) endogenous RNAs that are important regulators of gene expression, mostly through their binding to the 3'UTRs of specific genes and targeting them for degradation or translation inhibition (Friedman et al. 2009). The miRNAs are first transcribed as a primary miRNA (pri-miRNAs) and then cleaved to a precursor pre-miRNA of 60-70nt with a hairpin structure by the Drosha enzyme (RNA class III enzyme). In the cytoplasm, the pre-miRNA is cleaved by Dicer, another RNA III enzyme, and matured into miRNA that will target genes in the form of a ribonucleoprotein complex known as miRISCs (miRNA-induced silencing complex). The miRNA field has expanded exponentially since their discovery, as they regulate a wide range of biological processes and have been implicated in many disease etiologies. In 2014, Elzein and Goodyer showed for the first time that miRNAs are potentially important regulators of human *GHR* gene expression. Using co-transfection assays with luciferase reporter *GHR*-

3'UTR constructs and miR mimics, followed by validation on endogenous *GHR* mRNA expression, they reported significant inhibitory effects of four miRNAs on both *GHR* mRNA and protein expression in HEK293 and breast as well as prostate cancer cell lines (Elzein and Goodyer 2014).

While all of the epigenetic regulatory processes discussed above can potentially regulate *GHR* gene transcription, to date, only the one miRNA study has been reported.

C) Repeated elements and potential functional relevance for *GHR* expression

It has been known for decades that only a small fraction of eukaryotic genomes code for proteins and that the genome is replete with non-coding and repetitive DNA (Britten and Kohne 1968). Interspersed repeats are the predominant type of repeat and are derived from the activity of transposable elements (sequences that can replicate and move within the genome); they represent almost half of the human genome (Lander et al. 2001). These elements are an important source of mutations and are recognized as a driving force in vertebrate genome evolution. They influence the genomic structure *via* sequence-mediated chromosomal rearrangements and can affect gene expression regulation and transcriptional regulatory networks by insertion mechanisms (Bourque 2009; Feschotte 2008).

Interspersed repeat elements are classified according to their replicative strategy which involves RNA (class 1 or retrotransposon) or DNA (class 2 or DNA transposon) intermediates. Class 1 repeats are divided into two subclasses, the long terminal repeat (LTR) retrotransposons, which are inserted using retroviral-like integrase, or the non-LTR retrotransposons, which include long and short interspersed elements (LINEs and SINEs) and use target-primed reverse transcription (Feschotte et al. 2009).

a) Class I interspersed repeat elements in *GHR*

Alu elements are primate-specific repeats and comprise 11% of the human genome (>1 million copies). They're the largest family of the ubiquitous class of SINE elements and use trans-factors from the LINE-1 retro-element for their amplification (Deininger 2011). They've been shown to have a strong impact on primate genome evolution through insertional mutagenesis and their enrichment in genes has linked them functionally to gene regulation (Deininger 2011). In the *GHR* gene, most of the Alu sequences are located in introns although a few have been found in exonic regions. One from the S_c subfamily is located in its 3'UTR in the antisense orientation and is flanked by direct repeats (Godowski et al. 1989). In

addition, alternatively spliced variants of V3 mRNA have been reported to contain *Alu*-like elements in their 5'UTR: the V3b exon contains *Alu* sequences from the S subfamily in the sense direction while the V3a/b *Alu* is from the J_o subfamily in the antisense orientation (Goodyer et al. 2001b). These *Alu* elements in *GHR* first exons have the potential to alter its transcriptional regulation, alternative splicing and/or translation (Goodyer et al. 2001b). As mentioned earlier, another class of repeated elements, LTRs of endogenous retroviruses, have been implicated in the evolution of the *GHR* gene and explain the presence of two different *GHR* alleles either including exon 3 or not (Pantel et al. 2003; Pantel et al. 2000).

b) Short Tandem Repeats (STRs) in *GHR*

Another common form of repetitive DNA sequences are the short tandem repeats (STRs) or microsatellites which are tandem repeats of short (1-6bp) DNA motifs (Bagshaw 2017). Mini-satellites are also tandem repeats but with a repeated motif of >6-10bp and <100bp while extremely long repeated motifs (>100bp) are called DNA satellites or macrosatellites (Bhargava and Fuentes 2010; Ellegren 2004). Microsatellites are ubiquitous in both eukaryotic and prokaryotic genomes and represent ~1-3% of the human genome (Bhargava and Fuentes 2010; Gymrek et al. 2016; Lander et al. 2001); their density differs amongst species and tends to correlate with genome size, with a higher density in mammals (Ellegren 2004; Toth et al. 2000). Among the microsatellites, the dinucleotide repeats are the most common with the $(GT/CA)_n$ the most frequent followed by $(AT)_n$, $(GA)_n$ and $(GC)_n$, the last being rare (Ellegren 2004).

The main characteristic of the microsatellites is their polymorphic nature, with high levels of heterozygosity attributed to strand-slippage during replication by the DNA polymerase; this results from transient dissociation of the replicating DNA strands followed by misaligned re-association (Ellegren 2004). They are highly mutable, up to 10 orders of magnitude greater than point mutations, and their mutation rate generally increases with repeat number and purity, with slippage contributing to their expansion and contraction (Bhargava and Fuentes 2010; Ellegren 2004; Gemayel et al. 2012). Because of their polymorphic nature they've been widely used as markers for genetic mapping, for studying genomic instability in cancer, parentage and forensic analysis, molecular anthropology and population genetics (Bhargava and Fuentes 2010). Expansion in tandem repeat size has been associated with more than 40 diseases, mostly neurodegenerative and neuromuscular, including Huntington's disease and

the fragile-X syndrome. These diseases are due to tandem repeats with trinucleotides and are called trinucleotide expansion diseases (Mirkin 2007; Paulson 2018).

Since their discovery in the early 1980s in the β-globin gene, these " junk DNAs" have been viewed as non-functional neutral markers with no phenotypic consequences (Orgel and Crick 1980), as they were found primarily in non-coding regions. Later on, however, they were shown to be non-randomly distributed in the human genome, enriched in regulatory genes encoding transcription factors, DNA-RNA binding proteins and chromatin modifiers, and associated with specific properties depending on their location (Katti et al. 2001; Legendre et al. 2007). As they represent a high genetic variability due to their multi-allelic nature, they potentially provide more information than the bi-allelic SNPs and have been viewed as fine-tuning regulators in different cellular processes (Bagshaw 2017; Gemayel et al. 2012). Recently, more large-scale analyses of microsatellites in the human genome have been undertaken to investigate their potential association with gene expression variation and complex traits. In 2016, Gymrek et al. evaluated the contribution of STRs in gene expression (eSTRs) in lymphoblastoid cell lines (LCLs) and showed that they explained 10-15% of the cis-heritability attributed to common polymorphisms, with an enrichment in clinically relevant phenotypes (Gymrek et al. 2016).

GHR contains 48 microsatellites scattered throughout the gene. In 2001, Hadjiyannakis et al. reported that a GT dinucleotide repeat located 81bp upstream of the transcription start site of one of the main first exons of *GHR* (V9) is a polymorphic microsatellite (Hadjiyannakis et al. 2001). In 2004, Orlovskii et al. found a second polymorphic GT microsatellite in intron 2 ~1.8kb downstream of exon 2 (Orlovskii et al. 2004). As discussed below, these elements have the potential to be involved in *GHR* gene regulation.

c) Promoter microsatellites modulate gene expression

Analysis of the human genome has determined that a higher heterozygosity is found in the cis-regulatory sites of genes than in coding regions, due to an overrepresentation of polymorphic microsatellites, especially those with the GT/AC motif (Rockman and Wray 2002). In genome-wide studies, eSTRs have been shown to localize in conserved regions near transcription start sites and predicted enhancers, with strong enrichment of histone modifications (Gymrek et al. 2016). In human promoters, microsatellites near transcription start sites have been shown to be often highly conserved, with the distance to the transcription start site being a good predictor for its conservation score (Sawaya et al. 2012b). Their high

conservation with low repeat variance between individuals suggests that microsatellite length may be under stabilizing selection for cis-regulatory function and, interestingly, serve as a source of genetic variation (Rockman and Wray 2002; Sawaya et al. 2012a; Sawaya et al. 2012b).

Changes in the repeat length of promoter microsatellites can have significant effects on phenotypes by altering levels of gene expression. In humans, several studies have shown this link with GT/AC regulatory microsatellites (Agarwal et al. 2000; Itokawa et al. 2003b; Ng et al. 2009; Rife et al. 2009; Searle and Blackwell 1999; Shimajiri et al. 1999; Wang et al. 2005; Yamada et al. 2000). Many of these studies found an enrichment of the microsatellites in genes that regulate growth and development (Sawaya et al. 2012b). Several have reported their functional roles in modulating gene expression through different mechanisms: microsatellite repeat number and composition can alter chromatin remodeling and accessibility by transcription factors, spacing between potential regulatory elements, and chromatin organization and nucleosome positioning (Bagshaw 2017; Bayele et al. 2007; Chen et al. 2016a; Liu et al. 2001; Taka et al. 2013). On a larger scale, a mechanistic link between tandem repeat variation and local genome function has been suggested by the study of Quilez and colleagues: they showed that genes containing polymorphic tandem repeats in their promoters had a higher variance of gene expression and DNA methylation (eQTL/meQTL). They identified >100 polymorphic tandem repeats associated with expression/methylation of adjacent genes that were overlapping with transcription binding and DNAse hypersensitivity sites (Quilez et al. 2016).

d) Promoter microsatellites can take alternative secondary structures: the Z-DNA helix

Under specific physiological conditions, most of the repetitive sequences, including promoter microsatellites, have the potential to take alternative structures other than the canonical B-DNA form. These secondary conformations (e.g. hairpins, cruciforms, triplexes, quadruplexes and Z-DNA) have been linked to physiological and pathophysiological states as well as genomic instability (Wells 2007). In 1979, Wang and al. reported the first single-crystal X-ray structure of Z-DNA (Wang et al. 1979). The structural properties of this conformation are characterized by a rotation of every other base around the N-glycosydic bond so that there is an alternation of *syn*- and *anti*-conformations; this results in a zig-zagging arrangement of the sugar-phosphate backbone instead of the smooth coil observed in B-DNA (Rich and Zhang

2003). The most favorable sequence to take a Z-conformation is the rare CG-repeat while the others follow this order: d(TG)=d(CA)n>d(TA)n (Rich et al. 1984).

In vivo, Z-DNA forms mostly during transcription as a consequence of negative supercoiling accumulating behind the progressing RNA polymerase II and then decreases due to the relaxation by topoisomerase I, thus linking transient formation of Z-DNA and transcriptional activity (Rich and Zhang 2003; Wittig et al. 1989; Wittig et al. 1991; Wittig et al. 1992; Wolfl et al. 1995). Several genes have been shown to be modulated by a Z-DNA element in their promoter acting as an enhancer or a repressor: e.g. the rat *nucleolin* gene (Rothenburg et al. 2001), the human colony stimulating factor 1 (*CSF1*) gene (Liu et al. 2001), the macrophage immune response gene *SLC11A1* (Bayele et al. 2007), the desintegrin and metalloproteinase *ADAM-12* gene (Ray et al. 2011) and the human heme-oxygenase *HO-1* gene (Maruyama et al. 2013). In the case of the GT repeat located in the *CSF1* promoter, transcriptional activation of the gene was dependent on Z-DNA formation through the action of chromatin remodelers, emphasizing the role of Z-DNA in chromosomal remodeling (Liu et al. 2006).

Surveillance of the human genome for Z-DNA forming regions has been done primarily by *in silico* analyses with the use of different algorithms. For example, in 1983, Ho and colleagues developed Z-Hunt, the first algorithm that predicted a non-random distribution of Z-forming sequences at the 5' end of the human genes and highly concentrated near transcription start sites (Schroth et al. 1992). Recently, Shin et al. used a combination of two Z-DNA binding domains as a probe followed by ChIP sequencing to generate a map of Z-DNA forming regions in HeLa cells. They found an accumulation in promoter domains that correlated with an enrichment of RNA polymerase II and histone marks associated with actively transcribed genes (H3K4me3 and H3K9ac) (Shin et al. 2016).

RATIONALE AND OBJECTIVE OF THE RESEARCH

Previous studies have shown the complexity of the *GHR* gene 5' flanking region in terms of organization and regulation, with multiple *GHR* mRNAs derived from 13 unique first exons. Efforts have been made to understand the different mechanisms involved in the regulation of the ubiquitous *vs* tissue specific *GHR* variant mRNAs. Defects in this gene have implications as extreme as the Laron syndrome, when the receptor is not functional, to tumor progression when it is overexpressed. Milder phenotypes characterized by growth hormone insensitivity show variable degrees of short stature and obesity. The potential association of genetic variants with GHR activities and their possible implications in the etiology of complex traits such as idiopathic short stature and obesity have not been investigated.

My working hypothesis is that there is less *GHR* mRNA and protein at the surface of GH target cells in individuals with idiopathic short stature or obesity, resulting in GH insensitivity. My goal has been to investigate the genetic contribution of a set of SNPs and a GT polymorphic microsatellite in the *GHR* gene in these two complex traits by defining their association with GHR expression and, thus, GH responsiveness.

The results of my PhD work are presented in the three following chapters that delineate the different questions that were asked throughout the progress of my research: chapter II focused on whether there is a potential association of the different selected polymorphisms within regulatory regions of the *GHR* gene with the short stature phenotype; in chapter III, I investigated the effects of several adiposity indices in the association of *GHR* as well as other height-related genes with short stature; finally, in chapter IV, I examined the potential functional significance of one significant polymorphism, the GT microsatellite located in the promoter of a ubiquitously expressing *GHR* variant exon.

Table I-1 Main defects of the GH/IGF-1 axis causing GHI

Gene Affected	Disorder, Phenotype	Clinical features
GHR Extracellular mutations Transmembrane mutations Intracellular mutations	Laron syndrome Variable height deficit and midfacial hypoplasia	variable GHBP decrease ↑GH ↓IGF-1, IGFBP-3 and ALS
STAT5b	GHI with immunodeficiency Severe growth failure, midfacial hypoplasia, immunodeficiency	↑GH ↓IGF-1, IGFBP-3 and ALS
IGF-1 defects causing IGF-1 deficiency bio-inactive IGF-1	IGF-1 insensitivity SGA, microcephaly, deafness	↑GH and IGFBP-3 variable IGF-1
IGFALS	Mild height deficit	GH? ↓IGF-1, IGFBP-3 and ALS
IGF1-R	IGF-1 resistance SGA, microcephaly	↑ GH, IGF-1, and IGFBP-3

Adapted from (Savage et al. 2011; Wit et al. 2016)

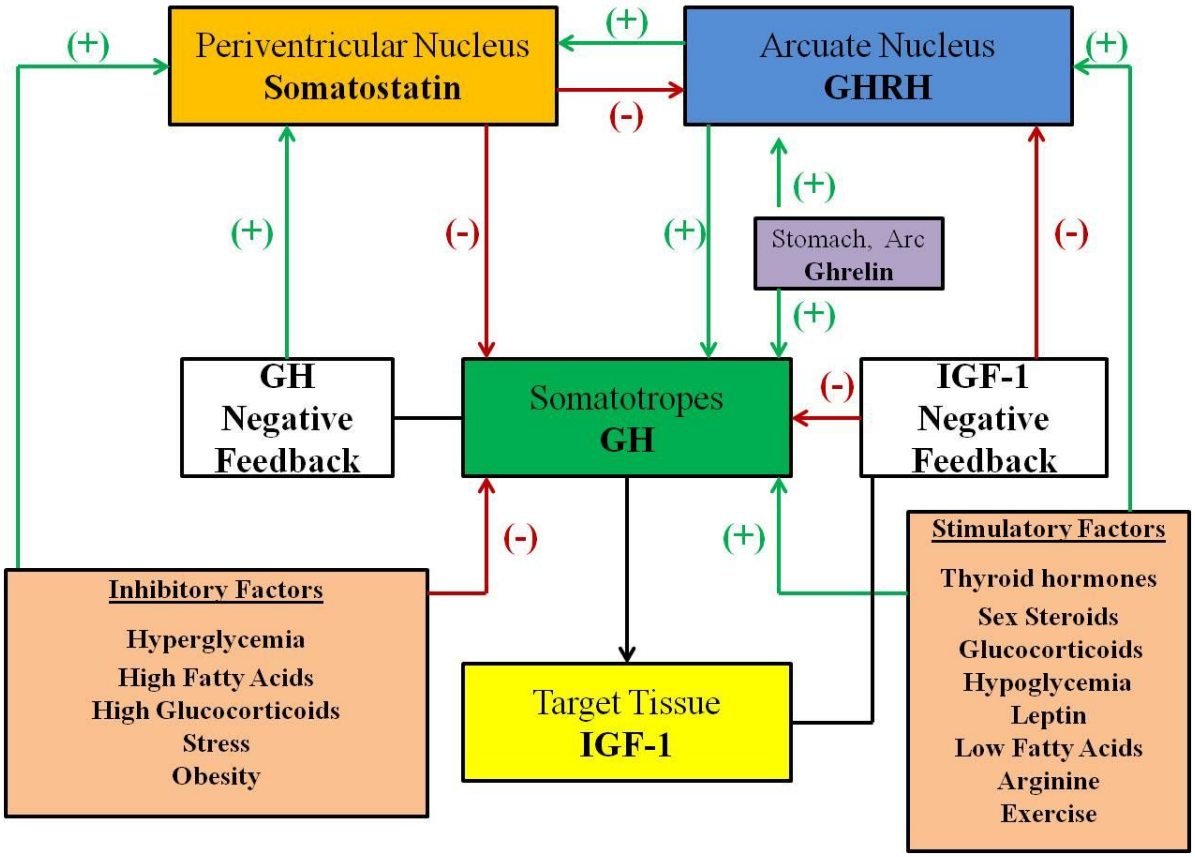


Figure I-1: Feedback regulation of GH release

.

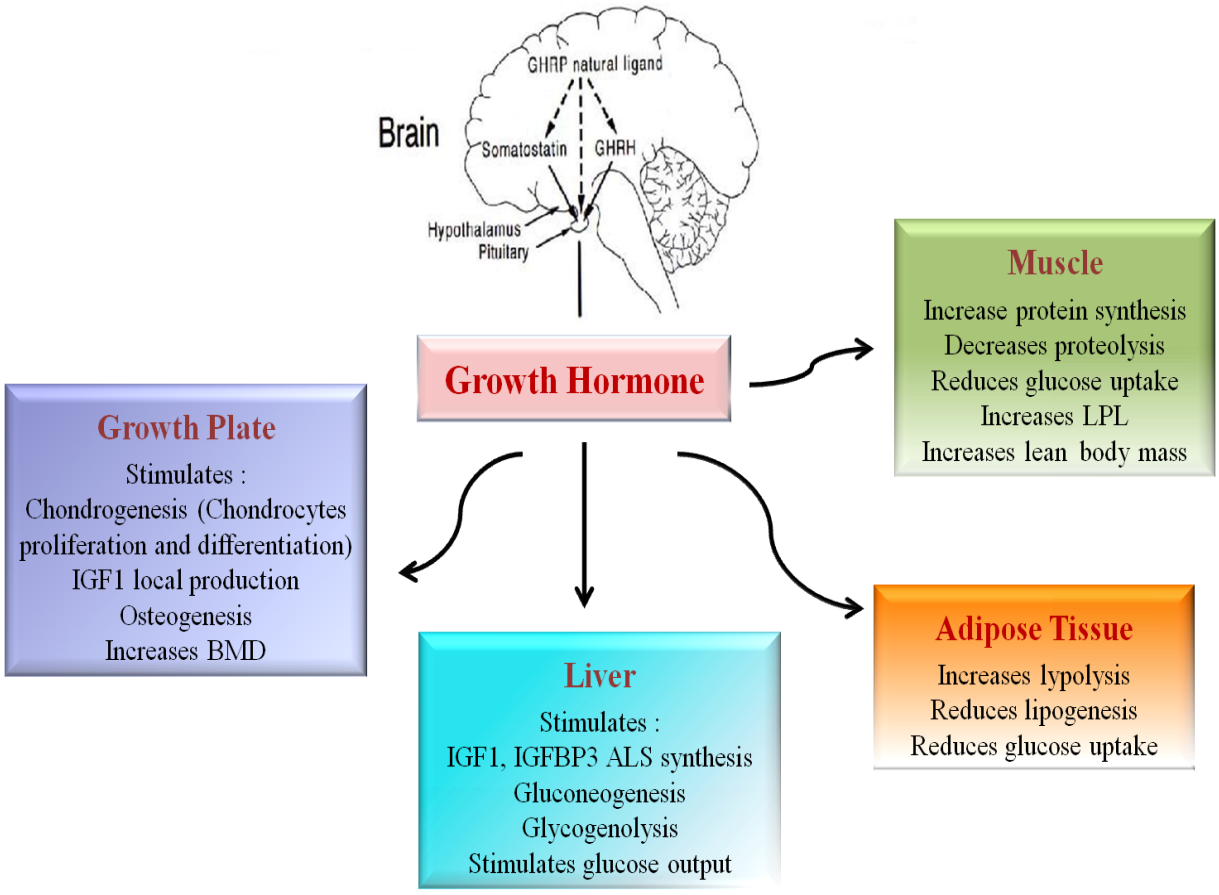


Figure I-2: GH main target organs and actions

Once released into the blood circulation, GH will bind with high affinity the GHR on target organs. In the liver, GH stimulates the production of IGF-1 as well as ALS proteins of the IGF-1 ternary complex. GH increases glucose levels in the context of hypoglycemia by increasing hepatic glucose production via gluconeogenesis and glycogenolysis; GH also regulates hepatic lipid metabolism *via* stimulation of triglyceride secretion (Fan et al. 2009). GH has a central role in the acquisition of longitudinal growth through stimulation of the chondrocytes, the cell unit of the growth plate, leading to their proliferation and maturation and also stimulates the local production of IGF-1. In the adipose tissue, GH exerts mostly direct effects *via* lipolytic and anti-lipogenic actions on adipocytes resulting in increased fatty acid concentrations in the blood. In skeletal muscle, both GH and IGF-1 promote growth, maintenance and regeneration of muscles through anabolic effects by increasing protein synthesis and decreasing the rate of protein breakdown (Chikani and Ho 2014).

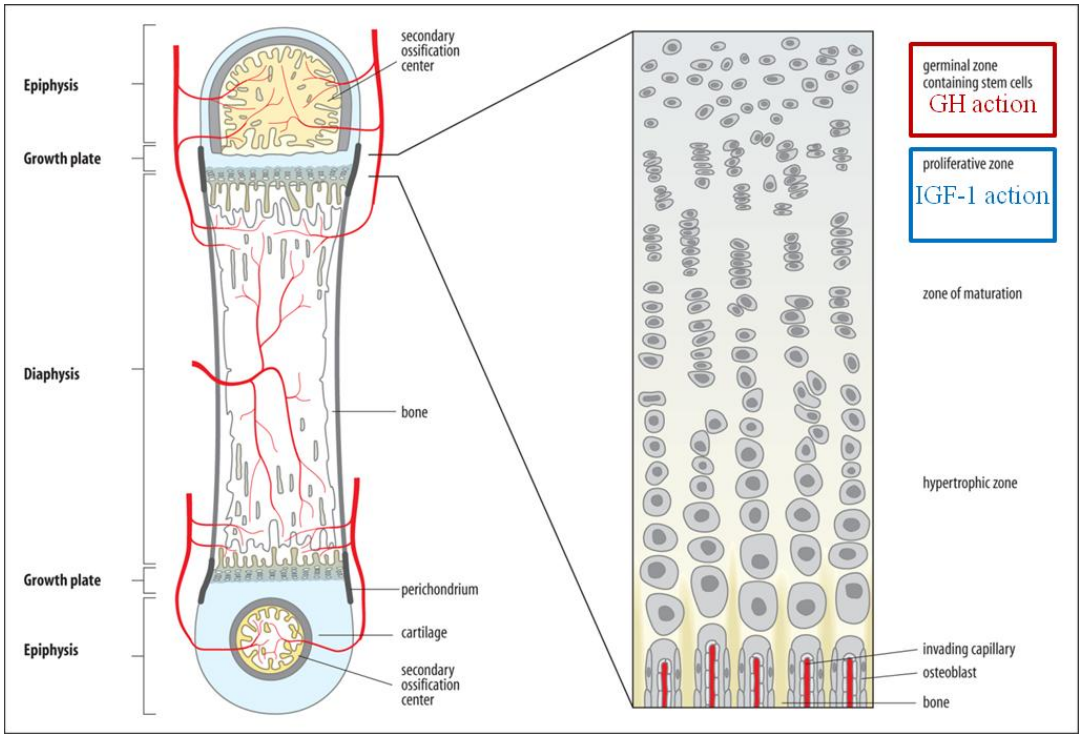


Figure I-3: Representation of endochondral ossification and GH/IGF-1 sites of action

Adapted from (Wolpert 2010)

Endochondral ossification is the process of mineralization undergone by most of the skeletal bones which involves replacement of cartilage by bone tissue (Lindsey and Mohan 2016). The cartilage is formed by the chondrocytes in the three zones of the growth plate representing different states of differentiation: the resting/germinal zone contains progenitor cells (pre-chondrocytes); the proliferating zone with dividing chondrocytes forming a column structure; and the hypertrophic/maturing zone where differentiated chondrocytes enlarge and are metabolically active and characterized by expression of the collagen X gene. These hypertrophic chondrocytes end their life cycle by undergoing apoptosis, enabling the process of osteogenesis to take place with calcification and the formation of new endochondral bone. Chondrocytes have been shown to express GHR and IGF1-R and studies have shown an indirect effect (through IGF-1) of GH on bone elongation (Barnard et al. 1988; Oberbauer and Peng 1995; Werther et al. 1993). IGF-1 then increases proliferation (clonal expansion) of the chondrocytes through autocrine/paracrine stimulation which ultimately enhances linear growth (Yakar and Isaksson 2016).

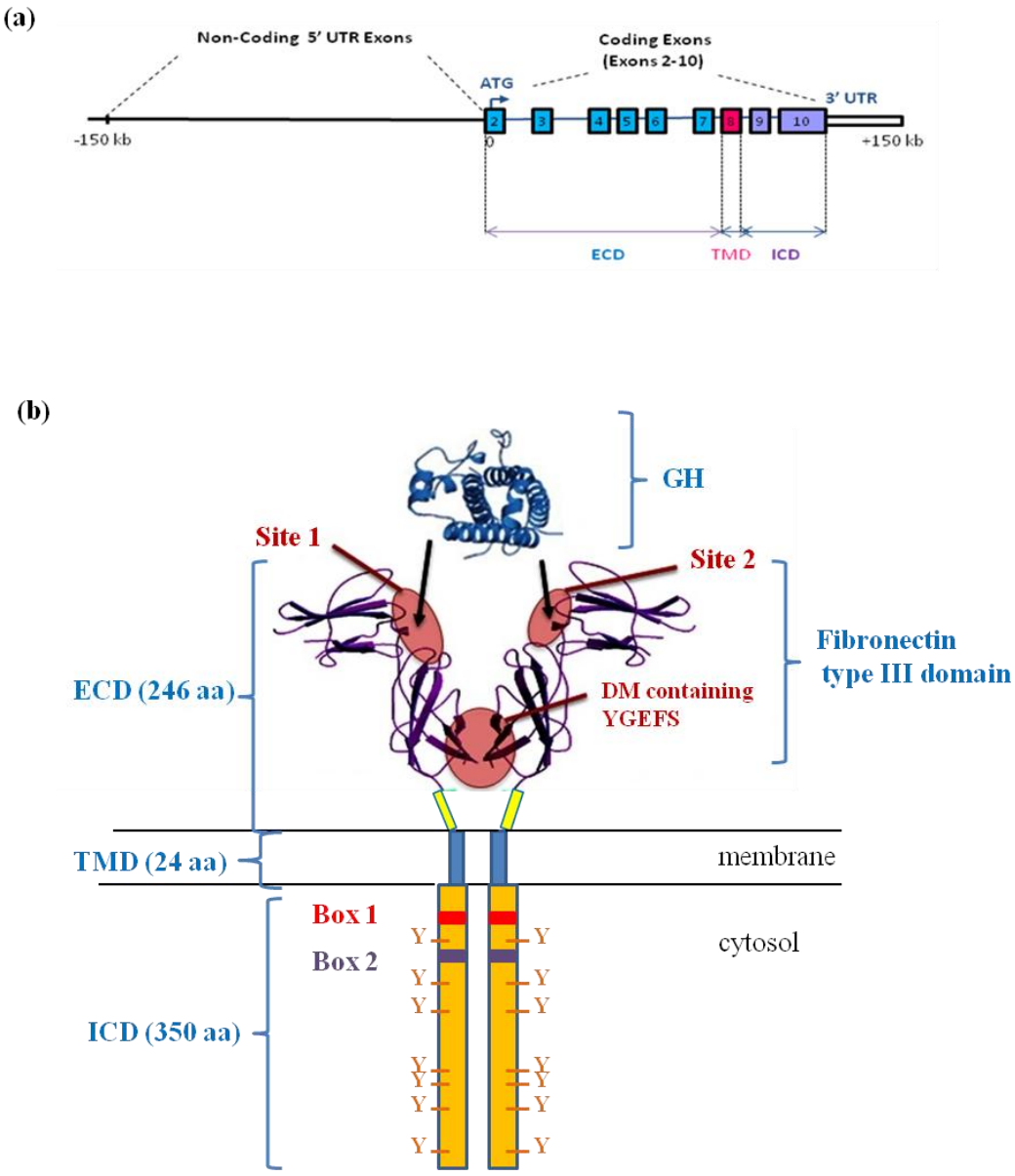


Figure I-4: The GHR gene and protein domain organizations

Adapted from (Waters and Brooks 2015)

Figure I-4: (a) The *GHR* gene spans ~300kb on the short arm of chromosome 5. The coding region is defined as exon 2 to 10 encoding a 638 aa protein that is divided into three domains: an extracellular domain (ECD) consisting of two β-sandwich domains with a fibronectin type III topology encoded by exons 2-7; a single pass helical transmembrane domain (TMD) mostly encoded by exon 8 and an intracellular domain (ICD) encoded by exons 9 and 10. **(b)** Schematic of the crystal structure of GH and the GHR (adapted from (Waters and Brooks 2015). GH, which comprises a four helix bundle structure, binds the dimerized GHR through asymmetrically positioned binding sites (site 1 and 2) stabilized by the extracellular receptor-receptor dimerization domain (DM). The binding of GH results in the rotation of one subunit of receptor relative to the other; this structural change is transmitted to the intracellular domain (ICD) through the TMD which ultimately results in the repositioning of the associated JAK2 to Box 1. JAK2 activation will lead to tyrosine phosphorylation (Y) on the intracellular domain, initiating several signaling pathways.

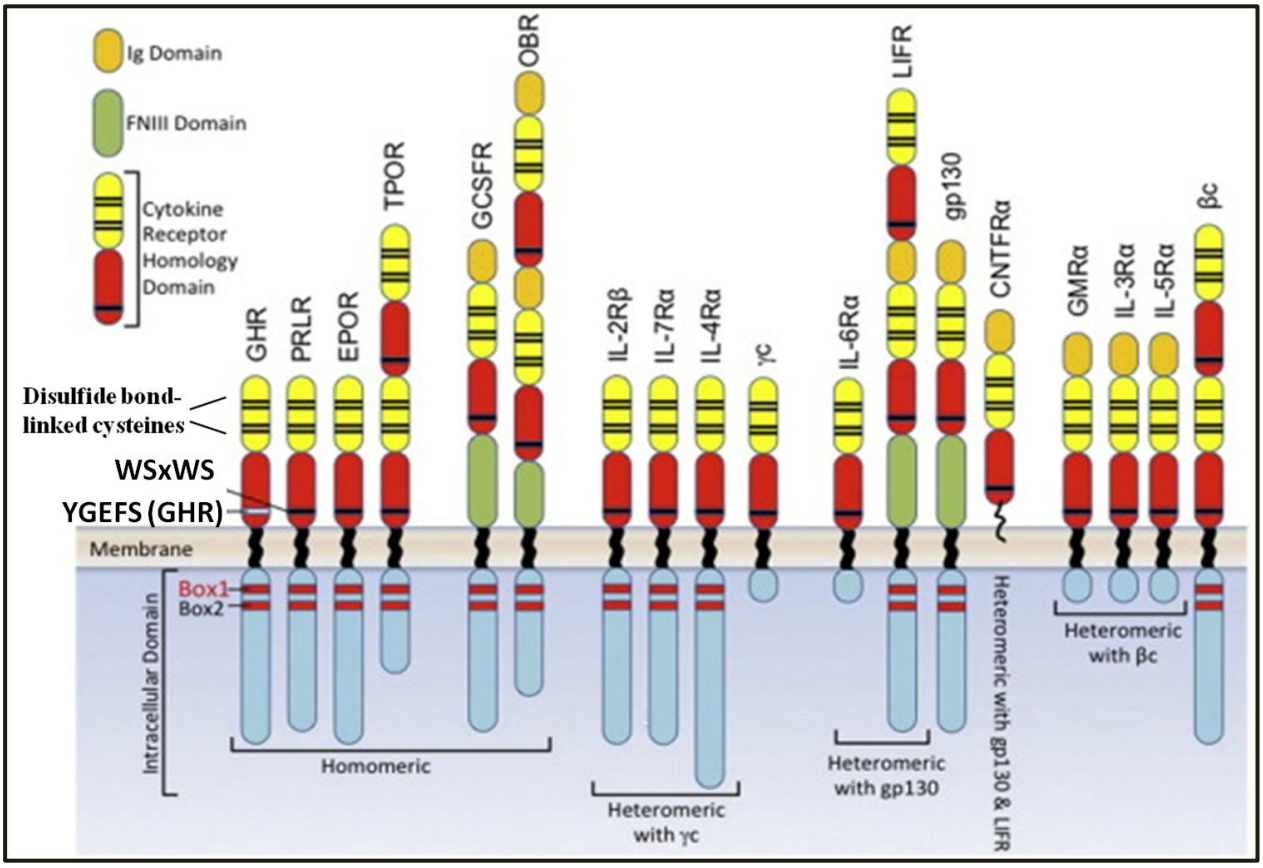


Figure I-5: The class I cytokine receptor family

Adapted from (Waters and Brooks 2015)

A schematic of the cytokine receptor domain structure, including the disulphide bond-linked cysteines (black thin lines) in the extracellular domains (three pairs for GHR and one free cysteine); the conserved WSxWS motif adjacent to the transmembrane domain (YGEFS for GHR); Box 1 and Box 2 sequences in the intracellular domain. The receptors of the class I cytokine family can form homodimers or heterodimers and common subunits shared by heteromeric receptors are indicated.

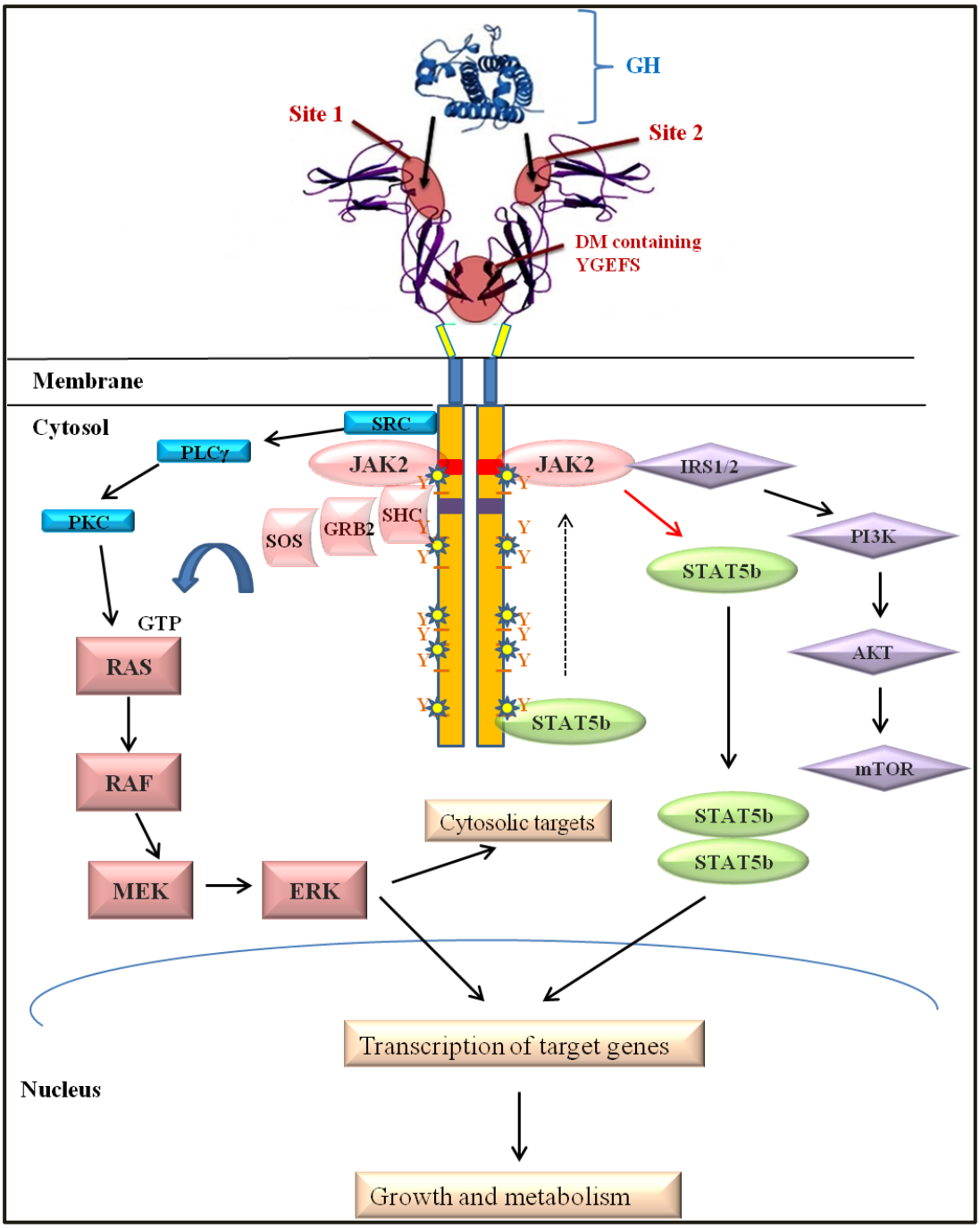


Figure I-6: Major signaling pathways of GH

Adapted from (Brooks and Waters 2010b)

Representation of the major signaling pathways in response to GH stimulation: the canonical JAK2-STAT5 pathway as well as the PI3K-AKT, MAPK/ERK and the SRC pathways. The relative activation of each pathway is dependent on the cell type.

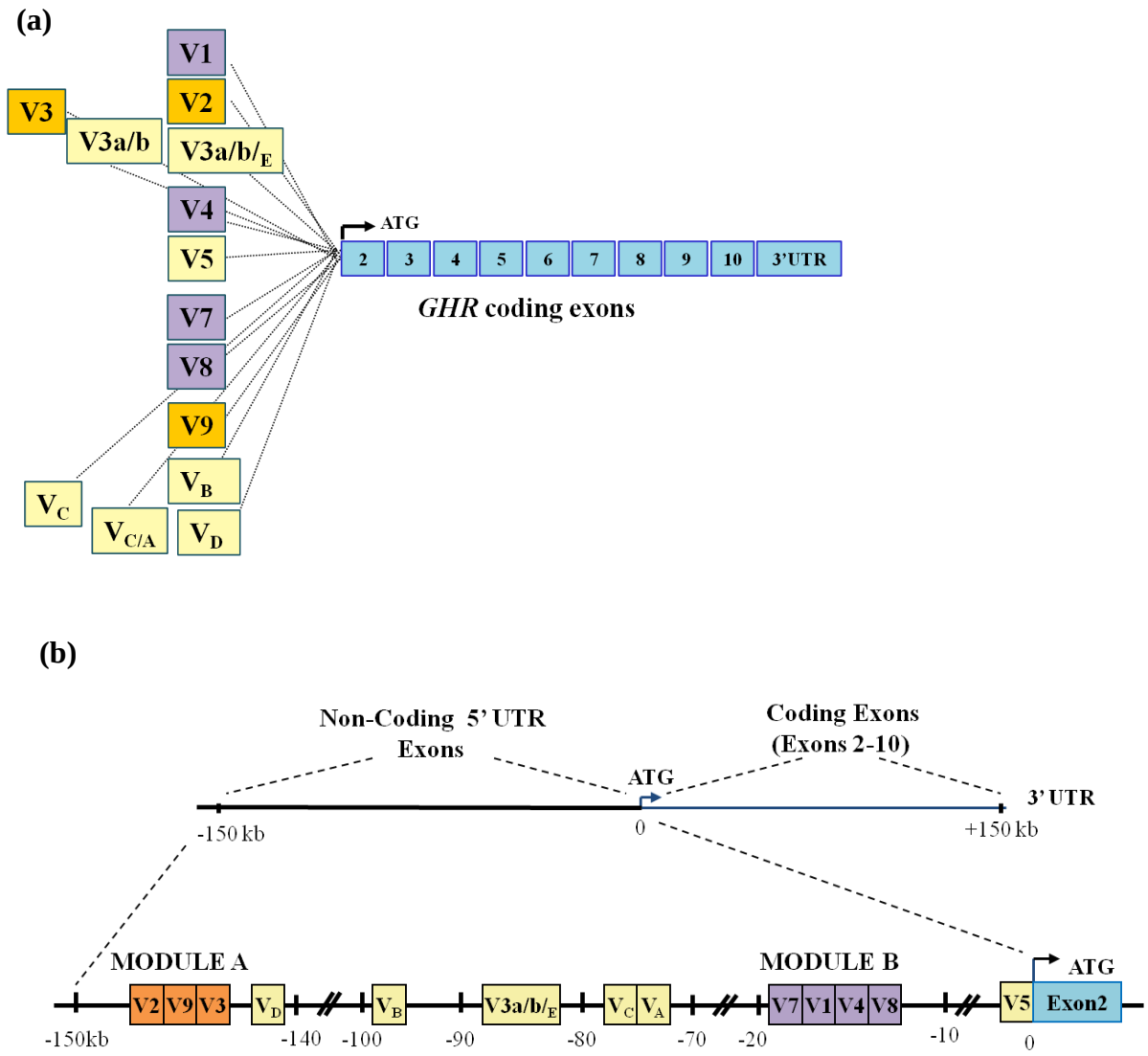


Figure I-7: *GHR* gene organisation

(a) Fourteen mRNAs for the full-length GHR are produced from the human *GHR* gene due to multiple 5'UTR variant (V) exons and alternative splicing. They all splice into the same site 11bp upstream of the translation start site in exon 2 and, thus, code for the same protein. **(b)** The *GHR* gene is located on the short arm of chromosome 5. Exons 2-10 code for the GHR protein. Seven of the non-coding variant (V) exons are clustered in 2 small regions defined as Module A (~1.6kb) and Module B (~2kb). VA-VD and V3a/b/VE exons are found between the two Modules. V5 is adjacent to the first coding exon, exon 2. V6 was an artifact.

CHAPTER II

Clinical studies have shown that the abnormal growth of children with Idiopathic Short Stature is due to growth hormone insensitivity. Their endocrine profile suggests a lower abundance of the GHR at the surface of target cells. In this chapter, I have conducted a case-control study with pediatric and adult short stature to investigate if sequence variants in the human *GHR* gene regulatory regions are associated with this phenotype.

(*J Cell Mol Med* 21(11): 2985-2999, 2017)

Genetic variations at the human *growth hormone receptor (GHR)* gene locus are associated with idiopathic short stature (ISS).

Christel Dias[1], Mara Giordano[2], Rosalie Frechette[3], Simonetta Bellone[4], Constantin Polychronakos[5], Laurent Legault[6], Cheri L Deal[7] and *Cynthia Gates Goodyer[1,8]

[1]Division of Experimental Medicine, McGill University, Montreal, Quebec, Canada; [2]Laboratory of Human Genetics, Department of Health Sciences, University of Eastern Piedmont, Novara, Italy; [3]Genotyping Platform, Genome-Québec, Montreal, Quebec, Canada; [4]Division of Pediatrics, Department of Health Science, University of Eastern Piedmont, Novara, Italy; [5]Departments of Experimental Medicine, Human Genetics and Pediatrics, McGill University, Montreal, Quebec, Canada; [6]Department of Pediatrics, McGill University, Montreal, Quebec, Canada; [7]CHU Ste-Justine Research Centre and Department of Pediatrics, Université de Montréal, Montreal, Quebec, Canada; [8]Departments of Experimental Medicine and Pediatrics, McGill University, Montreal, Quebec, Canada

Running Title: *GHR* gene variants and ISS

Number of Figures & Tables: 4 Figures & 3 Tables (+ Supplementary 2 Figures & 4 Tables)

Corresponding author and person to whom reprint requests should be addressed:

Dr. Cynthia Gates Goodyer, The Research Institute of McGill University Health Centre
Block E, EM0.3211, 1001 Decarie Boulevard, Montreal, Quebec, Canada H4A 3J1
Telephone: +(514) 934-1934, ext. 22481
Email: cindy.goodyer@muhc.mcgill.ca

1. ABSTRACT (250 words)

Background: GH plays an essential role in the growing child by binding to the GH receptor (GHR) on target cells and regulating multiple growth promoting and metabolic effects. Mutations in the *GHR* gene coding regions result in GH insensitivity (dwarfism) due to a dysfunctional receptor protein. However, children with idiopathic short stature (ISS) show growth impairment without GH or GHR defects. We hypothesized that decreased expression of the *GHR* gene may be involved.

Aims: To investigate whether common genetic variants (microsatellites, SNPs) in regulatory regions of the *GHR* gene region were associated with the ISS phenotype.

Methods/Results: Genotyping of a GT-repeat microsatellite in the *GHR* 5'UTR in a Montreal ISS cohort (n=37 ISS, n=105 controls) revealed that the incidence of the long/short(L/S) genotype was 3.3x higher in ISS children than controls (p=0.04, OR=3.85). In an Italian replication cohort (n=143 ISS, n=282 controls), the medium/short(M/S) genotype was 1.9x more frequent in the male ISS than controls (p=0.017, OR=2.26). In both ISS cohorts, logistic regression analysis of 27 SNPs showed an association of ISS with rs4292454, while haplotype analysis revealed specific risk haplotypes in the 3' haploblocks. In contrast, there were no differences in GT genotype frequencies in a cohort of short stature (SS) adults vs. controls (CARTaGENE: n=168 SS, n=207 controls) and the risk haplotype in the SS cohort was located in the most 5' haploblock.

Conclusions: These data suggest that the variants identified are potentially genetic markers specifically associated with the ISS phenotype.

KEY WORDS: Growth hormone receptor, ISS, short stature, GT microsatellite, SNPs, haplotype.

2. INTRODUCTION

The growth hormone (GH)-insulin-like growth factor 1(IGF-1) axis is recognized as a key regulator of normal musculoskeletal development in the child. If any member of this axis is defective, the result is a short stature phenotype (Wit and de Luca 2016). Loss-of-function mutations in the *GH receptor* (*GHR*) gene are a prime example since the ability of GH to exert its pleiotropic effects is contingent on the availability of its receptor at the surface of target cells (David et al. 2011). Individuals with a dysfunctional GHR or loss of GHR do not respond normally to GH: they are not only extremely short, they have decreased bone mineral density and increased adiposity, with a greater risk of osteoporosis, lipid disorders and cardiovascular disease (Bachrach et al. 1998; Benbassat et al. 2003; Savage et al. 2006).

However, in ~60-80% of childhood short stature cases (defined as a height z-score below -2 SD) no etiology can be found; the children have normal birth length, are GH sufficient and there is no evidence of systemic, endocrine, nutritional or chromosomal abnormalities (Cohen et al. 2008; Oostdijk et al. 2009). These individuals are defined as having idiopathic short stature (ISS). This is a heterogeneous population. Approximately half will be subcategorized as having 'constitutional growth delay': their pubertal development and growth spurt are delayed but, when they do occur, due to increased production of sex steroids, average final height is achieved (Cohen et al. 2008; Wit and de Luca 2016). A smaller percentage will be diagnosed with familial short stature as their predicted final heights are within the expected range for parental target height. The remaining children, who never achieve a "catch-up" growth and who have limited responsiveness to GH therapy, represent ~1-2% of populations worldwide (Bryant et al. 2007; Collett-Solberg 2011; Deodati and Cianfarani 2011).

One previous explanation for ISS was *GHR* haploinsufficiency. However, when the *GHR* coding exons were examined for heterozygous deleterious mutations, only 2-5% of ISS individuals were found to have coding sequence changes and most of these were not functionally significant (Bonioli et al. 2005; Goddard et al. 1995; Hujeirat et al. 2006). On the other hand, serum levels of GH binding protein, the product of enzymatic cleavage of cell-surface GHR, are often low and ~20% of ISS children have levels >2 SD below the mean, suggesting that tissue levels of GHR are chronically low. Additional evidence comes from studies of the African Baka pygmies. These individuals, who have a similar phenotype and endocrine profile as the ISS children, show a ~80% decrease of *GHR* mRNA in their

lymphocytes (Bozzola et al. 2009). Together these data suggest low transcription of the *GHR* gene may be occurring.

Height is a complex polygenic trait with a high heritability estimate (h^2~0.80-0.90) (Perola et al. 2007; Silventoinen et al. 2003). Since 2007, an increasing number of gene variants have been shown to be associated with height variation in the general population (Gudbjartsson et al. 2008; Lango Allen et al. 2010; Lettre et al. 2008; Weedon et al. 2008; Weedon et al. 2007; Wood et al. 2014). The most recent GWAS identified 697 common variants that explain ~16% of the adult height variation and implicate many genes and pathways important for skeletal growth (Wood et al. 2014). Surprisingly, the early reports, including the initial study from the GIANT consortium, did not find any association of common genetic variants in *GH/IGF-1* axis genes with adult height variation (Lango Allen et al. 2010; Lettre et al. 2007; Weedon et al. 2007). Only the later use of more dense gene-centric arrays revealed significant associations with some axis members, including *GHR*: two SNPs, rs17574650 in intron 1 (MAF<3%) and rs6180 in exon 10, were identified within the *GHR* locus (Lanktree et al. 2011; Wood et al. 2014).

One explanation for the relative lack of detection of the *GHR* gene is that sampling of population height in previous studies has generally excluded the extreme tails of the normal distribution. In addition, most of the common variants identified through GWAS studies have a small effect size and, thus, cannot explain all of the height variation. To understand this 'missing heritability', other classes and possible combinations of genetic variations need to be explored, including microsatellites (Eichler et al. 2010; Hannan 2010). Microsatellite polymorphisms are a ubiquitous class of simple repetitive DNA sequences (Kelkar et al. 2010). For example, the $(GT)_n$ repeat is frequently observed in the human genome; population studies show high mutability due to slippage, leading to complex polymorphic characteristics (Bhargava and Fuentes 2010; Lee et al. 1999). Interestingly, the length of the GT repeat in promoter regions has been shown to modulate flanking response elements in several genes (Chen et al. 2016b; Gao et al. 2004; Hata et al. 2000; Itokawa et al. 2003b; Lapoumeroulie et al. 1999; Searle and Blackwell 1999). We previously reported that the GT repeat in the *GHR* V9 promoter region is a microsatellite polymorphism, with 19-32 repeats in the general population (Hadjiyannakis et al. 2001).

In the present study, we have tested the hypothesis that GHR, because of its important position within the GH/IGF-1 axis, has a role in the occurrence of ISS. To do this, we have analyzed both the V9 microsatellite polymorphism as well as multiple SNPs within the *GHR* locus in two ISS cohorts and a cohort of short stature (SS) adults, along with their respective

controls, to determine if there is a *GHR* regulatory haplotype associated with the ISS phenotype.

3. MATERIALS AND METHODS

A. Study populations

The Montreal ISS cohort was comprised of 37 ISS children recruited from the Montreal Children's Hospital (MCH) and l'Hôpital Ste-Justine (HSJ) from 2009 to 2010 (**Table II-1**). The major ISS inclusion criterion was a height score (SDS or Z-score) ≤ -2 SD and normal stimulated (clonidine, arginine) GH levels (cut-off values of 5 [MCH] and 6.5 [HSJ] µg/L), with no evidence of organic disease, malnutrition, psychosocial issues, intrauterine growth retardation (IUGR) or hypothyroidism. We also obtained genomic DNA from 105 adults with normal final adult heights; these individuals were initially enrolled in the Type I Diabetes Susceptibility Study but are non-diabetic. Exclusion criteria for the adults included IUGR, small for gestational age (SGA), a childhood co-morbid disease or specific syndromes.

Genomic DNA samples from the Novara cohort have been described elsewhere (Fusco et al. 2016). Briefly, 143 ISS children, along with 282 normal stature healthy children matched for age and sex, were recruited by the Unit of Paediatrics of the Department of Health Sciences of Novara (Italy) (**Table II-1**). Criteria for ISS diagnosis were similar to the Montreal ISS cohort, including a height SDS ≤ -2 and GH sufficiency (cut-off value of 8 µg/L).

CARTaGENE is the largest population biobank in Quebec, with ~20,000 recruits aged 40-69 years at the time that our study was initiated (www.cartagene.qc.ca) (Awadalla et al. 2013); the participants represent a random selection of individuals residing in the metropolitain areas of Quebec. Our CARTaGENE cohort consisted of 168 short stature (SS) individuals with a final height corresponding to severe short stature (males: maximum height of 159 cm, -2.4 SDS; females: maximum height of 147 cm, -2.47 SDS) and 207 controls of average height (30^{th}-70^{th} percentile [~±0.5 SDS]: males were 173-180 cm; females were 160-167 cm) (**Table II-1**). There were no data available on whether the SS individuals had ever been diagnosed with ISS.

Recruitment of these patients was approved by local institutional review boards, including the Azienda Ospedaliera Universitaria Maggiore della Carita for the Italian cohort, the Research Ethics Committee of CHU Ste-Justine for CARTaGENE participants, and the

Research Ethics Boards at both CHU Ste-Justine and the McGill University Health Centre for the Montreal ISS cohort. In all cases, information on the recruits was anonymized prior to receiving the DNA samples for analysis. The majority (>95%) of the subjects in the three cohorts were from European ancestry. Asian, African and South American individuals were excluded to minimize stratification of the genetic results.

Analyses were also carried out on a pool of the two ISS cohorts. In addition, we created a pool of Montreal and CARTaGENE adult controls that conformed to the 30^{th}-70^{th} height percentiles on the WHO Growth Charts for Canada (2014; www.whogrowthcharts.ca).

B. Microsatellite genotyping

The GT microsatellite from the *GHR* V9 promoter region (chr5:42424274-42424321 hg19 Genome Assembly) was genotyped using fluorogenic probes followed by capillary electrophoresis (ABI Genetic Analyzer 3730 XL, Life Technologies, Foster City, CA, USA) to discriminate allele size (ABI GeneMapper Version 4.1). Primers were designed to amplify a 155 bp fragment containing a 24 GT repeat that was used as reference; deviation from this size allowed us to deduce the GT length of the different alleles (e.g. 157 bp = 25 repeats). The primers (Forward: 5'-6-FAM-TCCTCCTTGCGAAGAAGTTG-3' and Reverse: 5'-GTGTGATGGTTCGTCTGTCG-3') were used in a PCR reaction with Phusion enzyme (ThermoFisher Scientific, MA, USA) and 3% DMSO at an annealing temperature of 60ºC. Samples were then processed and analyzed at the Genotyping Platform at Génome Québec (Montreal, QC, Canada).

We classified the alleles arbitrarily into 3 categories in order that the cut-offs fall at ~± 1SD around the median. Thus, the shortest (S) alleles were <24 repeats (representing ~16% of the individuals), the medium (M) alleles were 24-28 repeats (~68%), and the longest (L) alleles were >28 repeats (~16%) (**Table II-S1).**

C. SNP selection

27 SNPs with minor allele frequencies >5% were selected to span the *GHR* gene region from ~200kb upstream of the major *GHR* (V2) transcriptional start site to ~120kb downstream of the 3'UTR (**Tables II-S3 and II-S4**). We prioritized SNPs previously shown to be in association with height (de Graaff et al. 2013; Lettre et al. 2007; Wood et al. 2014), transcription regulation (Verlaan et al. 2009) or disease risk (e.g. non-small cell lung cancer (Van Dyke et al. 2009) or prostate cancer (McKay et al. 2007)).

D. SNP genotyping and quality control

Genotyping was performed at Génome Québec using Sequenom iPLEX Gold Technology and the MassARRAY system (Agena Biosciences, CA, USA). To assess robustness of the technology ~20% of the total samples were replicated with 100% success rate. Quality controls were conducted prior to the analysis: all variants used for association analysis had a genotyping efficiency call rate >95% and showed no departure from Hardy-Weinberg Equilibrium in combined controls and cases ($p>0.001$) and in controls and cases separately. A table of the SNP probes is available on request.

E. Statistical analysis

Allelic frequency calculations of the GT polymorphisms were performed using GraphPad Prism v7.0 software (La Jolla, CA, USA). Significance was calculated using 2 × 2 contingency tables and Fisher's exact tests to obtain p-values, odds ratios (ORs) and 95% confidence intervals (95% CIs) as well as two-tailed unpaired t-tests and 2-way ANOVAs. Bonferroni corrections were applied following Fisher exact tests to account for the multiple testing of the GT genotype categories (p-corr <0.01 [0.05/5] and sex $p\text{-}corr_{sex}$ <0.005). ANOVA tests were followed by a Tukey post-hoc test; $p<0.05$ was considered significant. For case-control single marker as well as haplotype logistic regressions, we used the PLINK v1.07 software package (Barrett et al. 2005). Results of the logistic regression were adjusted for sex but also calculated for each gender separately. The GT genotypes were also used as covariates in the regression analysis after coding them as dummy (binary) variables. Two genetic models were tested: additive and recessive (Bush and Moore 2012). Measures of pairwise LD between SNPs (D' and r^2) and LD plots were computed using Haploview v4.2 (Broad Institute, Cambridge, MA, USA) and haplotype blocks were defined using the solid spine algorithm of LD (D'>0.8). Because of strong LD between certain SNPs in our panel, we used Haploview's Tagger software with a pairwise approach (r^2 >0.8) to calculate that 18 SNPs represented the effective number of independent SNPs to use in the Bonferroni correction method for multiple comparison. For single marker analysis, the significant p-value after correction by the number of effective SNPs was p-corr <0.0028 (0.05/18). For haplotype analysis, a permutation procedure was applied to correct for multiple comparisons.

4. RESULTS

A. Specific GT microsatellite genotypes are associated with Idiopathic Short Stature.

The *GHR* gene spans ~300kb on chromosome 5 (**Fig. II-1A**) (Barton et al. 1989b; Godowski et al. 1989). The coding region is defined by exons 2-10 where exon 2 contains the translation start site (Leung et al. 1987). Fourteen different *GHR* mRNAs encoding the full-length GHR have been reported to date (Goodyer et al. 2001a; Orlovskii et al. 2004; Pekhletsky et al. 1992; Wei et al. 2006). They each have a unique 5'UTR, derived from different first exons, but all splice into the same site in exon 2, 11bp upstream from the ATG translation start site and, thus, code for the same protein.

A $(GT)_n$ repeat polymorphism is located in the proximal promoter of V9, one of the major ubiquitously expressing *GHR* 5'UTR exons (**Figs. II-1C-D)**. This was genotyped in all three of our cohorts (**Fig. II-2**). The allelic distribution profile was similar across the cohorts and showed a (GT) repeat number ranging from 15 to 37, with a median average at 26 and an isolated peak at 19 (**Figs. II-2A-B, II-2D-E, II-2G-H**). The repeat length cutoffs for allelic categorization were defined arbitrarily as <24 repeats for the short (S) alleles, 24-28 for the medium (M) and >28 for the long (L) alleles; there were no significant differences in the frequencies of each category (S, M, L) amongst the three populations (**Table II-S1**). Following allele classification, we could attribute to each individual one of the six bi-allelic genotypes: L/L, L/M, L/S, M/M, M/S or S/S. Genotype distribution frequencies are shown for each cohort in **Figs. II-2C**, **2F** and **2I and Table S2**.

The Montreal ISS children showed a nominally significant 3.3-fold increase in L/S genotypes compared to their adult controls (**Fig. II-2C**) (Fisher exact test: p=0.04, OR= 3.85, 95% CI=1.28-12.92). Interestingly, the Montreal ISS L/S genotype carriers had a significantly lower average height z-score at diagnosis when compared to the non-L/S children: -3.16 ± 0.79 *vs.* -2.5 ± 0.37 (two-tailed unpaired t-test: p=0.0026) (**Fig. II-S1A**). The ISS individuals also showed an association between the GT genotypes and the height z-score at diagnosis, with the L/S genotype carriers being the shortest and the S/S the tallest, with a significant difference between the L/S and M/S carriers (one-way Anova and Tukey post-hoc test: p=0.026) (**Fig. II-S1A)**. At the time of recruitment, the L/S genotype children tended to remain clustered at the lowest part of the growth curves whereas the other genotype categories showed a percentage of individuals showing catch-up growth ($\geq 3^{rd}$ percentile) (**Fig. II-S1B**). In the Novara replication ISS cohort, there were 1.5-fold more M/S genotype carriers than in the normal height control children (**Fig. II-2F**) (Fisher exact test: p=0.053, OR=1.64,

95% CI=1.02-2.69). This difference was shown to be driven by the males (Fisher exact test: p=0.017, OR=2.26, 95% CI=1.19-4.26) although this was only nominally significant after correction for multiple comparisons and sex (**Fig. II-S2**). There were no differences in average height scores at the time of diagnosis between subjects of the M/S genotype compared to non-M/S carriers; there was no information available on heights at the time of recruitment.

To validate the significance of our results in the two ISS cohorts, and because of the potential bias in the Montreal cohort due to a high proportion of female adult controls, we performed an analysis using a pool of control adults (composed of the CARTaGENE controls and a matching restricted set of Montreal adult controls [30^{th}-70^{th} percentile]). For the Montreal ISS group, the result confirmed the L/S genotype as being significantly associated with the ISS phenotype (Fisher exact test: p=0.004, OR=4.9, 95% CI=1.88-13.39). For the Novara ISS cohort, there were still more males presenting with the M/S genotype than in the male controls (Fisher exact test: p=0.038, OR=2.1, 95% CI=1.14-4.03) although this difference was only nominally significant.

We subsequently pooled the two ISS groups and used the same control adult pool. Again, there were more subjects with the M/S genotype in the ISS group (Fisher exact test: p=0.055, OR=1.6, 95% CI=1.02-2.52) and this was driven by the males (Fisher exact test: p=0.0265, OR=2.08, 95% CI=1.14-3.96), similar to what we found in the original Novara ISS cohort. Interestingly, there were 2.3 times more females presenting with the L/S genotype in the pooled ISS cohorts compared to the pooled controls (9.8% *vs.* 4.3%, respectively) but this result did not reach significance.

For the CARTaGENE SS cohort, no significant genotype frequency differences were observed between the short stature adult group and their controls even when the sexes were analyzed separately. It is noteworthy that the homozygous L/L and S/S genotype frequencies were the lowest across all three cohorts; there were no significant differences between the sexes (**Figs. II-2C, II-2F, II-2I**).

B. Common variants in the *GHR* gene region are associated with idiopathic short stature.

To assess the association of common variants (MAF>5%) in the *GHR* gene region with short stature, we selected a panel of 27 SNPs that were previously used in different association studies (**Tables II-S3 and II-S4**) (de Graaff et al. 2013; Lettre et al. 2007; McKay et al. 2007; Van Dyke et al. 2009; Verlaan et al. 2009; Wood et al. 2014). Logistic regression

analyses were conducted using additive or recessive penetrance models within each of our three cohorts and the most significant association results are shown in **Tables II-2** and **II-3** as well as **Fig. II-3**. Because height is sexually dimorphic in humans, we also analyzed male and female groups separately.

In the Montreal ISS cohort, two SNPs showed the strongest association with the ISS phenotype: (i) the A allele of rs4273617 in intron 5 (*p=0.0003, OR=4.8) and (ii) the T allele of rs6873545 in intron 3 (*p=0.0025, OR=3.3); these remained significant after correction for multiple comparisons (**Table II-2 and Fig. II-3A**). The latter SNP, rs6873545, has been used to tag a common deletion that removes *GHR* exon 3 (Lettre et al. 2007). During evolution, a homologous recombination event resulted in *GHR* alleles in ~35% of humans that differ by the deletion of exon 3 (3-) (Pantel et al. 2000). In the present study, heterozygous (3+/3-) individuals in the Montreal ISS cohort represented ~32% of the total cohort and no children were homozygous for the exon 3 deletion; these results were confirmed by multiplex PCR assays (data not shown).

When the Montreal ISS cohort was analyzed using the larger pool of control adults, rs4273617 remained significant (A: *p=0.001, OR=3.2) and rs12153009 became significant, with a recessive penetrance (A: *p=0.001, OR=7.1) (**Table II-3**). These two SNPs delimit a cluster of nominally significant SNPs spanning ~141kb, from ~12kb upstream of the V5 5'UTR exon to intron 5 of the *GHR* gene.

When we compared regression analyses for the Montreal and Novara ISS cohorts that used the same pooled adult control group (**Table II-3**), two SNPs were replicated. Not only was there an increased risk associated specifically with the Novara males carrying the G allele of rs12233949 (p=0.013, OR=2.1) but the Montreal ISS risk allele (C) was now nominally significant (p=0.029, OR=1.9). In addition, rs4292454, also in intron 2, was significant for both ISS cohorts: the major T allele in the Montreal group was the risk allele (p=0.015, OR 2.1) while the minor C allele was the risk allele in the Novara ISS (p=0.007, OR 2.0). When the Montreal and Novara ISS cohorts were pooled and compared to the pooled adult controls (**Table II-3**), the risk C allele of rs4292454 remained significant in the recessive model (p=0.027, OR=1.7). Thus, rs4292454 showed the strongest association with the ISS phenotype.

In addition, in the Novara ISS cohort, four SNPs were nominally significant in the recessive model; the same SNPs showed similar association strengths in the CARTaGENE SS female cohort in the additive model (**Table II-2**). However, the allelic effects were opposite: minor alleles were protective in the Novara ISS cohort whereas major alleles were increasing

the risk in the CARTaGENE SS female cohort. The first three SNPs are clustered in the V2-V9-V3 region (**Fig. II-1C**) whereas rs2972419 is located in intron 1 (**Table II-S3**). This last SNP has been associated with pygmy short stature, with a higher proportion of the ancestral allele (G) in the pygmy population compared to non-pygmies, and with the derived allele (A) associated with taller stature (Becker et al. 2013). In the Novara cohort, the minor allele A showed a nominal association with the ISS phenotype but it had a protective odds ratio (p=0.024, OR=0.2). We looked for an effect on increased height in the Novara ISS children who were carrying the A allele (~45%), but did not observe a significant association with a higher height z-score; in addition, there was no significant difference between the z-scores of the three children with the AA genotype (-2.0 ±0.11 [M±SD]) and those of the total ISS group (-2.3±0.4). Thus, it is unlikely that this SNP is a major influence on the ISS phenotype.

For the CARTaGENE SS females, the major G allele of rs2972419 was associated with an increased risk of being short (p=0.031, OR=2.0) (**Table II-2**). However, the G allele was not associated with a smaller height z-score in the female group. The most significant SNP in the CARTaGENE SS cohort was rs666581 (A allele: p=0.024, OR=1.9) which is located in the distal promoter ~200kb upstream of the *GHR* V2 transcription start site (**Fig. II-3C, Table II-S3**), downstream of a lncRNA.

In order to assess a possible GT-SNP combinatorial effect, we conducted an independent set of logistic regression analyses adjusted for each GT genotype (data not shown). Using the GT genotypes as covariates did not alter the association results indicating that none of the GT genotypes have a significant effect in combination with our tested SNPs.

C. Haplotype variation in the *GHR* gene region and differences in the ISS and SS cohorts.

The *GHR* gene region has a relatively simple haplotype structure comprised of several large LD blocks (McKay et al. 2007). When we defined the LD structure of the *GHR* gene in the three cohorts based on our 27 genotyped SNPs, each group displayed a specific architecture, from 5 blocks in the Montreal and Novara ISS cohorts (**Figs. II-3A-B and II-4A-B**) to 4 blocks in the CARTaGENE SS group (**Figs. II-3C and II-4C**). The first 2 blocks span ~370kb of the *GHR* gene locus, comprising the distal promoter and 5' untranslated region. The third block (fourth for the Novara ISS cohort) contains the V5 promoter and V5 exon, exon 2 and intron 2. The last block (blocks 4 and 5 for Montreal ISS) starts from intron 3 and spans the remainder of the *GHR* coding region to downstream of the 3'UTR.

The two ISS cohorts showed significant haplotype variations in the last blocks: 4 and 5 for the Montreal ISS and 5 for the Novara ISS (**Figs. II-4A-B**). In the Montreal ISS cohort, the most significant risk haplotype was in block 4 (**Fig. II-4A**): the TTA haplotype (formed by the major alleles of rs4292454, rs6873545 and rs6886047) represented half of the haplotypes in block 4 and was associated with an increased risk of 2.5 times for the ISS phenotype (p=0.0043, OR=2.5, 95% CI=1.33-4.72; p=0.04 after permutation test) ; it is noteworthy that the rs4292454 and rs6873545 major T alleles were associated with the ISS phenotype as single markers (**Tables II-2 and II-3**). A second risk haplotype in block 5, AATT, was significantly associated with the ISS phenotype (p=0.043, OR=2.0, 95% CI=1.02-3.91) and included SNPs spanning from intron 5 (rs4273617) to ~120kb downstream of the *GHR 3'UTR*. In the Novara ISS group, individuals presenting with the TAGAT haplotype in block 5 showed a nominally increased risk for being ISS (p=0.038, OR=3.4; 95% CI=1.07-11.04); this haplotype was rare, as the frequency was 0.9% in the controls and 2.8% in the ISS group (**Fig. II-4B**). In the CARTaGENE SS cohort, an at-risk promoter haplotype in block 1 was nominally significant (p=0.046, OR=1.8, 95% CI=1.01-3.13) with no difference between the sexes (**Fig. II-4C**).

5. DISCUSSION

Our study examined the possible association of two types of genetic variations within the *GHR* gene locus, a GT microsatellite and SNPs, with the ISS phenotype. Our first goal was to investigate the highly polymorphic GT repeat located in the core promoter of the V9 5'UTR in a small exploratory cohort of ISS children and control adults recruited in Montreal. Genotyping demonstrated a high heterogeneity of the alleles, with lengths ranging from 15 to 37 repeats, confirming what we had previously shown in the general population (Hadjiyannakis et al. 2001). In addition, we identified a specific L/S genotype as being significantly more represented in the Montreal ISS cohort than adult controls. Interestingly, we also found that L/S children were the shortest group at the time of diagnosis and that none of the L/S children showed signs of catch-up growth compared to non L/S children at the time of recruitment.

To validate these results, we examined the GT microsatellite in a larger cohort of ISS children from Novara, Italy (Fusco et al. 2016), separately as well as in combination with our Montreal ISS group. In the Novara cohort, there were 1.9 fold more M/S males within the ISS group than in the control male children; when we combined the two ISS groups and

compared them with adult controls, this result was confirmed but was only nominally significant. In contrast, the L/S genotype was more frequent in the female ISS group, although this result also did not reach significance. The absence of an association of a GT genotype with short stature in the CARTaGENE adult cohort suggests a possible specificity of this polymorphism for the ISS phenotype.

Although our study is the first to indicate potential sex-specific differences for association of a microsatellite in *GHR* with the ISS phenotype, a CA repeat in the *IGF-1* gene promoter has previously been associated with short stature as well as constitutionally tall stature in a sex-specific manner (Hendriks et al. 2011; Rietveld et al. 2004). In the future, it would be of interest to investigate, in parallel, the *GHR* GT and the *IGF-I* CA repeat polymorphisms (and potentially others within the *GH-IGF-1* axis genes) to determine if there are separate or combinatorial mechanisms regulating the growth of ISS children.

Microsatellites represent ~3% of the human genome, with AC/GT repeats the second most common form. Promoter microsatellites, by expanding and contracting in length, are often polymorphic, mainly due to slippage during DNA replication (Sawaya et al. 2012a). They are found non-randomly distributed at a high density within promoters and, the more proximal they are to the transcription start site, the more likely it is that they are conserved (Sawaya et al. 2012b). Our V9 GT microsatellite is highly conserved in primates as well as the cow, the rat, the mouse and even the opossum, although the number of GT repeats varies across these species (UCSC browser).

Because their main location is within non-coding DNA, microsatellites have been traditionally considered to be neutral markers. However, many recent reports have shown that these polymorphic repeats can modulate gene transcriptional activity, particularly of genes involved in the regulation of growth and development (e.g. *PAX-6, COL1A2, gamma gene IV52S*) (Akai et al. 1999; Gemayel et al. 2012; Lapoumeroulie et al. 1999; Okladnova et al. 1998). Modulation of promoter activity by these microsatellites can ultimately lead to phenotypic alteration and disease states (e.g. *NRAMP1, COL1A2, GRIN2A, STAT6, heme-oxygenase 1*) (Chen et al. 2016b; Gao et al. 2004; Hata et al. 2000; Itokawa et al. 2003b; Searle and Blackwell 1999). For example, specific GT genotypes in the promoter of the *HO-1* (*heme-oxygenase 1*) gene have been linked to altered transcriptional activity (Chen et al. 2002; Hirai et al. 2003), due to modulation of flanking response elements (Chen et al. 2016b), and are associated with increased risk for several different diseases (cardiovascular disease, rheumatoid arthritis, type 2 diabetes mellitus, cancers) (Bao et al. 2010; Daenen et al. 2016; Kikuchi et al. 2005; Rueda et al. 2007). The cis-regulatory effects may also be explained, in

part, by the intrinsic property of the GT repeat sequences to form an alternative Z-DNA structure, primarily due to the alternation of purine-pyrimidine nucleotides (Rich and Zhang 2003; Wang and Vasquez 2007). The left-handed Z-DNA motif has been found to modulate promoter activity, likely due to the binding of specific Z-DNA binding proteins (Ray et al. 2011; Searle and Blackwell 1999). We are presently investigating the biological significance of the V9 GT microsatellite and its potential impact on *GHR* expression.

Many investigations have examined the SNPs associated with variation in height, primarily through GWAS studies of adults (Gudbjartsson et al. 2008; Lango Allen et al. 2010; Lettre et al. 2008; Weedon et al. 2008; Weedon et al. 2007; Wood et al. 2014). In our study, we focused on evaluating the possible contribution of common variants found within the *GHR* locus in ISS children and SS adults at the extreme tail of the height distribution. We selected a panel of 27 common SNPs (average MAF~27%) used in previous association studies that showed a strong potential for having a functional impact on *GHR* gene regulation (de Graaff et al. 2013; Lettre et al. 2007; McKay et al. 2007; Van Dyke et al. 2009; Verlaan et al. 2009). Indeed, we found two, rs6873545 and rs4273617, for which the major alleles were significantly associated with ISS in our Montreal cohort. Rs6873545 has been used as a tag for exon 3 deletion in *GHR*; the minor 3- allele codes for a GHR that is missing 22 amino acids in its extracellular domain, N-terminal to the GH binding domain (Pantel et al. 2000). Although there have been controversies regarding the physiological role of the exon 3-polymorphism, its major association appears to be restricted to an increased baseline height and growth velocity during the first year of GH treatment in GH-deficient (GHD) children; there was no relationship in non-GHD children, including ISS children (Wassenaar et al. 2009). Our study is in line with these results: the minor exon 3- allele was not associated with the ISS phenotype in the Montreal or Novara children (even when they were combined) or the SS phenotype in the CARTaGENE adults.

Interestingly, rs12233949 showed the same directional effect across the three different cohorts, even though we saw the opposite risk allele in the Novara ISS cohort. Baas et al (Baas et al. 2012) have reported similar opposite associations for three *SLC2A1* SNPs in two highly comparable populations. In addition, Lin et al (Lin et al. 2007) have observed what they called a "flip-flop" phenomenon for SNPs in the *COMT* and *GAPDH* genes in two different ethnic populations. This type of "flip-flop" association is difficult to explain but may be due to different genetic backgrounds and haplotypes within each ethnic group (Lin et al. 2007).

In addition, rs4292454 was consistently associated with ISS in the two ISS cohorts, alone or combined, when compared to the pooled adult controls. Both rs4292454 and rs12233949 are located in intron 2 and are in low LD (r^2~0.35 in the CEU population), suggesting that this region could be a site of variation associated with ISS. *GHR* intron 2 spans ~63kb and contains multiple DNase I hypersensitive sites and ChIP-validated binding sites for transcription factors, including for CTCF, a polyfunctional regulator that can mediate long-range chromatin looping and has been implicated in transcriptional regulation (Ong and Corces 2014).

In order to investigate the possibility of a multiple marker effect associated with ISS and SS, we undertook a haplotype analysis in our three cohorts. Our results showed a consistent association of the risk haplotypes located in the 3' haploblocks of the *GHR* gene with the ISS phenotype. These haplotypes encompass intron 2 to ~120kb downstream of the *GHR* 3'UTR. In a previous haplotype-based analysis, the same region was shown to be associated with prostate cancer risk in elderly men and with decreased BMI in the controls (McKay et al. 2007). Interestingly, in the CARTaGENE cohort, the risk haplotype for the adult SS phenotype showed a shift to the 5' region of the *GHR* locus, >200kb upstream of V2. The results of the haplotype analyses suggest distinct genetic variations for the ISS *vs.* SS phenotype and potentially different mechanisms for regulating *GHR* expression.

CONCLUSION AND LIMITATIONS

The goal of the present study was to assess the potential role of the *GHR* gene in ISS by examining two different types of genetic polymorphisms. Our case-control analysis revealed a significant association of a GT microsatellite in the *GHR* promoter with ISS, while SNP analysis showed a consistent direction of effect for one common variant in intron 2 of *GHR* (rs4292454). There are two major limitations of this study: the size of the ISS cohorts and the fact that the majority of our ISS participants were recruited before they had achieved their final height. It is also important to recognize that, because the mean height of some of our ISS individuals at diagnosis was <1st percentile, their short stature phenotype may not be explained by only common variants in the *GHR* gene locus. When Chan et al. (Chan et al. 2011) examined cohorts of extreme short and tall stature, they found that common genetic variants associated with height in the general population are also associated with height of individuals at the ~1st percentile but are less predictive at the most extreme end (~0.25 percentile, <-2.8 SDS). Finally, even though the participants of our cohorts were unrelated and the majority was from European ancestry, we cannot exclude the possibility of a certain

level of stratification that was not taken into account in this analysis. These findings underline the necessity to look in larger ISS cohorts for additional genetic contributors (e.g. rare variants, additional tandem repeats, structural polymorphisms, gene-gene or gene-environment interactions) to better define the association of the *GHR* gene with the idiopathic short stature phenotype.

ACKNOWLEDGEMENTS

We would like to acknowledge with thanks the contributions of Diane Laforte (recruiting the Montreal Children's Hospital ISS and control participants), Maryse Lagacé (recruiting l'Hopital Ste-Justine ISS participants), Deepak Babu (obtaining the Novara samples), the CARTaGENE Biobank and Databank (making samples and data available for the CARTaGENE participants), Robert Eveleigh (statistical support) and Dr. Aimee Ryan (critical review of the manuscript). CD received doctoral studentship awards from the Montreal Children's Hospital Research Institute and the Fonds de recherche du Québec – Santé (FRQ-S). The project was funded by the Canadian Institutes of Health Research (CGG MOP-119313). This research was undertaken at the Research Institute of the McGill University Health Centre which is supported in part by the FRQ-S. CD was responsible for data collection and analysis as well as writing of the manuscript, as part of her PhD program; RF contributed the genotyping analyses; MG and SB provided the Novara ISS and control cohorts while CP, LL and CLD provided the Montreal ISS children and control adults; CGG was the principal investigator and coordinator of the project, reviewer of the data, analyses and manuscript; all authors reviewed the manuscript.

CONFLICT OF INTEREST STATEMENT: the authors do not have anything to disclose.

6. REFERENCES

Akai J, Kimura A, and Hata RI. 1999. Transcriptional regulation of the human type I collagen alpha2 (COL1A2) gene by the combination of two dinucleotide repeats. Gene 239(1):65-73.

Awadalla P, Boileau C, Payette Y, Idaghdour Y, Goulet JP, Knoppers B, Hamet P, Laberge C, and Project CA. 2013. Cohort profile of the CARTaGENE study: Quebec's population-based biobank for public health and personalized genomics. International journal of epidemiology 42(5):1285-1299.

Baas DC, Ho L, Tanck MW, Fritsche LG, Merriam JE, van het Slot R, Koeleman BP, Gorgels TG, van Duijn CM, Uitterlinden AG et al. . 2012. Multicenter cohort association study of SLC2A1 single nucleotide polymorphisms and age-related macular degeneration. Molecular vision 18:657-674.

Bachrach LK, Marcus R, Ott SM, Rosenbloom AL, Vasconez O, Martinez V, Martinez AL, Rosenfeld RG, and Guevara-Aguirre J. 1998. Bone mineral, histomorphometry, and body composition in adults with growth hormone receptor deficiency. Journal of bone and mineral research : the official journal of the American Society for Bone and Mineral Research 13(3):415-421.

Bao W, Song F, Li X, Rong S, Yang W, Wang D, Xu J, Fu J, Zhao Y, and Liu L. 2010. Association between heme oxygenase-1 gene promoter polymorphisms and type 2 diabetes mellitus: a HuGE review and meta-analysis. Am J Epidemiol 172(6):631-636.

Barton DE, Foellmer BE, Wood WI, and Francke U. 1989. Chromosome mapping of the growth hormone receptor gene in man and mouse. Cytogenet Cell Genet 50(2-3):137-141.

Becker NS, Verdu P, Georges M, Duquesnoy P, Froment A, Amselem S, Le Bouc Y, and Heyer E. 2013. The role of GHR and IGF1 genes in the genetic determination of African pygmies' short stature. Eur J Hum Genet 21(6):653-658.

Benbassat CA, Eshed V, Kamjin M, and Laron Z. 2003. Are adult patients with Laron syndrome osteopenic? A comparison between dual-energy X-ray absorptiometry and volumetric bone densities. The Journal of clinical endocrinology and metabolism 88(10):4586-4589.

Bhargava A, and Fuentes FF. 2010. Mutational dynamics of microsatellites. Mol Biotechnol 44(3):250-266.

Bonioli E, Taro M, Rosa CL, Citana A, Bertorelli R, Morcaldi G, Gastaldi R, and Coviello DA. 2005. Heterozygous mutations of growth hormone receptor gene in children with idiopathic short stature. Growth Horm IGF Res 15(6):405-410.

Bozzola M, Travaglino P, Marziliano N, Meazza C, Pagani S, Grasso M, Tauber M, Diegoli M, Pilotto A, Disabella E et al. . 2009. The shortness of Pygmies is associated with

severe under-expression of the growth hormone receptor. Mol Genet Metab 98(3):310-313.

Bryant J, Baxter L, Cave CB, and Milne R. 2007. Recombinant growth hormone for idiopathic short stature in children and adolescents. Cochrane Database Syst Rev(3):CD004440.

Bush WS, and Moore JH. 2012. Chapter 11: Genome-wide association studies. PLoS Comput Biol 8(12):e1002822.

Chan Y, Holmen OL, Dauber A, Vatten L, Havulinna AS, Skorpen F, Kvaloy K, Silander K, Nguyen TT, Willer C et al. . 2011. Common variants show predicted polygenic effects on height in the tails of the distribution, except in extremely short individuals. PLoS Genet 7(12):e1002439.

Chen WJ, Chen YH, Lai YJ, Hsu YJ, Yeh YH, Tsai CS, and Lin CY. 2016. GT-repeat length polymorphism in heme oxygenase-1 promoter determines the effect of cilostazol on vascular smooth muscle cells. Int J Cardiol 222:407-415.

Chen YH, Lin SJ, Lin MW, Tsai HL, Kuo SS, Chen JW, Charng MJ, Wu TC, Chen LC, Ding YA et al. . 2002. Microsatellite polymorphism in promoter of heme oxygenase-1 gene is associated with susceptibility to coronary artery disease in type 2 diabetic patients. Hum Genet 111(1):1-8.

Cohen P, Rogol AD, Deal CL, Saenger P, Reiter EO, Ross JL, Chernausek SD, Savage MO, Wit JM, and participants ISSCW. 2008. Consensus statement on the diagnosis and treatment of children with idiopathic short stature: a summary of the Growth Hormone Research Society, the Lawson Wilkins Pediatric Endocrine Society, and the European Society for Paediatric Endocrinology Workshop. The Journal of clinical endocrinology and metabolism 93(11):4210-4217.

Collett-Solberg PF. 2011. Update in growth hormone therapy of children. The Journal of clinical endocrinology and metabolism 96(3):573-579.

Daenen KE, Martens P, and Bammens B. 2016. Association of HO-1 (GT)n Promoter Polymorphism and Cardiovascular Disease: A Reanalysis of the Literature. Can J Cardiol 32(2):160-168.

David A, Hwa V, Metherell LA, Netchine I, Camacho-Hubner C, Clark AJ, Rosenfeld RG, and Savage MO. 2011. Evidence for a continuum of genetic, phenotypic, and biochemical abnormalities in children with growth hormone insensitivity. Endocr Rev 32(4):472-497.

de Graaff LC, Clark AJ, Tauber M, Ranke MB, Johnston LB, Caliebe J, Molinas C, Amin N, van Duijn C, Wollmann H et al. . 2013. Association analysis of ten candidate genes in a large multinational cohort of small for gestational age children and children with idiopathic short stature (NESTEGG study). Horm Res Paediatr 80(6):466-476.

Deodati A, and Cianfarani S. 2011. Impact of growth hormone therapy on adult height of children with idiopathic short stature: systematic review. BMJ 342:c7157.

Eichler EE, Flint J, Gibson G, Kong A, Leal SM, Moore JH, and Nadeau JH. 2010. Missing heritability and strategies for finding the underlying causes of complex disease. Nat Rev Genet 11(6):446-450.

Fusco I, Babu D, Mellone S, Barizzone N, Prodam F, Fanelli A, Muniswamy R, Petri A, Bellone S, Bona G et al. . 2016. Variations in the high-mobility group-A2 gene (HMGA2) are associated with idiopathic short stature. Pediatr Res 79(2):258-261.

Gao PS, Heller NM, Walker W, Chen CH, Moller M, Plunkett B, Roberts MH, Schleimer RP, Hopkin JM, and Huang SK. 2004. Variation in dinucleotide (GT) repeat sequence in the first exon of the STAT6 gene is associated with atopic asthma and differentially regulates the promoter activity in vitro. J Med Genet 41(7):535-539.

Gemayel R, Cho J, Boeynaems S, and Verstrepen KJ. 2012. Beyond junk-variable tandem repeats as facilitators of rapid evolution of regulatory and coding sequences. Genes (Basel) 3(3):461-480.

Goddard AD, Covello R, Luoh SM, Clackson T, Attie KM, Gesundheit N, Rundle AC, Wells JA, and Carlsson LM. 1995. Mutations of the growth hormone receptor in children with idiopathic short stature. The Growth Hormone Insensitivity Study Group. The New England journal of medicine 333(17):1093-1098.

Godowski PJ, Leung DW, Meacham LR, Galgani JP, Hellmiss R, Keret R, Rotwein PS, Parks JS, Laron Z, and Wood WI. 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(20):8083-8087.

Goodyer CG, Figueiredo RM, Krackovitch S, De Souza Li L, Manalo JA, and Zogopoulos G. 2001. Characterization of the growth hormone receptor in human dermal fibroblasts and liver during development. Am J Physiol Endocrinol Metab 281(6):E1213-1220.

Gudbjartsson DF, Walters GB, Thorleifsson G, Stefansson H, Halldorsson BV, Zusmanovich P, Sulem P, Thorlacius S, Gylfason A, Steinberg S et al. . 2008. Many sequence variants affecting diversity of adult human height. Nat Genet 40(5):609-615.

Hadjiyannakis S, Zheng H, Hendy GN, and Goodyer CG. 2001. GT repeat polymorphism in the 5' flanking region of the human growth hormone receptor gene. Mol Cell Probes 15(4):239-242.

Hannan AJ. 2010. Tandem repeat polymorphisms: modulators of disease susceptibility and candidates for 'missing heritability'. Trends Genet 26(2):59-65.

Hata R, Akai J, Kimura A, Ishikawa O, Kuwana M, and Shinkai H. 2000. Association of functional microsatellites in the human type I collagen alpha2 chain (COL1A2) gene with systemic sclerosis. Biochem Biophys Res Commun 272(1):36-40.

Hendriks AE, Brown MR, Boot AM, Oostra BA, de Jong FH, Drop SL, and Parks JS. 2011. Common polymorphisms in the GH/IGF-1 axis contribute to growth in extremely tall subjects. Growth Horm IGF Res 21(6):318-324.

Hirai H, Kubo H, Yamaya M, Nakayama K, Numasaki M, Kobayashi S, Suzuki S, Shibahara S, and Sasaki H. 2003. Microsatellite polymorphism in heme oxygenase-1 gene

promoter is associated with susceptibility to oxidant-induced apoptosis in lymphoblastoid cell lines. Blood 102(5):1619-1621.

Hujeirat Y, Hess O, Shalev S, and Tenenbaum-Rakover Y. 2006. Growth hormone receptor sequence changes do not play a role in determining height in children with idiopathic short stature. Horm Res 65(4):210-216.

Itokawa M, Yamada K, Yoshitsugu K, Toyota T, Suga T, Ohba H, Watanabe A, Hattori E, Shimizu H, Kumakura T et al. . 2003. A microsatellite repeat in the promoter of the N-methyl-D-aspartate receptor 2A subunit (GRIN2A) gene suppresses transcriptional activity and correlates with chronic outcome in schizophrenia. Pharmacogenetics 13(5):271-278.

Kelkar YD, Strubczewski N, Hile SE, Chiaromonte F, Eckert KA, and Makova KD. 2010. What is a microsatellite: a computational and experimental definition based upon repeat mutational behavior at A/T and GT/AC repeats. Genome Biol Evol 2:620-635.

Kikuchi A, Yamaya M, Suzuki S, Yasuda H, Kubo H, Nakayama K, Handa M, Sasaki T, Shibahara S, Sekizawa K et al. . 2005. Association of susceptibility to the development of lung adenocarcinoma with the heme oxygenase-1 gene promoter polymorphism. Hum Genet 116(5):354-360.

Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S et al. . 2010. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 467(7317):832-838.

Lanktree MB, Guo Y, Murtaza M, Glessner JT, Bailey SD, Onland-Moret NC, Lettre G, Ongen H, Rajagopalan R, Johnson T et al. . 2011. Meta-analysis of Dense Genecentric Association Studies Reveals Common and Uncommon Variants Associated with Height. Am J Hum Genet 88(1):6-18.

Lapoumeroulie C, Castiglia L, Ruberto C, Fichera M, Amata S, Labie D, and Ragusa A. 1999. Genetic variations in human fetal globin gene microsatellites and their functional relevance. Hum Genet 104(4):307-314.

Lee JS, Hanford MG, Genova JL, and Farber RA. 1999. Relative stabilities of dinucleotide and tetranucleotide repeats in cultured mammalian cells. Hum Mol Genet 8(13):2567-2572.

Lettre G, Butler JL, Ardlie KG, and Hirschhorn JN. 2007. Common genetic variation in eight genes of the GH/IGF1 axis does not contribute to adult height variation. Hum Genet 122(2):129-139.

Lettre G, Jackson AU, Gieger C, Schumacher FR, Berndt SI, Sanna S, Eyheramendy S, Voight BF, Butler JL, Guiducci C et al. . 2008. Identification of ten loci associated with height highlights new biological pathways in human growth. Nat Genet 40(5):584-591.

Leung DW, Spencer SA, Cachianes G, Hammonds RG, Collins C, Henzel WJ, Barnard R, Waters MJ, and Wood WI. 1987. Growth hormone receptor and serum binding protein: purification, cloning and expression. Nature 330(6148):537-543.

Lin PI, Vance JM, Pericak-Vance MA, and Martin ER. 2007. No gene is an island: the flip-flop phenomenon. Am J Hum Genet 80(3):531-538.

McKay JD, Kaaks R, Johansson M, Biessy C, Wiklund F, Balter K, Adami HO, Boillot C, Gioia-Patricola L, Canzian F et al. . 2007. Haplotype-based analysis of common variation in the growth hormone receptor gene and prostate cancer risk. Cancer Epidemiol Biomarkers Prev 16(1):169-173.

Okladnova O, Syagailo YV, Tranitz M, Stober G, Riederer P, Mossner R, and Lesch KP. 1998. A promoter-associated polymorphic repeat modulates PAX-6 expression in human brain. Biochem Biophys Res Commun 248(2):402-405.

Ong CT, and Corces VG. 2014. CTCF: an architectural protein bridging genome topology and function. Nat Rev Genet 15(4):234-246.

Oostdijk W, Grote FK, de Muinck Keizer-Schrama SM, and Wit JM. 2009. Diagnostic approach in children with short stature. Horm Res 72(4):206-217.

Orlovskii IV, Sverdlova PS, and Rubtsov PM. 2004. [Fine structure, expression and polymorphism of the human growth hormone receptor gene]. Mol Biol (Mosk) 38(1):29-39.

Pantel J, Machinis K, Sobrier ML, Duquesnoy P, Goossens M, and Amselem S. 2000. Species-specific alternative splice mimicry at the growth hormone receptor locus revealed by the lineage of retroelements during primate evolution. J Biol Chem 275(25):18664-18669.

Pekhletsky RI, Chernov BK, and Rubtsov PM. 1992. Variants of the 5'-untranslated sequence of human growth hormone receptor mRNA. Mol Cell Endocrinol 90(1):103-109.

Perola M, Sammalisto S, Hiekkalinna T, Martin NG, Visscher PM, Montgomery GW, Benyamin B, Harris JR, Boomsma D, Willemsen G et al. . 2007. Combined genome scans for body stature in 6,602 European twins: evidence for common Caucasian loci. PLoS Genet 3(6):e97.

Pruim RJ, Welch RP, Sanna S, Teslovich TM, Chines PS, Gliedt TP, Boehnke M, Abecasis GR, and Willer CJ. 2010. LocusZoom: regional visualization of genome-wide association scan results. Bioinformatics 26(18):2336-2337.

Ray BK, Dhar S, Shakya A, and Ray A. 2011. Z-DNA-forming silencer in the first exon regulates human ADAM-12 gene expression. Proc Natl Acad Sci U S A 108(1):103-108.

Rich A, and Zhang S. 2003. Timeline: Z-DNA: the long road to biological function. Nat Rev Genet 4(7):566-572.

Rietveld I, Janssen JA, van Rossum EF, Houwing-Duistermaat JJ, Rivadeneira F, Hofman A, Pols HA, van Duijn CM, and Lamberts SW. 2004. A polymorphic CA repeat in the IGF-I gene is associated with gender-specific differences in body height, but has no effect on the secular trend in body height. Clin Endocrinol (Oxf) 61(2):195-203.

Rueda B, Oliver J, Robledo G, Lopez-Nevot MA, Balsa A, Pascual-Salcedo D, Gonzalez-Gay MA, Gonzalez-Escribano MF, and Martin J. 2007. HO-1 promoter polymorphism associated with rheumatoid arthritis. Arthritis Rheum 56(12):3953-3958.

Savage MO, Attie KM, David A, Metherell LA, Clark AJ, and Camacho-Hubner C. 2006. Endocrine assessment, molecular characterization and treatment of growth hormone insensitivity disorders. Nat Clin Pract Endocrinol Metab 2(7):395-407.

Sawaya SM, Bagshaw AT, Buschiazzo E, and Gemmell NJ. 2012a. Promoter microsatellites as modulators of human gene expression. Advances in experimental medicine and biology 769:41-54.

Sawaya SM, Lennon D, Buschiazzo E, Gemmell N, and Minin VN. 2012b. Measuring microsatellite conservation in mammalian evolution with a phylogenetic birth-death model. Genome Biol Evol 4(6):636-647.

Searle S, and Blackwell JM. 1999. Evidence for a functional repeat polymorphism in the promoter of the human NRAMP1 gene that correlates with autoimmune versus infectious disease susceptibility. J Med Genet 36(4):295-299.

Silventoinen K, Sammalisto S, Perola M, Boomsma DI, Cornes BK, Davis C, Dunkel L, De Lange M, Harris JR, Hjelmborg JV et al. . 2003. Heritability of adult body height: a comparative study of twin cohorts in eight countries. Twin Res 6(5):399-408.

Van Dyke AL, Cote ML, Wenzlaff AS, Abrams J, Land S, Iyer P, and Schwartz AG. 2009. Chromosome 5p Region SNPs Are Associated with Risk of NSCLC among Women. J Cancer Epidemiol 2009:242151.

Verlaan DJ, Ge B, Grundberg E, Hoberman R, Lam KC, Koka V, Dias J, Gurd S, Martin NW, Mallmin H et al. . 2009. Targeted screening of cis-regulatory variation in human haplotypes. Genome Res 19(1):118-127.

Wang G, and Vasquez KM. 2007. Z-DNA, an active element in the genome. Front Biosci 12:4424-4438.

Wassenaar MJ, Dekkers OM, Pereira AM, Wit JM, Smit JW, Biermasz NR, and Romijn JA. 2009. Impact of the exon 3-deleted growth hormone (GH) receptor polymorphism on baseline height and the growth response to recombinant human GH therapy in GH-deficient (GHD) and non-GHD children with short stature: a systematic review and meta-analysis. The Journal of clinical endocrinology and metabolism 94(10):3721-3730.

Weedon MN, Lango H, Lindgren CM, Wallace C, Evans DM, Mangino M, Freathy RM, Perry JR, Stevens S, Hall AS et al. . 2008. Genome-wide association analysis identifies 20 loci that influence adult height. Nat Genet 40(5):575-583.

Weedon MN, Lettre G, Freathy RM, Lindgren CM, Voight BF, Perry JR, Elliott KS, Hackett R, Guiducci C, Shields B et al. . 2007. A common variant of HMGA2 is associated with adult and childhood height in the general population. Nat Genet 39(10):1245-1250.

Wei Y, Rhani Z, and Goodyer CG. 2006. Characterization of growth hormone receptor messenger ribonucleic acid variants in human adipocytes. J Clin Endocrinol Metab 91(5):1901-1908.

Wit JM, and de Luca F. 2016. Atypical defects resulting in growth hormone insensitivity. Growth Horm IGF Res 28:57-61.

Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, Chu AY, Estrada K, Luan J, Kutalik Z et al. . 2014. Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet 46(11):1173-1186.

Table II-1: Baseline characteristics of study participants.

		Montreal Cohort			Novara Cohort		CARTaGENE Cohort	
		ISS	Controls	Controls[a]	ISS	Controls	SS	Controls
Total number of individuals		37	105	57	143	282	168	207
Sex	**Males**	23	36	16	76	159	92	108
	Females	14	69	41	67	123	76	99
Age[b] (years)	**Median [range]**	9.91 [2.5, 14.5]	43 [28, 60]	42.5 [28, 58]	10.9 (5.1,17.8) [c]		59.34 [41.1, 69.5]	57.85 [40.3, 70.1]
Height z-score[d] (SDS)	**Median [range]**	-2.49 [-4.56,-1.84]	-0.02 [-1.83,1.57]	0.01 [-0.55, 0.57]	-2.2 [-4.6, -1.5]	n/a	-2.675 [-3.83, -2.4]	0.01 [-0.49, 0.59]
	Males	-2.41 [-3.81,-1.84]	-0.085[-1.36,1.57]	-0.01 [-0.55, 0.49]	-2.2 [-3.5, -1.5]	n/a	-2.615 [-3.83, -2.4]	-0.08[-0.49, 0.47]
	Females	-2.8 [-4.56,-2.09]	-0.02[-1.83, 1.38]	0.02 [-0.54, 0.57]	-2.2 [-4.6, -1.8]	n/a	-2.49 [-3.8, -2.47]	0.08 [-0.48, 0.59]

[a] Montreal adult control group restricted to heights matching the CARTaGENE adult controls (30^{th} -70^{th} percentile).

[b] Age at diagnosis for the two ISS cohorts and age at enrolment for the CARTaGENE cohort.

[c] Median (range) age for Novara ISS and control groups (age-matched controls) (34).

[d] Height z-scores (Standard Deviation Scores) were calculated using the World Health Organization Growth Charts for Canada (2014). Height z-scores were established at the time of diagnosis for the two ISS pediatric cohorts and at enrolment for the CARTaGENE adult cohort. n/a: not available

Table II-2: SNP association analysis of short stature in three different cohorts

	Montreal ISS						Novara ISS						CARTaGENE SS					
SNP	**Risk allele[a]**	**Model[b]**	**Sex[c]**	**P value[d]**	**OR[e]**	**95% CI[f]**	**Risk allele**	**Model**	**Sex**	**P value**	**OR**	**95% CI**	**Risk allele**	**Model**	**Sex**	**P value**	**OR**	**95% CI**
rs666581	C	A	B	0.248	3.4	0.43-26.47	C	A	F	0.352	1.5	0.63-3.73	A	A	B	**0.024**	1.9	1.09-3.29
rs3764451	C	A	B	0.695	1.2	0.53-2.58	C	A	F	0.140	1.7	0.85-3.28	C	A	M	0.232	1.4	0.80-2.54
rs66487711	T	A	B	0.253	3.3	0.42-26.06	T	A	F	0.549	1.3	0.52-3.43	C	A	B	**0.048**	1.7	1.0-2.97
rs2940927	G	R	F	0.106	2.7	0.81-8.99	A	R	F	0.255	0.7	0.32-1.36	G	A	F	0.184	1.3	0.88-1.99
rs1876790	C	A	M	0.330	1.6	0.62-4.1	C	R	B	**0.035**	0.3	0.11-0.92	T	A	F	0.053	1.8	1.0-3.35
rs7732059	G	A	B	0.129	1.8	0.85-3.7	G	A	B	0.421	1.2	0.82-1.61	C	R	F	0.062	3.2	0.9-10.8
rs2972400	A	A	M	0.352	1.6	0.6-4.13	A	R	B	**0.036**	0.3	0.11-0.93	G	A	F	0.053	1.8	1.0-3.35
rs4642376	T	A	M	0.491	1.4	0.54-3.58	T	R	B	**0.035**	0.3	0.11-0.92	G	A	F	0.053	1.8	1.0-3.35
rs1509460	A	R	F	0.106	2.7	0.81-8.99	C	R	F	0.231	0.6	0.31-1.33	A	A	F	0.146	1.4	0.9-2.06
rs13171720	T	R	F	0.253	5.2	0.31-89.05	C	A	F	0.622	1.1	0.67-1.94	T	R	F	0.260	2.7	0.48-15.12
rs13156541	G	A	F	0.106	2.1	0.83-7.02	C	A	F	0.627	1.1	0.71-1.78	G	R	B	0.081	2.1	0.91-4.70
rs11744988	C	A	B	0.220	3.6	0.46-28.53	C	A	B	0.441	1.3	0.71-2.19	T	A	B	**0.047**	1.7	1.01-3.02
rs2972419	A	R	B	0.562	0.5	0.06-4.70	A	R	B	**0.024**	0.2	0.07-0.83	G	A	F	**0.031**	2.0	1.06-3.61
rs2972393	A	R	B	**0.043**	0.3	0.10-0.96	G	A	B	0.288	1.2	0.88-1.53	G	A	F	0.170	1.4	0.88-2.07
rs4509029	G	A	F	0.129	1.8	0.84-3.90	G	R	B	0.298	1.3	0.79-2.12	G	A	F	0.095	1.4	0.94-2.19
rs4129472	A	A	F	0.288	1.9	0.59-6.04	A	A	F	0.233	1.4	0.8-2.46	G	R	F	0.260	2.7	0.48-15.12
rs12153009	A	R	F	0.101	5.6	0.72-43.54	A	R	B	0.229	0.5	0.16-1.54	A	R	B	**0.025**	3.1	1.15-8.15
rs7735889	A	A	B	**0.019**	2.3	1.15-4.61	A	A	F	0.475	1.2	0.76-1.81	G	R	B	0.081	2.1	0.91-4.70
rs12233949	C	A	B	**0.080**	1.7	0.94-3.12	G	A	M	0.054	1.8	1.0-3.12	C	R	B	**0.043**	2.8	1.04-7.50
rs4866941	G	A	B	**0.009**	2.8	1.29-6.09	A	A	M	0.421	1.2	0.77-1.86	A	A	B	0.208	1.2	0.89-1.75
rs4292454	T	A	B	**0.01**	2.3	1.22-4.48	C	A	B	0.383	1.1	0.85-1.51	C	R	M	0.06	2.1	0.95-4.58
rs6873545	T	A	B	**0.0025***	3.3	1.52-6.98	C	R	M	0.424	0.6	0.20-1.98	C	R	B	0.418	1.4	0.65-2.83
rs6886047	A	A	B	**0.0061**	2.9	1.35-6.12	T	R	M	0.239	0.5	0.13-1.67	T	R	B	0.544	1.3	0.60-2.66
rs4273617	A	A	B	**0.0003***	4.8	2.07-11.14	G	A	M	0.304	1.3	0.82-1.92	G	R	B	0.324	1.5	0.69-3.08
rs6180	C	A	B	0.069	1.7	0.96-3.08	A	A	M	0.660	1.1	0.74-1.61	A	A	B	0.189	1.2	0.91-1.65
rs1559286	G	A	M	0.197	2.3	0.64-8.47	G	R	B	0.689	0.6	0.06-6.12	T	A	B	0.255	1.4	0.77-2.62
rs6880056	A	A	B	0.054	1.9	0.99-3.71	A	R	M	0.685	1.5	0.24-8.9	T	A	M	0.192	1.4	0.85-2.25

Table II-2:

[a] Risk allele: major or minor allele.

[b] Genetic model: A additive, R recessive.

[c] Sex: B Both, M Males, F Females.

[d] P values were calculated using logistic regression analysis. Highlighted in bold are tests nominally significant ($p<0.05$) and those with asterisks are tests that remained significant after Bonferroni correction ($p<0.0028$).

[e] OR: Odds Ratio

[f] 95% CI: 95% Confidence Intervals

Table II-3: SNP association analysis of idiopathic short stature vs. pooled adult controls.

	Montreal ISS[a]						Novara ISS[a]						Montreal + Novara ISS[a]					
SNP	Risk allele[b]	Model[c]	Sex[d]	P value[e]	OR[f]	95% CI[g]	Risk allele	Model	Sex	P value	OR	95% CI	Risk allele	Model	Sex	P value	OR	95% CI
rs666581	C	A	B	0.128	4.8	0.64-35.75	A	A	M	0.180	1.8	0.77-4.04	A	A	F	0.299	1.6	0.67-3.61
rs3764451	C	A	B	0.227	1.5	0.77-2.99	C	A	B	0.352	1.2	0.80-1.87	C	A	B	0.216	1.3	0.86-1.92
rs66487711	T	A	B	0.121	4.9	0.66-36.91	C	A	M	0.215	1.7	0.73-3.98	T	A	B	0.555	1.2	0.67-2.11
rs2940927	G	A	B	0.2	1.4	0.84-2.30	A	R	F	0.373	0.7	0.35-1.48	A	R	F	0.373	0.7	0.38-1.44
rs1876790	C	R	B	0.365	0.4	0.05-3.02	C	R	M	0.136	0.3	0.07-01.45	C	R	B	0.100	0.4	0.15-1.18
rs7732059	G	A	B	0.064	1.9	0.96-3.69	G	A	F	0.221	1.4	0.83-2.27	G	A	F	0.097	1.5	0.93-2.40
rs2972400	A	A	F	0.417	1.5	0.58-3.71	A	R	M	0.193	0.4	0.07-1.69	A	A	F	0.167	1.4	0.87-2.31
rs4642376	T	R	B	0.409	0.4	0.05-3.30	T	R	B	0.183	0.5	0.15-01.43	T	R	B	0.137	0.5	0.16-1.28
rs1509460	A	A	B	0.137	1.5	0.89-2.43	C	R	M	0.321	1.4	0.72-2.70	C	R	F	0.373	0.7	0.38-1.44
rs13171720	T	R	F	0.291	3.5	0.34-36.23	T	R	M	0.125	0.2	0.02-1.58	T	R	M	0.132	0.3	0.06-1.44
rs13156541	C	A	B	0.09	1.7	0.92-3.18	C	A	F	0.323	1.3	0.79-2.05	C	A	F	0.161	1.4	0.88-2.17
rs11744988	C	A	B	0.119	5.0	0.66-37.13	C	A	F	0.228	1.7	0.71-4.11	C	A	F	0.159	1.8	0.79-4.18
rs2972419	A	R	B	0.361	0.4	0.05-3.0	A	R	M	0.074	0.2	0.02-01.2	A	R	B	0.056	0.3	0.11-1.03
rs2972393	A	R	M	0.068	0.1	0.02-1.15	A	R	M	0.501	1.3	0.65-2.40	G	A	B	0.486	1.1	0.84-1.43
rs4509029	G	A	B	0.118	1.5	0.90-2.47	G	R	M	0.466	1.3	0.65-2.58	G	R	M	0.292	1.4	0.75-2.66
rs4129472	A	A	F	0.363	1.7	0.55-5.21	A	A	F	0.233	1.3	0.91-1.96	A	A	F	0.119	1.5	0.90-2.57
rs12153009	A	R	B	**0.001***	7.1	2.22-22.82	G	A	M	0.255	1.4	0.79-2.39	A	R	B	0.119	2.2	0.82-5.90
rs7735889	A	A	B	**0.023**	2.2	1.11-4.16	G	R	F	0.282	1.7	0.66-4.15	A	A	F	0.389	1.2	0.79-1.83
rs12233949	C	A	B	**0.029**	1.9	1.07-3.40	G	A	M	**0.013**	2.1	1.17-3.76	G	A	M	0.162	1.4	0.87-2.29
rs4866941	G	A	B	**0.009**	2.7	1.28-5.54	A	A	M	0.164	1.4	0.87-2.25	G	A	F	0.464	1.2	0.76-1.81
rs4292454	T	A	B	**0.015**	2.1	1.16-3.85	C	R	B	**0.007**	2.0	1.21-3.25	C	R	B	**0.027**	1.7	1.06-2.75
rs6873545	T	A	B	**0.014**	2.3	1.19-4.47	C	R	B	0.408	0.7	0.32-1.59	C	R	B	0.157	0.6	0.25-1.25
rs6886047	A	A	B	**0.022**	2.2	1.12-4.15	T	R	M	0.261	0.5	0.12-1.76	T	R	B	0.101	0.5	0.22-1.15
rs4273617	A	A	B	**0.001***	3.2	1.57-6.66	G	R	B	0.381	0.7	0.30-1.59	G	R	B	0.150	0.5	0.23-1.25
rs6180	C	A	F	0.118	2.0	0.84-4.80	C	A	M	0.709	1.1	0.71-1.64	C	R	B	0.551	1.1	0.73-1.80
rs1559286	G	A	M	0.106	2.3	0.84-6.29	G	A	M	0.628	1.2	0.53-2.84	G	A	M	0.303	1.5	0.71-3.03
rs6880056	T	A	B	0.089	1.7	0.93-2.96	A	A	B	0.656	1.1	0.74-1.60	T	R	B	0.446	1.5	0.52-4.41

Table II-3:

[a] For this analysis of the ISS groups separately as well as combined, the controls were a pool of Montreal adult controls with restricted height (n=57) and the CARTaGENE adult controls (n=207) (30^{th}-70^{th} percentile).

[b] Risk allele: major or minor allele.

[c] Genetic model: A additive, R recessive.

[d] Sex: B Both, M Males, F Females.

[e] P values were calculated using logistic regression analysis. Highlighted in bold are tests nominally significant ($p<0.05$) and those with asterisks are tests that remained significant after Bonferroni correction ($p<0.0028$).

[f]OR: Odds Ratio and [g]

95%CI:95%ConfidenceIntervals.

FIGURE LEGENDS

Figure II-1: GT microsatellite localization in the *GHR* gene. (A) The *GHR* gene spans ~300kb on the short arm of chromosome 5 close to the centromere. **(B)** Fourteen different *GHR* mRNAs encoding the full-length GHR have been reported to date (not all first exons are shown here). The coding region is defined by exons 2-10 where exon 2 contains the translation start site. **(C)** Three of the exons transcribing ubiquitous *GHR* mRNAs (V2, V9, V3) are located within a 1.6kb region of the 5'UTR. **(D)** A GT repeat lies within the V9 promoter and is polymorphic in the general population (Hadjiyannakis et al. 2001).

Figure II-2: (GT)n satellite polymorphism allele distribution and genotype frequency in three different cohorts. (A-C) Montreal ISS, **(D-F)** Novara ISS and **(G-I)** CARTaGENE SS *vs.* normal height controls for each cohort. All cohorts displayed similar allelic distribution profiles with a median of 26 GT repeats. The Montreal ISS children showed a 3.3-fold increase in L/S genotypes compared to their adult controls (Fisher exact test: p = 0.04). The Novara ISS cohort had 1.5-fold more M/S genotype carriers than their controls (Fisher exact test: p=0.053).

Figure II-3: Association plot results (additive model) with short stature and haploblock structure of the *GHR* gene region in each cohort. Upper panels were generated using Locus Zoom (Pruim et al. 2010), the 1000 Genomes LD European population (Nov 2014) and the hg19 genome build as references. Association results from logistic regression analysis (additive model) are shown as $-\log_{10}$ p values on the left y-axis for the 27 SNPs genotyped. The most significant SNP is indicated as a purple diamond and the others are color-coded according to the strength of their LD relationship as measured by r^2. Local recombination rates are shown (blue line, scale on the right y-axis). The intron–exon structures of local genes are provided below with direction of transcription and genomic coordinates in Mb. **(A)** In the Montreal ISS cohort, rs4273617 in intron 5 was the most significant ($p=3x10^{-4}$). **(B)** For the Novara cohort, rs12233949 in intron 2 was found to be the most significant (p=0.075). **(C)** For the CARTaGENE cohort, rs666581 >200kb upstream of V2 was nominally significant (p=0.0237). **Lower panels**: LD plots showing haplotype blocks for each of the populations. The heat map, based on D' values, was drawn using the SNP panel genotyped data with the Haploview software v4.2 using the solid spine algorithm.

Figure II-4: Haplotype analysis at the *GHR* locus. The **upper panels** represent the haplotype block structure in each of the populations studied. The haplotype blocks were determined using the solid spine algorithm from the Haploview software v4.2. Major and minor alleles are represented by blue and red squares, respectively. Frequency of each haplotype in cases and in controls are shown to the right of each haplotype and the D' value (indicating the level of linkage disequilibrium between two blocks) is also provided. Connections from one block to the next are shown for haplotypes of >10% frequency with thick lines and >1% frequency with thin lines. Each cohort displayed a specific haplotype block architecture in the *GHR* gene region, from 5 blocks in **(A)** Montreal ISS and **(B)** Novara ISS to 4 blocks in **(C)** CARTaGENE SS. The **lower panels** show p value, OR (odds ratio) and 95% CI results from haplotype logistic regression analyses. Frequencies (F) of each haplotype in cases and in controls are also provided. In the Montreal ISS cohort **(A)**, association of a common risk haplotype, TTA in block 4 (OR=2.5, p=0.0043), remained significant after correction for multiple testing (*p<0.05 after permutation test).

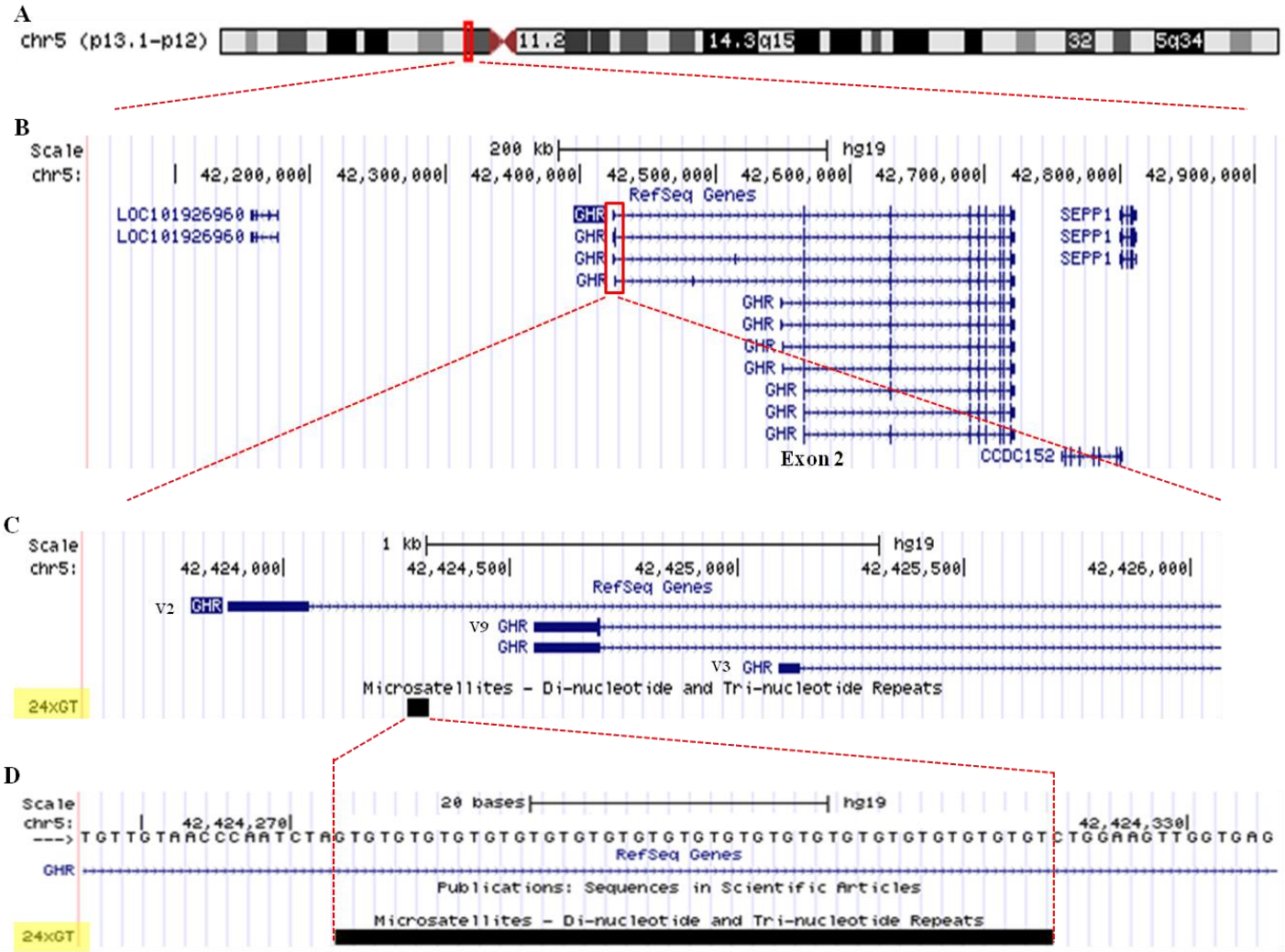


Figure II-1: GT microsatellite localization in the *GHR* gene.

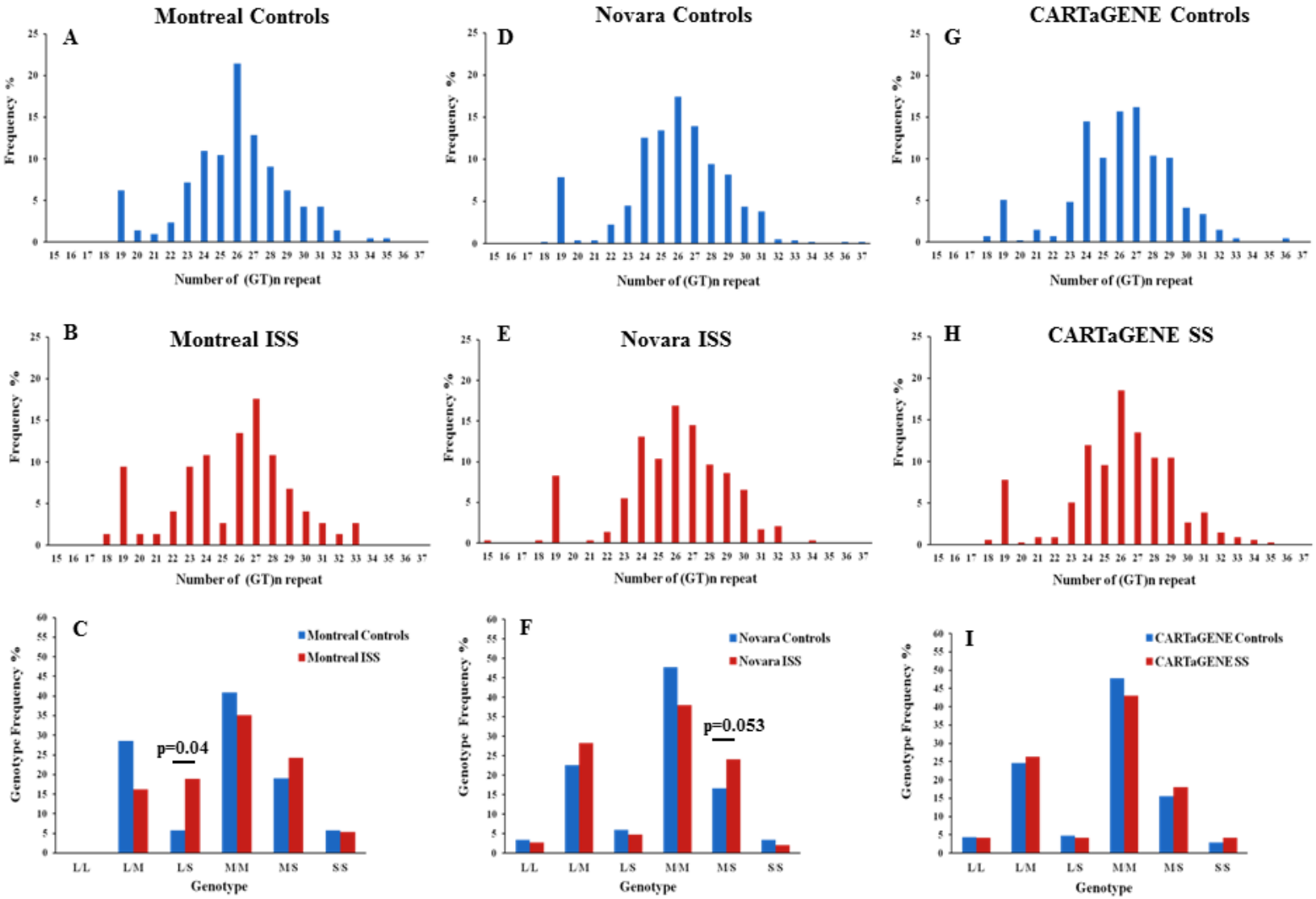


Figure II-2: (GT)n satellite polymorphism allele distribution and genotype frequency in three different cohorts.

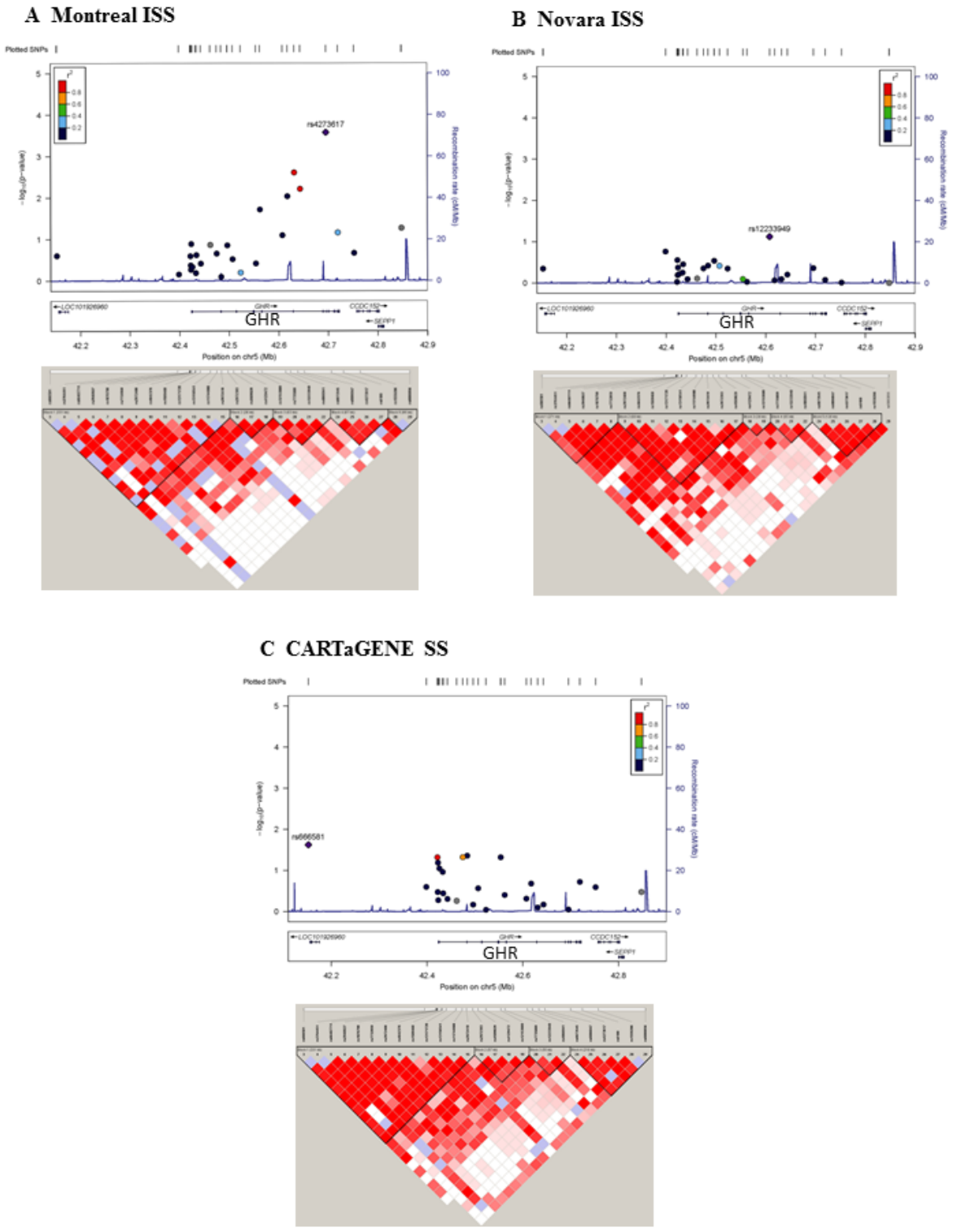
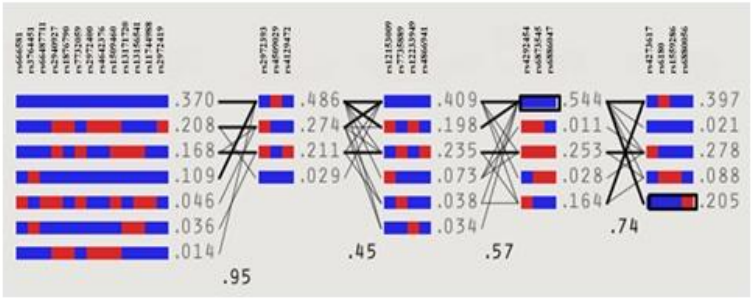


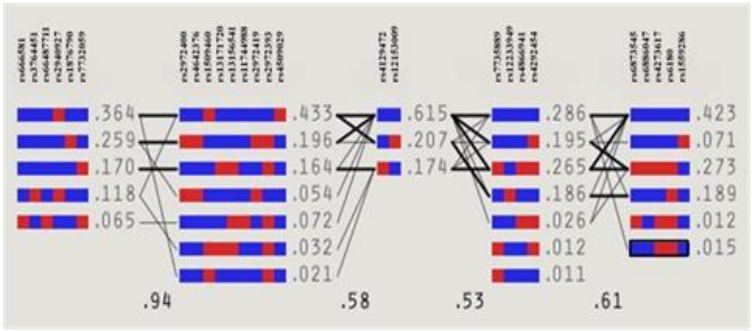
Figure II-3: Association plot results (additive model) with short stature and haploblock structure of the *GHR* gene region in each cohort.

A Montreal ISS



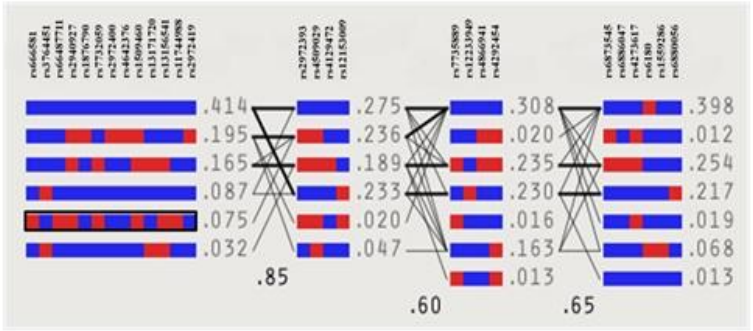
Block	Haplotype	F-cases	F-controls	OR	P-value	95% CI
4	**TTA**	0.688	0.493	2.5	**0.0043***	1.33-4.72
5	AATT	0.288	0.179	2.0	0.043	1.02-3.91

B Novara ISS



Block	Haplotype	F-cases	F-controls	OR	P-value	95% CI
5	TAGAT	0.028	0.009	3.4	0.038	1.07-11.04

C CARTaGENE SS



Block	Haplotype	F-cases	F-controls	OR	P-value	95% CI
1	ATCATCGGCCGTG	0.099	0.060	1.8	0.046	1.01-3.13

Figure II-4: Haplotype analysis at the GHR locus.

Table II-S1: Relative frequencies (%) for each GT allele category for each cohort

	Controls			Cases		
	S	**M**	**L**	**S**	**M**	**L**
Montreal	18	65	17	27	55	18
Novara	15	67	18	17	64	19
CARTaGENE	13	68	19	15	65	20

There were no significant differences in the frequencies of each category (S, M, L) amongst the three populations.

Table II-S2: Number of individuals per cohort and GT genotype

	Montreal		Montreal (filtered)	Novara *		CARTaGENE **	
GT genotype	Cases	Controls	Controls	Cases	Controls	Cases	Controls
L/L	0	0	0	4	10	7	9
L/M	6	30	18	41	65	44	51
L/S	7	6	2	7	17	7	10
M/M	13	43	18	55	136	72	99
M/S	9	20	13	36	48	30	32
S/S	2	6	6	2	11	7	6
Total	**37**	**105**	**57**	**145**	**287**	**167**	**207**

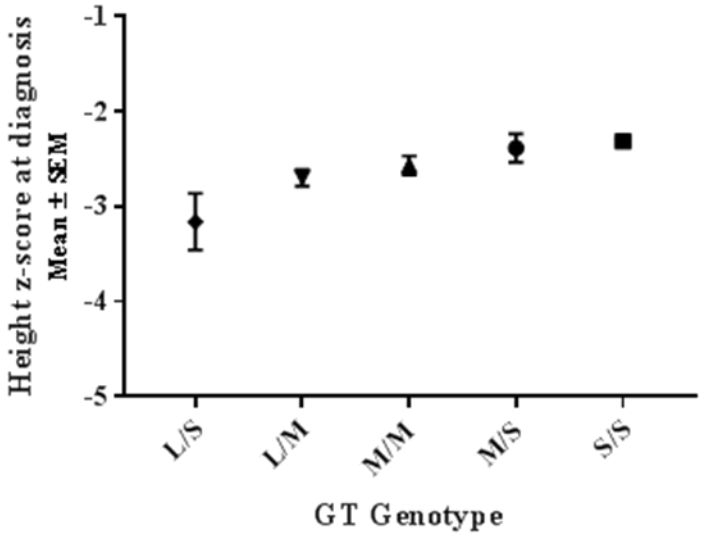
* For the Novara cohort, 2 cases and 5 controls have been excluded from the SNP analysis (see Table 1).
** For the CARTaGENE cohort, 1 case failed GT genotyping.

Table II-S3: SNP panel information

rs ID dbSNP147	Major:Minor Allele (1000 Genomes)	Position in GRCh37 hg19 assembly	MAF (1000 Genomes browser) European	Location in the *GHR* gene region
rs666581	C:A	42152335	A: 0.062	> 200kb upstream V2
rs3764451	T:C	42398465	C: 0.16	> 25kb upstream V2
rs66487711	T:C	42421689	C: 0.07	proximal V2 promoter
rs2940927	G:A	42422389	A: 0.48	proximal V2 promoter
rs1876790	T:C	42422855	C: 0.19	core V2 promoter
rs7732059	G:C	42423472	C: 0.29	core V2 promoter
rs2972400	G:A	42425830	A: 0.19	downstream of V3
rs4642376	G:T	42432739	T: 0.19	intron 1
rs1509460	A:C	42433819	C: 0.48	intron 1
rs13171720	C:T	42442715	T: 0.24	intron 1
rs13156541 *	C:G	42461964	G: 0.34	intron 1
rs11744988	C:T	42474696	T: 0.07	intron 1
rs2972419	G:A	42483877	A: 0.18	intron 1 (associated with short stature in Pygmy population)
rs2972393	G:A	42496257	A: 0.46	intron 1
rs4509029	G:A	42506775	A: 0.5	intron 1
rs4129472	A:G	42523172	G: 0.22	intron 1
rs12153009	G:A	42553879	A: 0.29	intron 1
rs7735889	A:G	42562199	G: 0.31	proximal V5 promoter
rs12233949	G:C	42607461	C: 0.27	intron 2
rs4866941	G:A	42617492	A: 0.30	intron 2
rs4292454	T:C	42623245	C: 0.46	intron 2
rs6873545	T:C	42631264	C: 0.31	intron 3 (used as a tag to genotype the *GHR* d3/fl polymorphism)
rs6886047	A:T	42643200	T: 0.30	intron 3
rs4273617	A:G	42695471	G: 0.32	intron 5
rs6180	A:C	42719239	C: 0.42	exon 10
rs1559286	T:G	42752037	G: 0.05	~ 25kb downstream of *GHR*
rs6880056 *	A:T	42847750	T: 0.22	~120 kb downstream of *GHR*

* CEU population (Utah residents with Northern and Western European ancestry from the CEPH collection)

(A)



(B)

	L/S	L/M	M/M	M/S	S/S
Age at enrolment (years):	9.4-16	6.8-21.5	6.9-21	8.4-20.8	13.8,16.7
% of individuals ≥3rd percentile:	0% (n=0/7)	16.7% (n=1/6)	38.5% (n=5/13)	66.7% (n=6/9)	100% (n=2/2)

Figure II-S1: GT genotype association with height z-score in the Montreal ISS cohort.

(A) At diagnosis, the association between the GT genotypes and the mean height z-score showed a significant difference between the L/S and the M/S genotypes (one-way Anova with Tukey correction: p=0.026) and between the L/S vs non-L/S genotypes (two-tailed unpaired t-test: p=0.0026**).** (SEM: standard error of the mean)

(B) At enrolment**,** the percentage of individuals showing catch-up growth by reaching the 3rd percentile. None of the L/S genotype carriers reached the 3rd percentile while carriers of the L allele (35% of the Montreal ISS cohort) showed the fewest number of individuals with height catch-up.

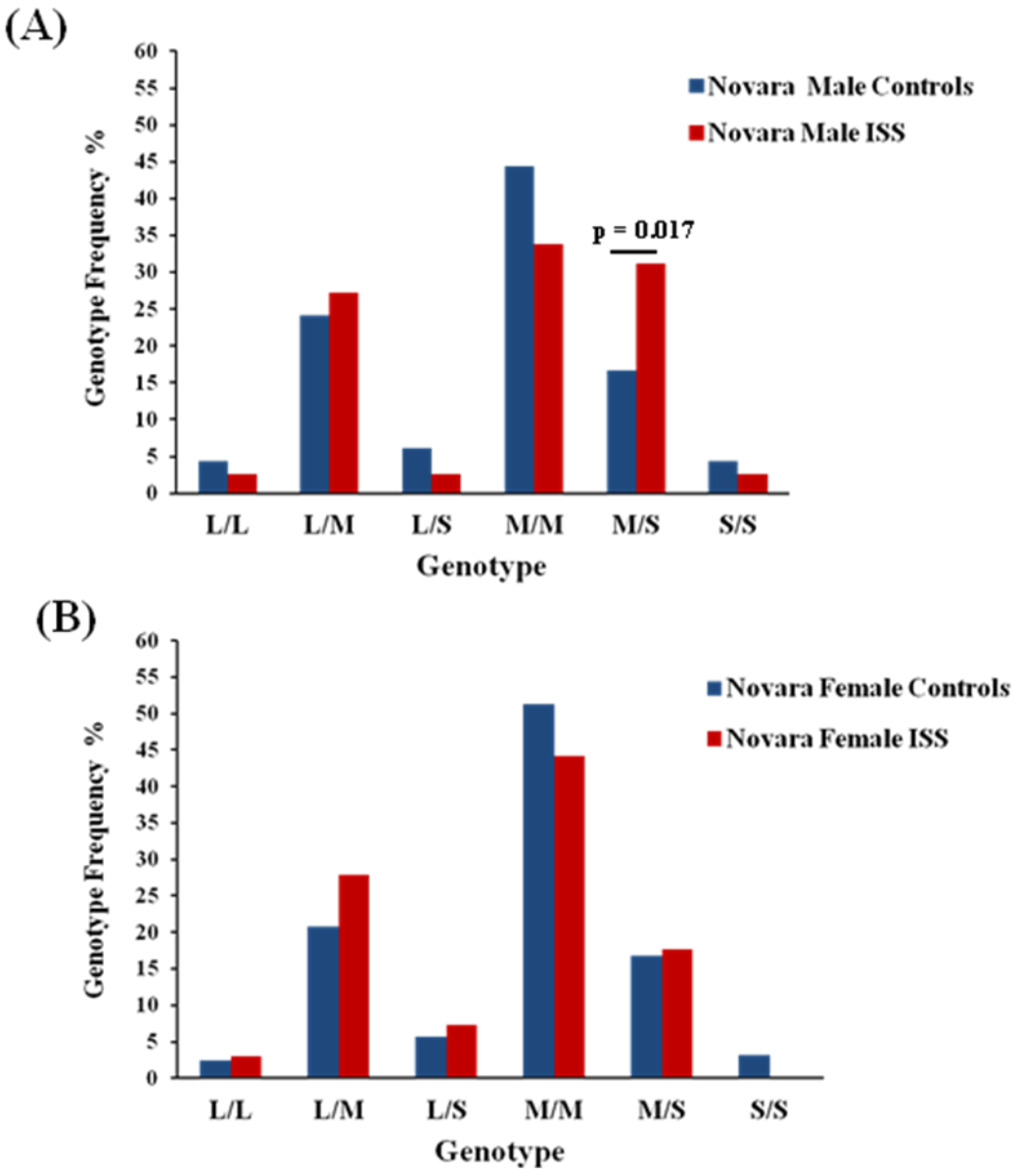


Figure II-S2: (GT)n microsatellite polymorphism allele genotype frequency in the Novara cohort by gender. (A) The Novara male ISS cohort has 1.9-fold more carriers of the M/S genotype compared to the normal height male controls (Fisher exact test: p=0.017), a difference not observed in the female cohort **(B)**.

CHAPTER III

It is well known that GH acts not only in the acquisition of longitudinal height but also in fat metabolism. Individuals with GHI can present with a variable degree of obesity depending on their endocrine profile. To investigate whether an increased obesity is observed in a cohort of short stature adults, I performed a comparative analysis of different adiposity indices with short stature *vs* normal height individuals and further investigated if obesity was a confounding factor in the association of *GHR* gene polymorphisms with short stature.

Sex-specific increase in the risk of obesity with short stature: insights from a Canadian cohort

Christel Dias[1], Agnihotram V Ramanakumar[2] and *Cynthia Gates Goodyer[1,3]

[1]Division of Experimental Medicine, McGill University, Montreal, Quebec, Canada; [2]Centre for Innovative Medicine Biostatistics, Research Institute of McGill University Health Centre, Montreal, Quebec, Canada; [3]Department of Pediatrics, McGill University, Montreal, Quebec, Canada

Author contributions: CD was responsible for data collection and analysis as well as writing of the manuscript as part of her Ph.D. program; AVR provided critical advice for the statistical analyses; CGG was the principal investigator of the project, reviewer of the data and analyses; all authors reviewed the manuscript.

***Corresponding author**: Dr. Cynthia Gates Goodyer, Research Institute of McGill University Health Centre, Block E, EM0.3211, 1001 Decarie Boulevard, Montreal, Quebec, Canada H4A 3J1; Telephone: 1 (514) 934-1934, ext. 22481, Fax: 1 (514) 933-4149; Email: cindy.goodyer@muhc.mcgill.ca

Short Title: Obesity, short stature and sex-specificity

Key Words: Obesity, short stature, sex-specific, *GHR*, microsatellite polymorphism

1. ABSTRACT (240 words)

Human Growth Hormone (GH) and its receptor (GHR) have significant effects on both the acquisition of height and fat metabolism, but little is known about their role at the extreme tails of height distribution. In the present study, we investigated the associations of polymorphisms (SNPs, microsatellite) in *GHR* and other growth-related genes (*HMGA2*, *GDF5, IGF1, IGF1R* and *IGFBP3*) with obesity in severe short stature individuals (<1st percentile). Using a Canadian cohort, we compared severe short stature males and females with normal height controls. Surprisingly, the short stature females exhibited significantly higher levels of several anthropometric traits of obesity (e.g. body mass index [BMI]; waist or hip circumference/height) than their male counterparts or controls ($p<0.0002$). Six SNPs located in the *GHR* promoter were more strongly associated with height following adjustment with BMI, fat mass index (FMI) or body adiposity index (BAI) only in the short stature women. In addition, the M/L genotype of the *GHR* GT repeat microsatellite polymorphism was significantly associated with almost all adiposity indices primarily in the short stature females ($p<0.009$-0.0001). Finally, *HMGA2* and *GDF5* SNPs showed significant associations only with the short stature women, with FMI the strongest confounding factor ($p<0.003$ and 0.009, respectively). In summary, our investigations revealed an increased risk of obesity in severe short stature women. In addition, we found that three major growth-related genes, *GHR*, *HMGA2* and *GDF5*, share associations with both extreme short stature and obesity-linked anthropometric traits in a sex-specific manner.

2. INTRODUCTION

The increasing prevalence of obesity is a global health challenge. The dramatic rise has been attributed to more sedentary lifestyles combined with changes in food consumption towards high-fat, energy-dense diets (Caballero 2007). However, obesity, like stature, is a complex trait with a strong genetic component. In 2007, the discovery of FTO (fat mass and obesity-associated) as a major obesity-associated gene paved the way for several other genome wide association studies that, to date, have identified >150 loci associated with obesity-related phenotypes (Sandholt et al. 2015).

Several studies have investigated the prevalence of obesity across height percentiles in different populations and showed similar results (Bosy-Westphal et al. 2009; Hermanussen et al. 2005; Lara-Esqueda et al. 2004): for example, in 2009, Bosy-Westphal et al. reported that, in a large cohort of German adults, the frequency of obesity (body mass index [BMI] >30 kg/m^2) gradually increased with decreasing height in both sexes (Bosy-Westphal et al. 2009). Given these findings, it is surprising that little attention has been paid to Growth Hormone [GH] and its potential role in the short stature link with obesity: GH is essential for normal musculoskeletal development but also has important regulatory effects on lipid metabolism at all stages of life.

GH functions by binding to its high-affinity receptor (GHR) on target cells and activating intracellular signaling pathways that lead to changes in gene expression. Individuals with low GH or GHR levels or a dysfunctional GHR are not only short, they have increased adiposity and a greater risk of lipid disorders and cardiovascular disease (Savage et al. 2006). Conversely, individuals suffering from GH excess have reduced fat depots (Bengtsson et al. 1989; Reyes-Vidal et al. 2015). Clinical and experimental studies have clearly demonstrated that GH is a major regulator of adipocyte proliferation, differentiation and function *in utero* as well as postnatally, with an important effect on lipolysis. GHR are present on both preadipocytes and mature adipocytes, and a significant up-regulation of *GHR* gene expression occurs during adipocyte differentiation (Nam and Lobie 2000; Wabitsch et al. 2001; Wei et al. 2006). GH- or GHR-deficient newborns have well-developed fat depots and they progress to obesity with age (Laron 2004a). Finally, GH treatments significantly reduce body fat in GH-deficient individuals (Beauregard et al. 2008; Boot et al. 1997; Kamel et al. 1995). Thus, GH and its receptor are closely linked to both human height and adiposity.

However, previous population studies have not investigated these pleiotropic relationships in individuals at the extreme tails of height distribution. In the present study, we examined severe short stature, using a Canadian cohort to compare adiposity in adults below the 1st percentile on the growth curve with normal height (30th – 70th percentile) adults. We first explored the relevance of BMI for detecting obesity in these two groups, as BMI has been used extensively as a marker of the obese state. However, because questions have been raised about the ability of BMI to predict obesity at the extreme tails of the height distribution (Lara-Esqueda et al. 2004), we also investigated the efficacy of several other adiposity indices. In addition, we studied a previously unstudied genetic component of obesity in short stature, looking at the potential effect of BMI and alternative adiposity indices on the association between short stature and polymorphisms of the *GHR* as well as other height-related genes (e.g. high mobility group AT-hook 2 [*HMGA2*]).

3. MATERIALS AND METHODS

A. The CARTaGENE Cohort.

The CARTaGENE Platform is the largest population-based cohort in Quebec, Canada, with ~20,000 recruits aged 40–69 years at the time our study was initiated; the participants represent a random selection of individuals residing in the metropolitan areas of Quebec (Awadalla et al. 2013) (https://cartagene.qc.ca/). The Research Ethics Committee of CHU Ste-Justine approved our access to the bio- and data-banks. Information on the recruits was anonymized prior to our receiving the DNA samples and anthropometric data for analysis. Our cohort consisted of 168 short stature individuals with a final height corresponding to severe short stature (<1st percentile: males: maximum height of 159 cm, −2.4 SDS; females: maximum height of 147 cm, −2.47 SDS) and 207 controls of average height (30th–70thpercentile [~±0.5 SDS]: males were 173−180 cm; females were 160–167 cm) (**Table III-1**). SDS scores and percentiles were calculated according to the World Health Organization (WHO) Growth Charts for Canada (2014 revision). The majority (>95%) of the participants were from European ancestry. Asian, African and South American individuals were excluded to minimize stratification of the genetic results.

B. Anthropometric Variables.

Standing height was measured in centimeters (cm) with a portable stadiometer (SECA 214). Body weight (kg), fat mass (FM, in kg), body fat % (BF%) and fat free mass (FFM, in kg) were obtained using bioelectrical impedance technology (TBF-310, TANITA). FM and FFM indexes were calculated: FMI = FM(kg)/height(m)^2 and FFMI = FFM(kg)/height(m)^2. Waist and hip circumferences (WC and HC, in cm) were an average of two measurements (SECA 200 circumference measuring tape). BMI was calculated as weight(kg)/height(m)^2 while Body Adiposity Index (BAI) was calculated using the following formula: [HC(cm)/ (height(m)*√height(m)] - 18.

C. Microsatellite Genotyping.

Detailed methodology was previously described (Dias et al. 2017). Briefly, the GT microsatellite from the *GHR* V9 promoter region (chr5:42424274-42424321 hg19 Genome Assembly) was genotyped using fluorogenic probes followed by capillary electrophoresis (ABI Genetic Analyzer 3730 XL; Life Technologies, Foster City, CA) to discriminate allele size (ABI GeneMapper Version 4.1). Primers were designed to amplify a 155bp fragment containing a 24 GT repeat that was used as reference; deviation from this size allowed us to deduce the GT length of the different alleles. We then classified the alleles arbitrarily into three categories in order that the cutoffs fall at ~±1S.D. around the median. Thus, the shortest (S) alleles were <24 repeats (~16% of the individuals), the medium (M) alleles were 24–28 repeats (~68%), and the longest (L) alleles were >28 repeats (~16%)**.** We then attributed one of the six possible genotypes to each individual: S/S, S/M, M/M, M/L, S/L or L/L.

D. Single nucleotide polymorphism (SNP) Selection and Genotyping.

At the GHR locus: Twenty-seven SNPs with minor allele frequencies >5% were selected to span the *GHR* gene region from ~200 kb upstream of the major *GHR* (V2) transcriptional start site to ~120 kb downstream of the 3′UTR (**Supplemental Table III-1**).

At height-related gene loci: A set of five growth-related genes were chosen: rs1042725 (HMGA2) (Weedon et al. 2007), rs224333 (growth differentiation factor 5 [GDF5]) (Shungin et al. 2015), rs4969035 (insulin-like growth factor 1 receptor [IGF1R])

(Haataja et al. 2011), rs1019731(IGF1) (Lettre et al. 2007) and rs2854744 (IGF binding protein 3 [IGFBP3]) (Patel et al. 2008) (**Supplemental Table III-2**).

Genotyping: Genotyping was performed at Génome Québec (Montreal, QC) using Sequenom iPLEX Gold Technology and the MassARRAY system (Agena Biosciences, San Diego, CA). To assess the robustness of the technology, ~20% of the total samples were replicated with 100% success rate. Quality controls were conducted prior to the analysis: all variants used for association analysis had a genotyping efficiency call rate >95% and showed no departure from Hardy–Weinberg equilibrium in combined controls and cases ($p > 0.001$) and in controls and cases separately. A table of the SNP probes is available on request.

E. Statistical Analysis.

Demographic and anthropometric data differences between groups were analyzed using the two-tailed Mann-Whitney test. Differences in frequencies of individuals were calculated using a Fisher exact test. A $p<0.05$ for combined sex and a $p<0.025$ after stratification were considered significant (Tables 1-3). Quantitative traits normalized for height were compared between the short and normal stature groups for each GT genotype using the two-tailed Mann-Whitney test (Table 4). Bonferroni corrections were applied to account for the multiple testing of the GT genotype categories and the stratification by sex (pcorr<0.00083 [0.05/60]); $p<0.05$ was considered nominally significant. Analyses were conducted with the GraphPad Prism v7.0 software (La Jolla, CA).

For the case-control SNP binary logistic regressions (Tables III-5 and III-6), we used the PLINK v1.07 software package and the SPSS statistical software v22.0 (Chicago, IL). Results of unadjusted and adjusted logistic regressions in the additive or recessive model were obtained for both sexes combined and for each gender separately. For the adjusted regressions, BMI, BAI or FMI were added as covariates; age was adjusted for all. The β regression coefficients of the different covariates are shown in Supplemental Table III-3. Because of strong LD between certain SNPs in the *GHR* locus, 18 SNPs represented the effective number of independent SNPs to use in the Bonferroni correction for multiple comparisons ($p_{corr}<0.0028$ [0.05/18]) (Table III-5). A Bonferroni correction was also applied for the height-related genes: $p_{corr}<0.01$ (0.05/5) was considered significant (Table III-6).

4. RESULTS

A. Demographics.

The total number of individuals in our CARTaGENE cohort was n=375, 46.6% of which were female. The case group was composed of n=168 short stature adults with heights below the 1st percentile of the growth curve (median SDS: -2.68). The control group had n=207 normal height adults (30th-70th percentiles; median SDS: 0.01) **(Table III-1)**. Their ages ranged from 40-70 years, with similar median ages for the males and females; 82.5 % of the men and 80.6% of the women were >50 years old.

B. Anthropometric Measures.

We first examined a number of anthropometric variables to determine the differences between our short stature and control height groups as well as any sex-related differences. Body weights were significantly lower in the short stature males and females compared to their normal height controls ($p<0.0001$ and $p=0.0004$, respectively). Interestingly, this translated into a significantly higher BMI in the short stature females compared to their controls ($p=0.0002$), whereas no differences were observed between the male groups **(Table III-2)**. The short stature females still had a significantly higher BMI than controls after removing the women presenting with morbid obesity (BMI≥40; $p=0.0045$). When the BMI data were stratified into three categories according to the WHO definition, the % of short stature women in the non-obese (BMI<25 kg/m^2) category was significantly lower compared to female controls (Fisher exact test: OR=3.4, 95% CI [1.67-7.03]; $p=0.0006$), with a significant shift towards the obese class (Fisher exact test: OR=1.64, 95% CI [1.1-2.47]; $p=0.0247$) (**Table III-3a**). In contrast, no significant differences were seen for the male groups. It should be noted that the majority of individuals in the cohort were in the overweight and obese BMI categories.

Analysis of the waist circumference (WC) and hip circumference (HC) data showed significantly lower medians for men with short stature compared to their male controls ($p<0.0001$) while there was no difference for the waist/hip ratios (WHR) (**Table III-2**). These data suggest proportional changes in the accumulation of fat in the short stature male. This outcome is supported by the WC/height and HC/height data: the short stature men have

slightly but significantly higher waist ($p<0.037$) and hip ($p<0.024$) circumferences when their heights are taken into account. In contrast, the short stature and control females had similar medians for WC, HC and WHR, indicating much higher levels of both central and gluteal adiposity in the short stature women. Again, this is corroborated by the WC/height and HC/height data: the short stature females have markedly higher ratios than their controls ($p<0.0001$).

WC and WHR are well-recognized measures of abdominal obesity. The WHO recommends the following obesity risk cut-offs for Caucasian populations: for males a WC≥102 cm and a WHR≥1; and for females a WC≥88 cm and a WHR≥0.85 (2008). In our cohort, more than 50% of both the short stature and control women were above these cut-offs for both WC and WHR (**Tables III-3b and III-3c**). Surprisingly, when we compared the male groups, there were almost twice as many control males above the WC cutoff compared to the short stature males (Fisher exact test: OR=2.67, 95% CI [1.47-4.89]; p=0.0019), while there was a similar proportion (~50%) of short stature and control males classified ≥1 for the WHR data. More than 50% of the short stature women were above the WC cut-off compared to ~25% of the short stature men (Fisher exact test: OR=3.68, 95% CI [1.9-7.1]; $p<0.0001$). Similar observations were noted for the WHR data: nearly 70% of the short stature women were above the cut-off compared to 50% of the short stature men (Fisher exact test: OR=2.27, 95% CI [1.2-4.27]; p=0.016). In contrast, there were no significant differences between the male and female control groups.

The Body Adiposity Index (BAI), based on hip circumference and height, has been used as an alternative to BMI (Bergman et al. 2011). Our short stature males and females both showed a significant increase in BAI compared to their respective controls ($p<0.0001$), with a much larger increase for the women (**Table III-2**). BAI is a measure of adiposity normalized for height while body fat % (BF%) is not; therefore we were not able to compare these two indices in the short stature groups. For the normal stature groups, BAI and BF% gave similar medians of overall adiposity in the men while, in the women, BAI seemed to underestimate the BF%. However, BAI and BF% correlated in both control groups ($r^2=0.77$ in men and $r^2=0.89$ in women; Spearman test).

Measurements of fat mass (FM) and fat-free mass (FFM) were significantly lower in the two short stature groups compared to their respective controls and this was more important

in the males (**Table III-2**). In contrast, the BF% data did not show any significant differences. There were also no differences in the percentage of individuals in the subgroups when the BF% data were stratified for above the cut-off for obesity (≥29% for men and ≥41% for women, aged 40-70) (**Table III-3d**) (Gallagher et al. 2000).

It has been suggested in previous studies that adiposity measurements, like FM and FFM, do not adjust appropriately for body size unless stature is taken into account (Heymsfield et al. 2011; VanItallie et al. 1990). Therefore, we divided the FM and FFM data by the square of height to obtain FM index (FMI) and FFM index (FFMI) results (**Table III-2**). The FMI medians were not significantly different between the groups. However, by setting a cut-off for obesity above 9 kg/m^2 for men and 13 kg/m^2 for women (Kelly et al. 2009), we found that, for the females above this cut-off (25-30% of our cohort), short stature individuals had an FMI median significantly higher than average height controls (17.1, 95% CI [13.3, 29.6] *vs.* 15.35, 95% CI [13.1, 20]; p=0.0055); in contrast, there was no significant difference for the males. The FFMI was increased significantly in short stature women compared to female controls (p<0.0001) but, again, there were no differences for the men.

These data demonstrate major sex-related differences in our short stature cohort, with the females exhibiting significantly stronger associations with most anthropometric traits of obesity.

C. Association of the *GHR* Gene GT Microsatellite Polymorphism with Short Stature, Different Anthropometric Traits Adjusted for Height, and Sex.

To assess the potential associations of the *GHR* gene with short stature and different anthropometric measures, we first conducted a series of case-control studies comparing different adiposity indices relative to the individual's GT microsatellite genotype and sex. Because of the stature differences in our case *vs.* control groups, each of the six GT genotypes was examined with anthropometric indices adjusted for height using the Mann-Whitney test (**Table III-4**). A higher BMI was associated with short stature females *vs.* controls presenting with the M/L (p=0.0087) or S/M (p=0.046) genotypes; no associations were noted for the males. Interestingly, both short stature females and males presenting with the M/L genotype had higher WC/Height and HC/Height ratios compared to their respective controls, with a highly significant HC/Height increase for the females (p=0.0003). Several of the GT genotype categories showed a significant association with a higher BAI, including M/L, M/M and S/M,

for both the short stature females and males; the S/M genotype showed a stronger association with BAI in the women (p=0.0003) than the men (p=0.01). The FMI data were negative for each GT genotype and the two sexes; however, by setting a cut-off for obesity above 13 kg/m^2 (Kelly et al. 2009), short stature females with the M/M genotype showed a higher FMI median than their controls (17.1, 95% CI [14.3-21.2] *vs.* 14.1, 95% CI [13.6-16.2]; p=0.034). Finally, short stature women with the M/L, the M/M and, to a lesser degree, the S/M genotype, were associated with a significantly higher FFMI. In contrast, no significant associations were observed for the male groups.

In summary, we observed a significant association of the M/L *GHR* GT genotype with almost all of the anthropometric traits examined, primarily in short stature women.

D. Association of SNPs at the *GHR* Locus with Short Stature, Indices of Adiposity, and Sex.

We previously tested the association of 27 common variants (MAF>5%) in the *GHR* gene region with short stature in three different cohorts (Dias et al. 2017). In the present study, we conducted a series of logistic regression analyses of the same SNP set in the CARTaGENE cohort to examine possible associations following adjustment with three adiposity indices: BMI, BAI and FMI. **Table III-5** shows only results for which the associations had at least a nominal significance.

Of the 27 SNPs, six showed a significant change in the odds ratios after adjustment for BMI, selectively in the female group: rs666581, rs7732059, rs1876790, rs2972400, rs13156541 and rs2972419. Interestingly, the first five are located in the same linkage disequilibrium (LD) block of the *GHR* gene that includes distal, proximal and core regulatory elements, a small cluster of three first exons transcribing the most abundant and ubiquitous *GHR* mRNAs (V2,V9,V3), and part of intron 1 (**Supplemental Table III-1**). The sixth SNP, rs2972419, is located further downstream in intron 1 and has been associated with pygmy short stature (Becker et al. 2013). Parallel findings were obtained after adjustment for FMI with most of the SNPs, except for the intron 1 SNP, rs12153009, for which the effect of FMI was greater than for BMI, again in the females. BAI had no association with distal and proximal promoter SNPs but strong ones with core promoter SNP rs1876790 (p=0.007), intron 1 SNP rs2972400 (p=0.007) and the pygmy SNP rs2972419 (p=0.0049), all three

showing increased protective odds only in the females. For the males, 2 SNPs showed a trend of association with short stature that did not reach significance (rs66487711, rs11744988); BMI, BAI and FMI had little effect. Within our initial SNP set, rs6873545 was used to tag a common genomic deletion that signifies loss of exon 3 from the *GHR* gene (Lettre et al. 2007); no association was detected with either short stature or obesity traits in our cohort.

Thus, three different adiposity indices were associated with SNPs located primarily in regulatory regions of the *GHR* gene selectively in the short stature females.

E. Association of SNPs in Height-related Genes with Short Stature, Indices of Adiposity, and Sex.

To compare whether other height-related loci were also affected by adiposity markers and sex in our cohort, we chose SNPs for 5 genes that are associated with variation in height in the human population (**Table III-6, Supplemental Table III-2**) (Haataja et al. 2011; Lettre et al. 2007; Patel et al. 2008; Shungin et al. 2015; Weedon et al. 2007). Interestingly, *GDF5* and *HMGA2* SNPs both showed a significant association only with short stature females (p=0.012 and p=0.037, respectively), with protective odds ratios (0.55, 0.43), suggesting that the minor alleles are associated with a taller stature specifically for women. For both of these gene SNPs, all three adiposity indices had confounding effects specifically in the female group, but FMI had the strongest (*GDF5*: p=0.0033; *HMGA2*: p=0.0092). Males presenting with the T allele at rs4966035 in *IGF1R* showed a 58% increased risk of short stature (p=0.034); adjustment by BMI, BAI and FMI had little effect. The SNPs for *IGF1* and *IGFBP3* genes did not show any significant association with short stature in our cohort, under unadjusted or adjusted conditions.

Our data show that variants in three major growth-related genes, *GHR*, *HMGA2* and *GDF5*, share associations with short stature as well as anthropometric traits linked to obesity in a sex-specific fashion.

5. DISCUSSION

According to the Canadian Community Health Survey, in 2015, 61.3% of Canadian adults were overweight or obese while obese individuals represented 26.7% of the total population (2015). Among individuals aged 45-64 years, 32% of men and 27.6% of women were obese, which is in line with the proportions we observed in our control population (**Table III-3a**). The prevalence of obesity for that age range is higher than for both younger adults and the elderly. Thus, although high BMI values were expected for our cohort, the striking sex-specific increase in obesity in our female short stature group (44.7% *vs.* 27.3% for control females and 27.8% for short stature males) was surprising and suggested a sex-specific association between severe short stature and obesity phenotypes.

Previous studies have shown a positive association between adult short stature and obesity (Bosy-Westphal et al. 2009; Hermanussen et al. 2005; Lara-Esqueda et al. 2004) with a few reporting a higher prevalence in women (Asao et al. 2006; Guerrero-Igea et al. 2001). Others have documented that short stature is a risk factor for women with very low incomes, often in parallel with malnutrition as a child (Alvarez et al. 2013; Florencio et al. 2001; Tyrrell et al. 2016). Although we cannot exclude the possibility of some cases of short stature due to childhood nutritional deficiency in our cohort, the CARTaGENE socioeconomic/demographic data show that more than 80% of the participants have an income >25,000$/year and 2x more individuals have a university level of education than the national average (Awadalla et al. 2013), making it unlikely to be a major confounding factor.

To better characterize the obesity phenotype in our female short stature group, we investigated whether the high BMI was paralleled by other adiposity indices. Because BMI does not discriminate between fat and lean body mass (Frankenfield et al. 2001; Peltz et al. 2010), it was important to determine these components separately. Calculating FMI and FFMI enabled us to make a direct comparison with BMI and to interpret the differences inherent only to the fat mass component regardless of height (VanItallie et al. 1990).

Women and men differ in the proportion of body fat: women have a higher total percentage (reflected in FMI values) for all BMI categories (Power and Schulkin 2008). Body fat distribution is also sex- and age specific. Premenopausal women show a gluteo-femoral pattern while men tend to have greater abdominal adiposity due to more visceral fat (Pulit et al. 2017). The majority of our female cohort is >50 years of age and, thus, likely to be

menopausal or post-menopausal; menopause transition has been associated with an increase in visceral fat due to decreased levels of endogenous estradiol and reduction in energy expenditure (Lovejoy et al. 2008). The elevated level of abdominal adiposity in our short stature men and women was apparent through both WC/Height and WHR indices, although a significant difference between the short and normal stature females was apparent only through WC/Height. WHR is a proxy measure for central adiposity independent of height but gave less information than WC/Height as both HC and WC seemed to increase proportionally in the short stature group. On the other hand, a large HC relative to WC has been shown to be a protective factor for a number of health endpoints (Heitmann et al. 2004). Thus, our data suggest that it may be better to take HC and WC into account independently, rather than in a ratio, to study obesity-related outcomes (Cameron et al. 2012) and that both WC/Height and HC/Height are useful markers of obesity in a short stature cohort.

Our second goal for this study was to determine if BMI and alternative adiposity indices may affect associations between short stature and polymorphisms in *GHR* as well as other height-related genes. We first investigated a *GHR* gene microsatellite polymorphism, a GT repeat located in a core promoter region of the *GHR* gene, upstream of the transcription start site of one of the main ubiquitously expressing *GHR* first exons (V9) (Dias et al. 2017). We previously showed associations of specific GT repeat genotypes (S/L, S/M) with idiopathic short stature in two different pediatric cohorts, suggesting a potential functional role of this microsatellite in regulating expression of the *GHR* gene. Microsatellites are found throughout the human genome and represent ~3% of the entire genome (Subramanian et al. 2003); their distribution is non-random with a significant concentration in the non-coding regions, particularly at promoter sites (Sawaya et al. 2013). They function as transcriptional modulators for many genes (Sawaya et al. 2012a) but also are associated with regulation of alternative splicing (Gabellini 2001).

Although no significant associations of the GT repeat with adult short stature were observed in the CARTaGENE cohort in a previous study (Dias et al. 2017), we re-examined the cohort in the context of both stature and obesity and found that the *GHR* GT repeat polymorphism showed a much stronger association with adiposity indices than stature, predominantly in the short stature females. Keeping the dichotomization of the height distribution as short stature *vs.* normal height, each GT genotype was tested for association

with each of the adiposity indices. When BMI was examined, we found that short stature women presenting with the M/L genotype had a 5 kg/m^2 increase which translated into a BMI class shift from overweight to obese (**Table III-4)**. As no associations were observed for the M/M genotype, the L allele appeared to be the risk allele, selectively in the short stature females. Moreover, the M/L genotype association was maintained across other adiposity markers, including WC/height, HC/Height and BAI, suggesting a potential risk association of this genotype with elevated levels of all of these adiposity markers. In two cases (HC/Height, BAI), this was true for both females and males, but there was always a more significant increase for the women. The sexual dimorphism that we have observed in the present study is not unusual for this *GHR* microsatellite: our previous investigation showed a higher prevalence of the S/M genotype in an Italian male idiopathic short stature pediatric cohort (Dias et al. 2017).

Our analyses also aimed to capture a more complex relationship of SNPs in the *GHR* (Dias et al. 2017) and other height-related genes (Lango Allen et al. 2010; Lanktree et al. 2011; Sanna et al. 2008; Weedon et al. 2007), by testing associations following adjustment for different adiposity indices. Indeed, after adjustments for BMI, BAI and FMI, most of the associations with *GHR* SNPs in the unadjusted model were strengthened in the short stature females, highlighting the pleiotropic role of the *GHR* gene in regulating both height and adiposity. To characterize these effects better, we would have to perform another stratification (e.g. by BMI categories); however, we would need a much larger cohort to fully exploit this question without an important loss of statistical power.

Interestingly, our study revealed that all six of the *GHR* SNPs associated with the sex-specific extreme short stature phenotype and affected by the three adiposity markers were located within or near potential functional regulatory elements that could affect transcriptional activity of the *GHR* gene. Three are upstream of the initial cluster of three first exons (V2, V9, V3). Rs666581 had a nominally significant risk odds ratio of 2.21 in association with female short stature that increased slightly when BMI was adjusted as a covariate, even though it is located ~250kb upstream of the initial promoter region of the *GHR* gene. In contrast, both rs1876790 and rs7732059 are found in a core promoter region, only a few kb upstream of the *GHR* first exon, V2. Rs1876790 showed markedly increased protective odds ratios for the female short stature phenotype after adjustment for all three indices but especially BAI (0.33),

while rs7732059 showed increased risk odds ratios with BMI (4.19) and FMI (3.88). It should be noted that rs7732059 is found within a region enriched with transcription factor binding sites validated through ChipSeq experiments (UCSC genome browser): these include sites for CCCTC binding factor (CTCF) and two sub-units of the cohesin complex that often co-localize in the human genome to regulate gene expression by chromatin looping and long range DNA interactions (Wendt and Peters 2009).

The other three SNPs are found in intron 1, downstream of the initial first exon cluster but upstream of several alternative first exons of the *GHR* gene (Dias et al. 2017; Goodyer et al. 2008). Rs2972400 showed an important increase of the protective odds ratios in the short stature females once adjusted for the adiposity indices, especially BAI (0.33), while rs13156541 exhibited increased risk odds ratios (4.0); the latter SNP is found in a DNAse hypersensitive region. The sixth SNP, rs2972419, has been specifically associated with pygmy short stature: there is a higher proportion of the ancestral allele (G) in the pygmy population vs. non-pygmies, while the derived allele (A) is associated with taller stature (Becker et al. 2013). In the present study, the A allele showed a nominally significant protective odds for the short stature phenotype in females (0.50) that was strengthened considerably after adjustment for BMI (0.41) or BAI (0.31). Thus, several *GHR* gene SNPs and microsatellite genotypes linked to the obese female short stature phenotype are located in regulatory regions and could have functional relevance.

Previously, our laboratory reported significantly lower *GHR* mRNA levels in omental and subcutaneous fat depots of obese vs. lean individuals (Erman et al. 2011a). Obesity-related factors (tumor necrosis factor alpha [TNFα], hypoxia-inducible factor 1 alpha [HIF-1α], glucocorticoids) were also shown to regulate *GHR* mRNA expression in human adipocytes by binding to specific response elements in core promoters (Wei et al. 2006). Moreover, adipogenesis is a process tightly controlled by key transcription factors, including CCAAT/enhancer-binding protein family members (Mota de Sa et al. 2017). The effects of these transcription factors on *GHR* gene expression are currently under investigation in our laboratory, including a study of a CCAAT box at the 5' end of the GT repeat polymorphism in the promoter region of V9.

We were also interested in comparing our results for the *GHR* gene with other genes identified as being strongly associated with variations in human height. *HMGA2* and GDF5

were amongst the first such genes found by GWAS (Alyaqoub et al. 2012; Buysse et al. 2009; Sanna et al. 2008; Weedon et al. 2007). In the present study, SNPs for both of these genes showed a significant protective association with the short stature phenotype but only in the females. In addition, all three adiposity index adjustments, but especially FMI, enhanced the protective odds ratios, and again only in the women. These findings are surprising for two reasons: *HMGA2* has not been linked to sex-specific traits before and *GDF5* has been reported to demonstrate male-specificity in association with WHRadjBMI (Shungin et al. 2015) or similarity in both sexes for WHR (Lango Allen et al. 2010; Sanna et al. 2008). The genes of several members of the GH-IGF axis are also associated with variation in height (Lanktree et al. 2011; Wood et al. 2014). Our analyses showed a small increased risk of short stature for the *IGF1R* SNP in males but no associations of the *IGF1R*, *IGF1* or *IGFBP3* SNPs with female short stature or obesity indices.

Previous GWAS studies have reported strong sexual dimorphism in the genetic regulation of fat distribution traits such as WHR and WHRadjBMI (Heid et al. 2010; Lindgren et al. 2009; Randall et al. 2013; Shungin et al. 2015; Winkler et al. 2015) but not overall obesity as assessed by BMI or total fat percentage (Pulit et al. 2017; Willer et al. 2009; Winkler et al. 2016). Only one GWAS meta-analysis has focused on anthropometric traits in populations within the lower 5^{th} percentiles of height (Berndt et al. 2013): the authors reported similar loci as in the general population but with the caveat that their analysis "does not necessarily extend to more extreme cutoffs, such as the …bottom 1^{st} percentile." In addition they did not separate the data by sex. Thus, the present study is the first to examine gene associations with obesity in the $<1^{st}$ height percentile in both females and males, looking at confounding effects of BMI and multiple anthropometric traits and indices.

In conclusion, our investigations have revealed an increased prevalence of obesity in severe short stature women. Comparative analyses of different anthropometric traits suggest the need to use indices normalized to height when investigating individuals at the tail of the height distribution. In addition, we determined that variants in three major growth-related genes, *GHR*, *HMGA2* and *GDF5*, share associations not only with short stature but also anthropometric traits linked to obesity in a sex-specific fashion. Finally, performing sex-specific analyses in health outcome studies is of importance as overall case/control comparisons may mask critical sex-specific differences (Reusch et al. 2018).

FUNDING: This project was funded by the Canadian Institutes for Health Research (CGG MOP 119313). CD received doctoral studentships from the Montreal Children's Hospital Research Institute and the Fonds de recherche du Québec – Santé (FRQ-S). This research was undertaken at the Research Institute of the McGill University Health Centre which is supported in part by the FRQ-S.

DISCLOSURE: No conflicts of interest, financial or otherwise are declared by the authors.

6. REFERENCES

2008. World Health Organization (WHO)- Waist Circumference and Waist-Hip Ratio: Report of a WHO Expert Consultation.

2014 revision. World Health Organization (WHO) growth Charts for Canada- Anthropometric calculator. available at http://whogrowthchartsca.

2015. Statistics Canada. Measured adult body mass index (BMI) (World Health Organization classification), by age group and sex, Canada and provinces, Canadian Community Health Survey Nutrition, occasional, CANSIM Table 105-2023.

Alvarez LS, Alejandro E, Goez JD, Carreno C, and Mancilla LP. 2013. The effects of socioeconomic status and short stature on overweight, obesity and the risk of metabolic complications in adults. Colomb Med (Cali) 44(3):146-154.

Alyaqoub F, Pyatt RE, Bailes A, Brock A, Deeg C, McKinney A, Astbury C, Reshmi S, Shane KP, Thrush DL et al. . 2012. 12q14 microdeletion associated with HMGA2 gene disruption and growth restriction. Am J Med Genet A 158A(11):2925-2930.

Asao K, Kao WH, Baptiste-Roberts K, Bandeen-Roche K, Erlinger TP, and Brancati FL. 2006. Short stature and the risk of adiposity, insulin resistance, and type 2 diabetes in middle age: the Third National Health and Nutrition Examination Survey (NHANES III), 1988-1994. Diabetes Care 29(7):1632-1637.

Awadalla P, Boileau C, Payette Y, Idaghdour Y, Goulet JP, Knoppers B, Hamet P, Laberge C, and Project CA. 2013. Cohort profile of the CARTaGENE study: Quebec's population-based biobank for public health and personalized genomics. International journal of epidemiology 42(5):1285-1299.

Beauregard C, Utz AL, Schaub AE, Nachtigall L, Biller BM, Miller KK, and Klibanski A. 2008. Growth hormone decreases visceral fat and improves cardiovascular risk markers in women with hypopituitarism: a randomized, placebo-controlled study. J Clin Endocrinol Metab 93(6):2063-2071.

Becker NS, Verdu P, Georges M, Duquesnoy P, Froment A, Amselem S, Le Bouc Y, and Heyer E. 2013. The role of GHR and IGF1 genes in the genetic determination of African pygmies' short stature. Eur J Hum Genet 21(6):653-658.

Bengtsson BA, Brummer RJ, Eden S, Bosaeus I, and Lindstedt G. 1989. Body composition in acromegaly: the effect of treatment. Clin Endocrinol (Oxf) 31(4):481-490.

Bergman RN, Stefanovski D, Buchanan TA, Sumner AE, Reynolds JC, Sebring NG, Xiang AH, and Watanabe RM. 2011. A better index of body adiposity. Obesity (Silver Spring) 19(5):1083-1089.

Berndt SI, Gustafsson S, Magi R, Ganna A, Wheeler E, Feitosa MF, Justice AE, Monda KL, Croteau-Chonka DC, Day FR et al. . 2013. Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. Nat Genet 45(5):501-512.

Boot AM, Engels MA, Boerma GJ, Krenning EP, and De Muinck Keizer-Schrama SM. 1997. Changes in bone mineral density, body composition, and lipid metabolism during growth hormone (GH) treatment in children with GH deficiency. J Clin Endocrinol Metab 82(8):2423-2428.

Bosy-Westphal A, Plachta-Danielzik S, Dorhofer RP, and Muller MJ. 2009. Short stature and obesity: positive association in adults but inverse association in children and adolescents. Br J Nutr 102(3):453-461.

Buysse K, Reardon W, Mehta L, Costa T, Fagerstrom C, Kingsbury DJ, Anadiotis G, McGillivray BC, Hellemans J, de Leeuw N et al. . 2009. The 12q14 microdeletion syndrome: additional patients and further evidence that HMGA2 is an important genetic determinant for human height. Eur J Med Genet 52(2-3):101-107.

Caballero B. 2007. The global epidemic of obesity: an overview. Epidemiol Rev 29:1-5.

Cameron AJ, Magliano DJ, Shaw JE, Zimmet PZ, Carstensen B, Alberti KG, Tuomilehto J, Barr EL, Pauvaday VK, Kowlessur S et al. . 2012. The influence of hip circumference on the relationship between abdominal obesity and mortality. Int J Epidemiol 41(2):484-494.

Dias C, Giordano M, Frechette R, Bellone S, Polychronakos C, Legault L, Deal CL, and Goodyer CG. 2017. Genetic variations at the human growth hormone receptor (GHR) gene locus are associated with idiopathic short stature. J Cell Mol Med 21(11):2985-2999.

Erman A, Veilleux A, Tchernof A, and Goodyer CG. 2011. Human growth hormone receptor (GHR) expression in obesity: I. GHR mRNA expression in omental and subcutaneous adipose tissues of obese women. Int J Obes (Lond) 35(12):1511-1519.

Florencio TM, Ferreira HS, de Franca AP, Cavalcante JC, and Sawaya AL. 2001. Obesity and undernutrition in a very-low-income population in the city of Maceio, northeastern Brazil. Br J Nutr 86(2):277-284.

Frankenfield DC, Rowe WA, Cooney RN, Smith JS, and Becker D. 2001. Limits of body mass index to detect obesity and predict body composition. Nutrition 17(1):26-30.

Gabellini N. 2001. A polymorphic GT repeat from the human cardiac Na+Ca2+ exchanger intron 2 activates splicing. Eur J Biochem 268(4):1076-1083.

Gallagher D, Heymsfield SB, Heo M, Jebb SA, Murgatroyd PR, and Sakamoto Y. 2000. Healthy percentage body fat ranges: an approach for developing guidelines based on body mass index. Am J Clin Nutr 72(3):694-701.

Goodyer CG, Rhani Z, and Zheng H. 2008. Expression of the hepatic specific V1 messenger ribonucleic acid of the human growth hormone receptor gene is regulated by hepatic nuclear factor (HNF)-4alpha2 and HNF-4alpha8. Mol Endocrinol 22(2):485-500.

Guerrero-Igea FJ, Lepe-Jimenez JA, Garrido-Serrano A, and Palomo-Gil S. 2001. Association among hyperinsulinemia, family history of diabetes, and diminutive stature in normoglycemic premenopausal women. Diabetes Care 24(3):602-603.

Haataja R, Karjalainen MK, Luukkonen A, Teramo K, Puttonen H, Ojaniemi M, Varilo T, Chaudhari BP, Plunkett J, Murray JC et al. . 2011. Mapping a new spontaneous preterm birth susceptibility gene, IGF1R, using linkage, haplotype sharing, and association analysis. PLoS Genet 7(2):e1001293.

Heid IM, Jackson AU, Randall JC, Winkler TW, Qi L, Steinthorsdottir V, Thorleifsson G, Zillikens MC, Speliotes EK, Magi R et al. . 2010. Meta-analysis identifies 13 new loci associated with waist-hip ratio and reveals sexual dimorphism in the genetic basis of fat distribution. Nat Genet 42(11):949-960.

Heitmann BL, Frederiksen P, and Lissner L. 2004. Hip circumference and cardiovascular morbidity and mortality in men and women. Obes Res 12(3):482-487.

Hermanussen M, Sunder M, Voigt M, and Tresguerres JA. 2005. Morbid obesity is associated with short stature. J Pediatr Endocrinol Metab 18(7):647-650.

Heymsfield SB, Heo M, Thomas D, and Pietrobelli A. 2011. Scaling of body composition to height: relevance to height-normalized indexes. Am J Clin Nutr 93(4):736-740.

Kamel A, Margery V, Norstedt G, Thoren M, Lindgren AC, Bronnegard M, and Marcus C. 1995. Growth hormone (GH) treatment up-regulates GH receptor mRNA levels in adipocytes from patients with GH deficiency and Prader-Willi syndrome. Pediatr Res 38(3):418-421.

Kelly TL, Wilson KE, and Heymsfield SB. 2009. Dual energy X-Ray absorptiometry body composition reference values from NHANES. PLoS One 4(9):e7038.

Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S et al. . 2010. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 467(7317):832-838.

Lanktree MB, Guo Y, Murtaza M, Glessner JT, Bailey SD, Onland-Moret NC, Lettre G, Ongen H, Rajagopalan R, Johnson T et al. . 2011. Meta-analysis of Dense Genecentric Association Studies Reveals Common and Uncommon Variants Associated with Height. Am J Hum Genet 88(1):6-18.

Lara-Esqueda A, Aguilar-Salinas CA, Velazquez-Monroy O, Gomez-Perez FJ, Rosas-Peralta M, Mehta R, and Tapia-Conyer R. 2004. The body mass index is a less-sensitive tool for detecting cases with obesity-associated co-morbidities in short stature subjects. Int J Obes Relat Metab Disord 28(11):1443-1450.

Laron Z. 2004. Increasing incidence of childhood obesity. Pediatr Endocrinol Rev 1 Suppl 3:443-447.

Lettre G, Butler JL, Ardlie KG, and Hirschhorn JN. 2007. Common genetic variation in eight genes of the GH/IGF1 axis does not contribute to adult height variation. Hum Genet 122(2):129-139.

Lindgren CM, Heid IM, Randall JC, Lamina C, Steinthorsdottir V, Qi L, Speliotes EK, Thorleifsson G, Willer CJ, Herrera BM et al. . 2009. Genome-wide association scan meta-analysis identifies three Loci influencing adiposity and fat distribution. PLoS Genet 5(6):e1000508.

Lovejoy JC, Champagne CM, de Jonge L, Xie H, and Smith SR. 2008. Increased visceral fat and decreased energy expenditure during the menopausal transition. Int J Obes (Lond) 32(6):949-958.

Mota de Sa P, Richard AJ, Hang H, and Stephens JM. 2017. Transcriptional Regulation of Adipogenesis. Compr Physiol 7(2):635-674.

Nam SY, and Lobie PE. 2000. The mechanism of effect of growth hormone on preadipocyte and adipocyte function. Obes Rev 1(2):73-86.

Patel AV, Cheng I, Canzian F, Le Marchand L, Thun MJ, Berg CD, Buring J, Calle EE, Chanock S, Clavel-Chapelon F et al. . 2008. IGF-1, IGFBP-1, and IGFBP-3 polymorphisms predict circulating IGF levels but not breast cancer risk: findings from the Breast and Prostate Cancer Cohort Consortium (BPC3). PLoS One 3(7):e2578.

Peltz G, Aguirre MT, Sanderson M, and Fadden MK. 2010. The role of fat mass index in determining obesity. Am J Hum Biol 22(5):639-647.

Power ML, and Schulkin J. 2008. Sex differences in fat storage, fat metabolism, and the health risks from obesity: possible evolutionary origins. Br J Nutr 99(5):931-940.

Pulit SL, Karaderi T, and Lindgren CM. 2017. Sexual dimorphisms in genetic loci linked to body fat distribution. Biosci Rep 37(1).

Randall JC, Winkler TW, Kutalik Z, Berndt SI, Jackson AU, Monda KL, Kilpelainen TO, Esko T, Magi R, Li S et al. . 2013. Sex-stratified genome-wide association studies including 270,000 individuals show sexual dimorphism in genetic loci for anthropometric traits. PLoS Genet 9(6):e1003500.

Reusch JEB, Kumar TR, Regensteiner JG, and Zeitler PS. 2018. Identifying the Critical Gaps in Research on Sex Differences in Metabolism Across the Life Span. Endocrinology 159(1):9-19.

Reyes-Vidal CM, Mojahed H, Shen W, Jin Z, Arias-Mendoza F, Fernandez JC, Gallagher D, Bruce JN, Post KD, and Freda PU. 2015. Adipose Tissue Redistribution and Ectopic Lipid Deposition in Active Acromegaly and Effects of Surgical Treatment. J Clin Endocrinol Metab 100(8):2946-2955.

Sandholt CH, Grarup N, Pedersen O, and Hansen T. 2015. Genome-wide association studies of human adiposity: Zooming in on synapses. Mol Cell Endocrinol 418 Pt 2:90-100.

Sanna S, Jackson AU, Nagaraja R, Willer CJ, Chen WM, Bonnycastle LL, Shen H, Timpson N, Lettre G, Usala G et al. . 2008. Common variants in the GDF5-UQCC region are associated with variation in human height. Nat Genet 40(2):198-203.

Savage MO, Attie KM, David A, Metherell LA, Clark AJ, and Camacho-Hubner C. 2006. Endocrine assessment, molecular characterization and treatment of growth hormone insensitivity disorders. Nat Clin Pract Endocrinol Metab 2(7):395-407.

Sawaya S, Bagshaw A, Buschiazzo E, Kumar P, Chowdhury S, Black MA, and Gemmell N. 2013. Microsatellite tandem repeats are abundant in human promoters and are associated with regulatory elements. PLoS One 8(2):e54710.

Sawaya SM, Bagshaw AT, Buschiazzo E, and Gemmell NJ. 2012. Promoter microsatellites as modulators of human gene expression. Advances in experimental medicine and biology 769:41-54.

Shungin D, Winkler TW, Croteau-Chonka DC, Ferreira T, Locke AE, Magi R, Strawbridge RJ, Pers TH, Fischer K, Justice AE et al. . 2015. New genetic loci link adipose and insulin biology to body fat distribution. Nature 518(7538):187-196.

Subramanian S, Madgula VM, George R, Kumar S, Pandit MW, and Singh L. 2003. SSRD: simple sequence repeats database of the human genome. Comp Funct Genomics 4(3):342-345.

Tyrrell J, Jones SE, Beaumont R, Astley CM, Lovell R, Yaghootkar H, Tuke M, Ruth KS, Freathy RM, Hirschhorn JN et al. . 2016. Height, body mass index, and socioeconomic status: mendelian randomisation study in UK Biobank. BMJ 352:i582.

VanItallie TB, Yang MU, Heymsfield SB, Funk RC, and Boileau RA. 1990. Height-normalized indices of the body's fat-free mass and fat mass: potentially useful indicators of nutritional status. Am J Clin Nutr 52(6):953-959.

Wabitsch M, Brenner RE, Melzner I, Braun M, Moller P, Heinze E, Debatin KM, and Hauner H. 2001. Characterization of a human preadipocyte cell strain with high capacity for adipose differentiation. Int J Obes Relat Metab Disord 25(1):8-15.

Weedon MN, Lettre G, Freathy RM, Lindgren CM, Voight BF, Perry JR, Elliott KS, Hackett R, Guiducci C, Shields B et al. . 2007. A common variant of HMGA2 is associated with adult and childhood height in the general population. Nat Genet 39(10):1245-1250.

Wei Y, Rhani Z, and Goodyer CG. 2006. Characterization of growth hormone receptor messenger ribonucleic acid variants in human adipocytes. J Clin Endocrinol Metab 91(5):1901-1908.

Wendt KS, and Peters JM. 2009. How cohesin and CTCF cooperate in regulating gene expression. Chromosome Res 17(2):201-214.

Willer CJ, Speliotes EK, Loos RJ, Li S, Lindgren CM, Heid IM, Berndt SI, Elliott AL, Jackson AU, Lamina C et al. . 2009. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. Nat Genet 41(1):25-34.

Winkler TW, Justice AE, Graff M, Barata L, Feitosa MF, Chu S, Czajkowski J, Esko T, Fall T, Kilpelainen TO et al. . 2015. The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. PLoS Genet 11(10):e1005378.

Winkler TW, Justice AE, Graff M, Barata L, Feitosa MF, Chu S, Czajkowski J, Esko T, Fall T, Kilpelainen TO et al. . 2016. Correction: The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. PLoS Genet 12(6):e1006166.

Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, Chu AY, Estrada K, Luan J, Kutalik Z et al. . 2014. Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet 46(11):1173-1186.

Table III-1: Demographic characteristics of the CARTaGENE cohort

	Sex[ǂ]	SS Median (min, max)	Controls Median (min, max)	P value[‡]
Total Number of Indviduals	**B**	168	207	
	M	92	108	
	F	76	99	
Age (years)	**B**	59.3 (41.1, 69.5)	57.9 (40.3, 70.1)	0.27
	M	59.3 (41.9, 69.5)	59.6 (40.3, 69.6)	0.99
	F	59.3 (41.1, 69.5)	56.4 (41.8, 70.1)	0.11
Height (cm)	**B**	154.3 (138.3, 159)	173.1 (160, 180)	<0.0001
	M	157.5 (148.6, 159)	176 (173, 180)	<0.0001
	F	145.5 (138.3, 147)	163.7 (160, 167)	<0.0001
Height (SDS)	**B**	−2.68 (−3.83, −2.4)	0.01 (−0.49, 0.59)	<0.0001
	M	−2.62 (−3.83, −2.4)	−0.08 (−0.49, 0.47)	<0.0001
	F	−2.69 (−3.8, −2.47)	0.08 (−0.48, 0.59)	<0.0001
Weight (kg)	**B**	65.61 (35.55. 114.1)	79.21 (45.85, 144.4)	<0.0001
	M	66.85 (48.17, 108.3)	85.67 (67.91, 144.4)	<0.0001
	F	61.29 (35.55, 114.1)	69.22 (45.85, 110.2)	0.0004

[ǂ] Sex: B Both, M Males, F Females.

[‡] P values calculated using the two tailed Mann-Whitney test.

Table III-2: Anthropometric traits and indices of the CARTaGENE cohort

	Sex ǂ	SS Median (min, max)	Controls Median (min, max)	P value‡
BMI (kg/m^2)	**B**	27.9 (16.7, 54.1)	27 (16.6, 44.8)	0.043
	M	27.05 (19.1, 45.5)	27.7 (21.5, 44.8)	0.31
	F	29.7 (16.7, 54.1)	25.7 (16.6, 41.6)	0.0002
WC (cm)	**B**	92.15 (61.2, 127.8)	95.2(40.5, 135.4)	0.018
	M	92.8 (75, 124.5)	101.4 (40.5, 135.4)	<0.0001
	F	90.1 (61.2, 127.8)	90.6 (62.7, 128.1)	0.86
HC (cm)	**B**	96 (74.3, 152)	101.6(42.5, 182.3)	<0.0001
	M	93.5(82.8, 125.7)	101.8 (42.5, 137.9)	<0.0001
	F	100.8 (74.3, 152)	101.6 (84.3, 182.3)	0.50
WHR	**B**	0.95(0.72, 1.19)	0.94(0.51, 1.15)	0.27
	M	1.0 (0.89, 1.19)	1.0(0.81, 1.15)	0.52
	F	0.88 (0.72,1.05)	0.88(0.51, 1.05)	0.37
WC/Height	**B**	0.61 (0.42, 0.88)	0.57 (0.23, 0.79)	<0.0001
	M	0.6 (0.48, 0.8)	0.58 (0.23, 0.76)	0.037
	F	0.62 (0.42, 0.88)	0.55 (0.39, 0.79)	<0.0001
HC/Height	**B**	0.63[0.51, 1.05]	0.6 (0.24, 1.13)	<0.0001
	M	0.6(0.53, 0.81)	0.58 (0.24, 0.77)	0.024
	F	0.69 (0.51, 1.05)	0.62 (0.51, 1.13)	<0.0001
BAI	**B**	33.1 (23.7, 68.9)	27.7 (20.2,71)	<0.0001
	M	29.8 (23.7, 47.6)	25.9 (20.2, 39.3)	<0.0001
	F	39.4 (24.2, 68.9)	30.3 (21.5, 71)	<0.0001
FM (kg)	**B**	19.2(2, 62.4)	23.8 (8.8, 62.4)	<0.0001
	M	17.6 (5.4, 50.4)	22.5 (11.4, 62.4)	<0.0001
	F	22.6 (2.0, 62.4)	26.2(8.8, 54.2)	0.0373
FMI (kg/m^2)	**B**	8.3 (0.9, 29.6)	8.4 (3.2,20)	0.3
	M	7.1 (2.1, 20.7)	7.35 (3.6, 19.4)	0.35
	F	10.6 (0.9, 29.6)	9.9 (3.2, 20)	0.2
Body Fat (%)	**B**	29.95(5.6, 54.9)	30.7(16.2, 50.1)	0.43
	M	26.1(11.3, 50.3)	25.9(16.2, 43)	0.78
	F	37.0 (5.6, 54.9)	37.1(17.1, 50.1)	0.37
FFM (kg)	**B**	45.8 (31.6, 63.8)	56.8(35.8, 83.6)	<0.0001
	M	49.6 (41.6, 63.8)	63.3 (54.6, 83.6)	<0.0001
	F	40.2 (31.6, 56.8)	43.8 (35.8, 61.4)	<0.0001
FFMI (kg/m^2)	**B**	19.75 (15.1, 26.4)	18.8 (13.3, 25.9)	<0.0001
	M	20.1(16.7, 25.8)	20.4 (17.2, 25.9)	0.54
	F	18.9 (15.1, 26.4)	16.4 (13.4, 22.9)	<0.0001

ǂ Sex: B Both, M Males, F Females.

‡ P values calculated using the two tailed Mann-Whitney test.

BMI: Body Mass Index, WC: Waist circumference, HC: Hip circumference, WHR: Waist to Hip Ratio, BAI: Body Adiposity Index, FM: Fat Mass, FMI: Fat Mass Index, FFM: Fat Free Mass, FFMI: Fat Free Mass Index

Table III-3:Frequencies of short stature *vs.* normal height control individuals for different anthropometric traits stratified by sex and obesity cut-offs: for (a) BMI, (b) WC, (c) WHR and (d) Body Fat %

(a)

BMI	Classification	Male Controls		Short Stature		Female Controls		Short Stature	
		N	%	N	%	N	%	N	%
<25	normal weight	28	26.2	25	27.8	43	43.4	14	18.4
25-29.9	overweight	41	38.3	40	44.4	29	29.3	28	36.9
≥30	obese	38	35.5	25	27.8	27	27.3	34	44.7
	Total	107	100.0	90	100.0	99	100.0	76	100.0

(b)

WC Classification	Male Controls		Short Stature	
	N	%	N	%
<102cm	56	51.9	66	74.2
≥102cm	52	48.1	23	25.8
Total	108	100.0	89	100.0

WC Classification	Female Controls		Short Stature	
	N	%	N	%
<88cm	44	45.4	32	43.8
≥88cm	53	54.6	41	56.2
Total	97	100.0	73	100.0

(c)

WHR Classification	Male Controls		Short Stature	
	N	%	N	%
<1	53	49.1	44	49.4
≥1	55	50.9	45	50.6
Total	108	100.0	89	100.0

WHR Classification	Female Controls		Short Stature	
	N	%	N	%
<0.85	37	38.1	22	30.1
≥0.85	60	61.9	51	69.9
Total	97	100.0	73	100.0

(d)

Body Fat % Classification	Male Controls		Short Stature	
	N	%	N	%
<29%	69	69	55	67.9
≥29%	31	31	26	32.1
Total	100	100.0	81	100.0

Body Fat % Classification	Female Controls		Short Stature	
	N	%	N	%
<41%	63	67.7	52	71.2
≥41%	30	32.3	21	28.8
Total	93	100.0	73	100.0

Table III-4: GT genotype associations with short stature and different anthropometric traits

(a) Males

		BMI			WC/Height			HC/Height			WHR		
N§	G‡	SS‡	Controls ‡	P	SS	Controls	P	SS	Controls	P	SS	Controls	P
4/4	S/S	32.0 (18.3-41.6)	29.3 (23.7-33.1)	0.34	0.68 (0.45-0.84)	0.60 (0.48-0.69)	0.49	0.63 (0.52-0.70)	0.58 (0.54-0.62)	0.26	1.02 (0.87-1.14)	1.06 (0.87-1.23)	0.49
16/15	S/M	25.1 (23.8-27.8)	26.2 (25.2-28.2)	0.38	0.57 (0.53-0.61)	0.54 (0.53-0.59)	0.76	0.57 (0.56-0.60)	0.57 (0.56-0.60)	0.73	0.98 (0.94-1.03)	0.95 (0.94-1.01)	0.95
37/52	M/M	26.9 (26.1-28.8)	28.4 (27.7-30.2)	0.13	0.60 (0.58-0.62)	0.60 (0.56-0.61)	0.72	0.60 (0.58-0.61)	0.60 (0.58-0.61)	0.82	1.0 (0.98-1.03)	1.0 (0.98-1.02)	0.82
28/25	M/L	26.6 (26.4-30.1)	27.3 (26.1-29.5)	0.81	0.58 (0.58-0.65)	0.57 (0.54-0.59)	**0.026**	0.61 (0.59-0.63)	0.57 (0.56-0.59)	**0.0075**	0.97 (0.97-1.04)	1.0 (0.95-1.02)	0.75
4/7	S/L	30.5 (20.0-46.8)	32.4 (27.3-39.4)	>0.99	0.64 (0.51-0.82)	0.63 (0.57-0.69)	0.69	0.62 (0.50-0.81)	0.62 (0.57-0.69)	0.68	1.02 (0.96-1.08)	1.02 (0.95-1.05)	0.69
2/5	L/L	31.8 (-45.7-109.3)	24.8 (22.7-28.6)	0.38	0.71	0.54 (0.49-0.61)	0.095	0.69	0.57 (0.54-0.59)	0.57	1.05 (0.09-2.0)	0.96 (0.92-1.04)	0.38

(a) Females

		BMI			WC/Height			HC/Height			WHR		
N	G	SS	Controls	P	SS	Controls	P	SS	Controls	P	SS	Controls	P
3/2	S/S	26.4 (19.2-31.1)	23.85 (18.1-19.6)	0.8	0.53 (0.41-0.69)	0.48 (0.41-0.54)	0.2	0.63 (0.59-0.70)	0.60 (0.28-0.91)	0.1	0.80 (0.61-1.09)	0.79 (0.54-1.04)	0.8
14/16	S/M	31.45 (27.05-33.71)	25.1(23.6-29.02)	**0.046**	0.64 (0.59-0.69)	0.55 (0.51-0.61)	**0.01**	0.70 (0.66-0.76)	0.62 (0.59-0.67)	**0.01**	0.89 (0.87-0.94)	0.90 (0.85-0.94)	0.86
32/44	M/M	29.6 (27.13-32.26)	26.9(25.87-29.24)	0.19	0.61 (0.57-0.64)	0.56 (0.53-0.59)	0.053	0.68 (0.66-0.72)	0.64 (0.62-0.67)	**0.017**	0.87 (0.85-0.90)	0.88 (0.85-0.90)	0.76
18/29	M/L	30.95 (28.53-36.94)	25.6 (24.31-28.87)	**0.009**	0.66 (0.60-0.73)	0.53 (0.52-0.58)	**0.0013**	0.74 (0.68-0.80)	0.61 (0.60-0.69)	**0.0003***	0.88 (0.85-0.93)	0.87 (0.82-0.89)	0.33
3/4	S/L	28.15 (22.43-36.92)	25.2 (16.41-36.34)	0.34	0.59 (0.55-0.65)	0.54 (0.48-0.59)	0.086	0.65 (0.45-0.95)	0.61 (0.53-0.72)	0.22	0.91 (0.64-1.08)	0.87 (0.80-0.92)	0.64
5/4	L/L	31.3 (24.72-40.72)	26.1 (23.51-30.64)	0.13	0.67 (0.57-0.77)	0.54 (0.47-0.64)	**0.04**	0.72 (0.64-0.87)	0.65 (0.52-0.74)	**0.04**	0.88 (0.81-0.96)	0.89 (0.72-1.06)	>0.99

(a) Males

		BAI			FMI			FFMI		
N[§]	G[ǂ]	SS[ǂ]	Controls[ǂ]	P	SS	Controls	P	SS	Controls	P
4/4	S/S	32.0 (23.2-37.4)	25.9 (23.2-28.3)	0.23	10.7 (2.1-16.3)	7.9 (5.1-10.6)	0.49	21.1 (15.9-25.5)	21.2 (18.1-23.1)	>0.99
16/15	S/M	27.3 (26.8-29.9)	24.6 (23.8-27.03)	**0.01**	6.1 (4.9-7.6)	6.1 (5.6-7.5)	0.49	20.3 (19.3-21.0)	20.1 (19.5-20.8)	>0.99
37/52	M/M	29.7 (28.5-30.5)	26.4 (25.7-27.5)	**<0.0001***	7.6 (6.4-8.5)	7.8 (7.3-9.1)	0.30	19.9 (19.6-20.9)	20.5 (20.3-21.4)	0.16
28/25	M/L	30.2 (29.0-32.4)	24.6 (23.9-26.0)	**<0.0001***	6.7 (6.3-9.4)	7.1 (6.1-8.6)	0.81	20.1 (19.5-21.2)	20.0 (19.7-21.1)	0.99
4/7	S/L	30.8 (21.4-47.5)	28.5 (25.0-33.6)	**0.049**	8.9 (0.85-21.6)	9.6 (6.0-16.0)	0.91	22.2 (18.8-25.4)	21.1 (18.9-25.4)	0.61
2/5	L/L	37.1 (-97-171)	24.8 (22.7-26.5)	0.10	9.9 (-33.3-53.1)	5.7 (3.1-8.5)	0.27	21.9 (-11.1-54.9)	19.4 (17.5-21.3)	0.53

(b) Females

		BAI			FMI			FFMI		
N	G	SS	Controls	P	SS	Controls	P	SS	Controls	P
3/2	S/S	34.4 (30.9-39.7)	28.3 (8.0-39.7)	0.2	6.3 (3.11-11.03)	7.7 (-5.0-20.4)	0.8	18 (12.5-23.7)	16.2 (9.2-23.1)	0.4
14/16	S/M	39.7 (36.7-45.1)	30.2 (28.2-34.2)	**3E-04 ***	12.3 (9.41-14.58)	9.1 (7.76-11.85)	0.14	19.0 (17.7-20.0)	15.9 (15.7-17.5)	**0.0022**
32/44	M/M	39.1 (36.5-41.9)	31.4 (30.6-34.1)	**<0.0001***	9.2 (8.20-11.93)	10.8 (9.44-11.96)	0.380	18.9 (18.5-20.5)	16.4 (16.2-17.3)	**<0.0001***
18/29	M/L	42.8 (38.8-48.7)	29.4 (28.7-35.6)	**<0.0001***	12.2 (10.11-16.67)	8.8 (8.26-11.58)	0.079	19.8 (18.9-21.0)	16.9 (16.0-17.4	**<0.0001***
4/4	S/L	42.7 (-45.0-130.4)	29.1 (23.1-38.1)	0.27	9.9 (5.4-16.6)	8.9 (2.51-16.14)	0.490	18.4 (16.9-20.5)	16.3 (13.8-20.2)	0.2
5/4	L/L	42.2 (35.3-54.6)	32.5 (22.2-40.5)	**0.032**	12.9 (8.52-18.12)	10.05 (8.71-12.14)	0.29	18.3 (16.2-22.9)	16.3 (14.5-18.7)	0.11

Table III-4: For each anthropometric trait, the different GT genotypes were tested for their potential association with short stature using the Mann-Whitney test: highlighted in bold are tests that are nominally significant while those with asterisks are tests that remained significant after Bonferroni correction ($p<0.0005$). S = short (<24 repeats); M = medium (24-28 repeats); L = long (>28 repeats).

N[§] Number of individuals : Short stature (SS)[ǂ]/Controls[ǂ], G[ł] Genotype class of the GT repeat, SS[ǂ] and Controls[ǂ]: Mann-Whitney test results are presented as median (95%CI).

				Unadjusted			Adjusted[†] BMI			Adjusted BAI			Adjusted FMI		
SNP	A[§]	M[∥]	Sex[‡]	OR	95% CI	P value[‡]	OR	95% CI	P value	OR	95% CI	P value[‡]	OR	95% CI	P value[‡]
rs666581	A	A	B	1.84	1.06-3.2	**0.031**	1.89	1.08-3.30	**0.026**	1.62	0.89-2.97	**0.12**	1.86	1.04--3.31	**0.036**
			M	1.62	0.74-3.53	0.23	1.66	0.76-3.62	0.21	1.47	0.58-3.72	0.42	1.63	0.72-3.72	0.24
			F	2.11	0.96-4.62	0.064	2.36	1.05-5.32	**0.038**	2.42	0.97-6.05	0.059	2.21	0.97-5.03	0.058
rs66487711	C	A	B	1.70	0.99-2.93	0.054	1.76	1.02-3.06	**0.043**	1.66	0.92-2.99	0.093	1.77	1.0-3.12	**0.046**
			M	1.99	0.93-4.26	0.078	2.04	0.95-4.38	0.069	2.00	0.80-4.96	0.14	2.05	0.92-4.55	0.08
			F	1.44	0.66-3.14	0.37	1.65	0.73-3.71	0.23	1.81	0.74-4.56	0.21	1.59	0.71-3.56	0.26
rs1876790	C	A	B	0.73	0.51-1.03	0.076	0.71	0.49-1.02	0.06	0.71	0.47-1.05	0.088	0.74	0.51-1.06	0.097
			M	0.83	0.53-1.29	0.40	0.87	0.56-1.36	0.55	0.72	0.41-1.25	0.25	0.88	0.56-1.38	0.57
			F	0.54	0.29-1.0	0.05	0.44	0.23-0.85	**0.015**	0.33	0.14-0.74	**0.007**	0.49	0.25-0.92	**0.028**
rs7732059	C	R	B	1.94	0.84-4.43	0.12	2.10	0.91-4.86	0.082	1.75	0.70-4.38	0.23	2.06	0.89-4.74	0.091
			M	1.18	0.37-3.80	0.78	1.17	0.36-3.78	0.79	1.23	0.28-5.36	0.78	1.26	0.39-4.06	0.70
			F	3.38	0.99-11.57	0.052	4.19	1.19-14.7	**0.025**	3.59	0.91-14.1	0.067	3.88	1.11-13.56	**0.034**
rs2972400	A	A	B	0.74	0.52-1.06	0.1	0.72	0.50-1.04	0.080	0.72	0.48-1.07	0.1	0.75	0.52-1.0	0.13
			M	0.87	0.55-1.35	0.52	0.91	0.58-1.43	0.68	0.75	0.28-0.43	0.31	0.92	0.58-1.45	0.71
			F	0.54	0.29-1.0	0.05	0.44	0.23-0.85	**0.015**	0.33	0.14-0.74	**0.007**	0.49	0.25-0.92	**0.028**
rs13156541	G	R	B	2.11	0.93-4.80	0.074	2.37	1.03-5.45	**0.042**	2.40	0.99-5.83	0.054	2.29	1.0-5.26	0.050
			M	1.62	0.54-4.84	0.39	1.57	0.52-4.74	0.43	2.51	0.67-9.39	0.17	1.72	0.57-5.20	0.34
			F	3.25	0.92-11.47	0.067	4.03	1.1-14.71	**0.035**	4.04	1.03-15.9	**0.046**	3.77	1.04-13.68	**0.044**
rs11744988	T	A	B	1.72	0.99-2.97	0.054	1.78	1.02-3.10	**0.0 43**	1.63	0.90-2.97	0.11	1.79	1.00-3.16	**0.047**
			M	2.04	0.93-4.48	0.076	2.09	0.95-4.60	0.067	2.06	0.80-5.32	0.13	2.10	0.92-4.79	0.08
			F	1.44	0.66-3.14	0.37	1.65	0.73-3.71	0.23	1.81	0.72-4.56	0.21	1.59	0.71-3.56	0.26
rs2972419	A	A	B	0.70	0.49-1.0	0.049	0.68	0.48-0.98	**0.037**	0.68	0.46-1.01	0.059	0.72	0.50-1.04	0.077
			M	0.82	0.53-1.28	0.39	0.87	0.56-1.36	0.53	0.68	0.39-1.18	0.17	0.87	0.55-1.37	0.56
			F	0.50	0.27-0.92	**0.026**	0.41	0.21-0.80	**0.0083**	0.31	0.14-0.70	**0.0049**	0.47	0.24-0.89	**0.02**
rs12153009	A	A	B	1.41	1.0-1.98	0.052	1.42	1.00-2.00	0.05	1.34	0.91-1.96	0.13	1.50	1.05-2.13	**0.026**
			M	1.22	0.76-1.95	0.42	1.24	0.77-1.99	0.37	1.27	0.74-2.19	0.39	1.29	0.79-2.10	0.31
			F	1.65	1.0-2.75	0.052	1.68	0.99-2.86	0.055	1.71	0.92-3.15	0.088	1.77	1.05-2.99	**0.033**
rs12233949	C	R	B	2.79	1.04-7.52	**0.043**	2.77	1.02-7.52	**0.045**	2.72	0.94-7.81	0.064	2.87	1.06-7.76	**0.038**
			M	2.48	0.72-8.52	0.15	2.61	0.75-9.04	0.13	1.81	0.46-7.04	0.39	2.64	0.76-9.12	0.12
			F	3.23	0.61-17.22	0.17	3.71	0.66-20.84	0.14	1.65	0.87-3.11	0.12	3.30	0.61-17.9	0.17

Table III-5: Logistic regression analysis of SNPs at the *GHR* locus with short stature and markers of adiposity

Table III-5: OR: odds ratio, 95% CI: 95% confidence intervals. [§] A : Minor allele, [ɣ]M: Genetic Model : A Additive, R Recessive [ǂ] Sex: B Both, M Males, F Females, [‡] p value calculated from binary logistic regression analyses in additive or recessive genetic models with short stature as a dependent variable. ł Adjusted: refers to a binary logistic regression analysis with BMI, BAI or FFMI as a covariate; age has been adjusted for all. Highlighted in bold are tests that are nominally significant ($p<0.05$); tests significant after Bonferroni correction would be $p<0.0028$.

					Unadjusted			Adjusted[ł] BMI			Adjusted BAI			Adjusted FMI		
SNP	Gene	A[§]	M[ɣ]	Sex[ǂ]	OR	95% CI	P value[‡]	OR	95% CI	P value	OR	95% CI	P value	OR	95% CI	P value
rs224333	GDF5	A	A	B	0.64	0.46-0.87	**0.0053***	0.63	0.45-0.86	**0.0042***	0.57	0.40-0.81	**0.0018***	0.61	0.44-0.85	**0.0039***
				M	0.74	0.48-1.13	0.16	0.76	0.49-1.17	0.21	0.60	0.36-1.00	0.053	0.77	0.49-1.21	0.25
				F	0.53	0.32-0.85	**0.0086***	0.51	0.31-0.83	**0.0073***	0.52	0.30-0.93	**0.026**	0.47	0.28-0.78	**0.0033***
rs1042725	HMGA2	C	R	B	0.51	0.30-0.87	**0.013**	0.50	0.29-0.87	**0.013**	0.48	0.26-0.88	0.017	0.40	0.22-0.71	**0.0018***
				M	0.59	0.29-1.23	0.16	0.60	0.29-1.23	0.16	0.71	0.30-1.69	0.44	0.47	0.22-1.03	0.059
				F	0.43	0.19-0.95	**0.038**	0.38	0.16-0.87	**0.022**	0.38	0.15-0.97	**0.043**	0.31	0.13-0.75	**0.0092***
rs4966035	IGF1R	T	A	B	1.29	0.95-1.76	0.12	1.22	0.89-1.68	0.22	1.13	0.80-1.61	0.49	1.19	0.86-1.64	0.29
				M	1.58	1.04-2.41	**0.034**	1.51	0.99-2.32	0.058	1.58	0.96-2.58	0.071	1.40	0.91-2.16	0.13
				F	1.0	0.62-1.60	0.99	0.85	0.52-1.41	0.54	0.62	0.33-1.16	0.13	0.95	0.59-1.55	0.84
rs1019731	IGF1	A	A	B	0.83	0.55-1.27	0.40	0.81	0.53-1.24	0.32	0.84	0.53-1.35	0.48	0.90	0.58-1.39	0.62
				M	0.84	0.46-1.52	0.56	0.81	0.44-1.49	0.50	0.95	0.49-1.85	0.89	0.86	0.47-1.60	0.64
				F	0.85	0.46-1.55	0.59	0.81	0.43-1.52	0.50	0.78	0.36-1.69	0.53	0.94	0.51-1.75	0.85
rs2854744	IGFBP3	T	A	B	0.94	0.70-1.26	0.68	0.97	0.72-1.30	0.83	0.89	0.65-1.24	0.50	0.98	0.72-1.33	0.91
				M	0.86	0.58-1.27	0.45	0.93	0.62-1.38	0.71	0.68	0.43-1.09	0.11	0.92	0.60-1.39	0.68
				F	1.03	0.66-1.60	0.91	1.12	0.70-1.78	0.64	1.32	0.77-2.28	0.32	1.07	0.67-1.69	0.79

Table III-6: Logistic regression analysis of height-related gene SNPs with short stature and markers of adiposity

OR: odds ratio, 95% CI: 95% confidence intervals. [§] A : Minor allele, [ɣ]M: Genetic Model : A Additive, R Recessive [ǂ] Sex: B Both, M Males, F Females, [‡] p value calculated from binary logistic regression analyses in additive or recessive genetic models with short stature as a dependent variable. ł Adjusted: refers to a binary logistic regression analysis with BMI, BAI or FFMI as covariates; age has been adjusted for all. Highlighted in bold are tests that are nominally significant ($p<0.05$); those with asterisks are tests that remained significant after Bonferroni correction ($p<0.01$).

Supplemental Table III-1: *GHR* SNP panel information

rs ID dbSNP147	Major:Minor Allele	Position in GRCh37 hg19 assembly	MAF European	MAF CARTaGENE cohort (controls+cases)	Location in the *GHR* gene region
rs666581	C:A	42152335	A: 0.062	A: 0.08	> 200kb upstream V2
rs3764451	T:C	42398465	C: 0.16	C: 0.12	> 25kb upstream V2
rs66487711	T:C	42421689	C: 0.07	C: 0.08	proximal V2 promoter
rs2940927	G:A	42422389	A: 0.48	A: 0.46	proximal V2 promoter
rs1876790	T:C	42422855	C: 0.19	C:0.20	core V2 promoter
rs7732059	G:C	42423472	C: 0.29	C: 0.26	core V2 promoter
rs2972400	G:A	42425830	A: 0.19	A: 0.20	downstream of V3
rs4642376	G:T	42432739	T: 0.19	T: 0.20	intron 1
rs1509460	A:C	42433819	C: 0.48	C: 0.46	intron 1
rs13171720	C:T	42442715	T: 0.24	T: 0.20	intron 1
rs13156541 *	C:G	42461964	G: 0.34	G: 0.28	intron 1
rs11744988	C:T	42474696	T: 0.07	T: 0.08	intron 1
rs2972419	G:A	42483877	A: 0.18	A: 0.21	intron 1 (associated with short stature in Pygmy population)
rs2972393	G:A	42496257	A: 0.46	A:0.45	intron 1
rs4509029	G:A	42506775	A: 0.5	A: 0.49	intron 1
rs4129472	A:G	42523172	G: 0.22	G: 0.19	intron 1
rs12153009	G:A	42553879	A: 0.29	A: 0.25	intron 1
rs7735889	A:G	42562199	G: 0.31	G: 0.27	proximal V5 promoter
rs12233949	G:C	42607461	C: 0.27	C: 0.24	intron 2
rs4866941	G:A	42617492	A: 0.30	A: 0.26	intron 2
rs4292454	T:C	42623245	C: 0.46	C: 0.43	intron 2
rs6873545	T:C	42631264	C: 0.31	C: 0.28	intron 3 (used as a tag to genotype the *GHR* d3/fl polymorphism)
rs6886047	A:T	42643200	T: 0.30	T: 0.26	intron 3
rs4273617	A:G	42695471	G: 0.32	G: 0.29	intron 5
rs6180	A:C	42719239	C: 0.42	C: 0.48	exon 10
rs1559286	T:G	42752037	G: 0.05	G: 0.07	~ 25kb downstream of *GHR*
rs6880056 *	A:T	42847750	T: 0.22	T: 0.23	~120 kb downstream of *GHR*

* CEU population (Utah residents with Northern and Western European ancestry from the CEPH collection)

MAF: Minor Allele Frequency in the European population (1000 Genomes database). SNPs highlighted in blue are shown in the results section.

Supplemental Table III-2: Height-related gene SNP panel information

rs ID dbSNP147	Gene ID	Major:Minor Allele	Position in GRCh37 hg19 assembly	MAF European	MAF CARTaGENE cohort (controls+cases)
rs224333	*GDF5*	G:A	Chr20:34023962	A:0.37	A:0.36
rs1042725	*HMGA2*	T:C	Chr12:66358347	C:0.46	C: 0.44
rs4966035	*IGF1R*	G:A	Chr15:99429462	A:0.36	A: 0.31
rs1019731	*IGF1*	C:A	Chr12:102864425	A:0.13	A: 0.14
rs2854744	*IGFBP3*	G:T	Chr7:45961075	T:0.46	T: 0.48

MAF: Minor Allele Frequency in the European population (1000 Genomes database).

Supplemental Table III-3: Regression coefficients of the adiposity markers adjusted for height

	Variable	β Coefficient	Significance	OR	95% CI
Males					
	BMI	-.025	0.42	0.98	0.97-1.04
	BAI	0.329	.000	1.39	1.25-1.55
	FMI	-.006	0.89	0.99	0.91-1.09
	FFMI	-.052	0.532	.95	0.81-1.12
	WHR	0.026	0.21	1.03	0.99-1.07
Females					
	BMI	0.092	0.001	1.10	1.04-1.16
	BAI	0.186	.000	1.21	1.13-1.28
	FMI	0.047	0.20	1.05	0.98-1.13
	FFMI	0.71	.000	2.03	1.61-2.65
	WHR	0.021	0.3	1.02	0.98-1.06

OR: odds ratio. 95% CI: 95% confidence intervals

CHAPTER IV

My previous two studies showed significant associations of the GT microsatellite with short stature linked with obesity.

To explore the biological significance of this microsatellite, I conducted a series of functional studies to compare the effects of the length of the GT repeat in modulating *GHR* transcriptional activity.

A GT microsatellite polymorphism in the human *growth hormone (GH) receptor (GHR)* gene affects its transcriptional activity

Christel Dias[1], Samar Elzein[1], Robert Sladek[1,2] and *Cynthia Gates Goodyer[1,3]

[1]Division of Experimental Medicine, McGill University, Montreal, Quebec, Canada; [2]Department of Human Genetics, McGill University, Montreal, Quebec, Canada; [3]Department of Pediatrics, McGill University, Montreal, Quebec, Canada

Running Title: *GT microsatellite and GHR transcriptional activity*

Keywords: ***GH receptor* gene, GT repeat microsatellite, *cis/trans* effects, allelic imbalance**

Number of Figures & Tables: Five Figures and one Table + three Supplementary Figures and two Supplementary Tables

Number of words: 5995

***Corresponding author and person to whom reprint requests should be addressed**:

Dr. Cynthia Gates Goodyer, Research Institute of McGill University Health Centre,

EM0.3211, 1001 Decarie Boulevard, Montreal, Quebec, Canada H4A 3J1

Telephone: + (514) 934-1934, ext. 22481

Email: cindy.goodyer@muhc.mcgill.ca

1. ABSTRACT (250 words)

Growth hormone (GH) binding to its specific receptor (GHR) at the surface of target cells activates multiple signaling pathways, leading to changes in gene expression and function. Dysregulation of GHRs can lead to pathophysiological states that most commonly affect height and body composition. We previously showed the association of a polymorphic (n=15-37) GT microsatellite in the *GHR* gene promoter with short stature and obesity in a sex-specific manner. In the present study we evaluated the functional relevance of this microsatellite in regulating *GHR* expression. Using luc-reporter constructs containing different GT repeat lengths, we observed a significant GT allelic effect following HIF-1α stimulation and characterized a repressor role following C/EBPβ stimulation that was independent of GT length. We evaluated the effect of microsatellite length on *GHR* expression *in vivo* using a panel of 41 CEPH lymphoblastoid cell lines. Using a novel digital PCR assay to measure *GHR* allelic imbalance (AI), we showed a high prevalence of AI at rs6180 (~86%), a higher degree of AI in males (p=0.0045) and a nominal effect of the long (L: >28GT) allele in males (p=0.047). Total *GHR* expression analysis revealed significantly lower *GHR* expression in cells carrying medium (M: 24-28GT) alleles *vs*. cells with short (S: <24GT) or L alleles in females (p=0.029). Finally, L/M genotype cells showed significantly lower expression of *IGF1* in males and *BCL2* in females, suggesting a potential *trans*-acting effect. Our data suggest that this microsatellite may act as a fine-tuning modulator of *GHR* expression in a context- and sex-specific manner.

2. INTRODUCTION

Growth Hormone (GH) is essential for normal musculoskeletal development in children but also has important regulatory effects on protein, carbohydrate and lipid metabolism at all stages of life . It functions by binding to a dimer of its high-affinity receptor (GHR), activating multiple intracellular signaling pathways and modulating cellular gene expression and function (Brooks and Waters 2010b; Rowlinson et al. 2008b).

The ability of GH to exert its pleiotropic effects is contingent on the availability of its receptor at the surface of the target cell. Individuals with low GHR levels or a dysfunctional GHR do not respond normally to GH: they are not only short, but also have decreased bone mineral density and increased adiposity, leading to a greater risk of developing osteoporosis, lipid disorders and cardiovascular disease (Savage et al. 2006). Persons with an enhanced GH response, due to increased GH secretion or increased tissue levels of GHR, exhibit excessive growth and abnormal protein, lipid and carbohydrate metabolism, leading to an increased incidence of cardiomyopathies, hypertension, diabetes and several types of cancers (Ben-Shlomo and Melmed 2008). Therefore, the GHR has a central position within the GH-IGF1 axis and dysregulation of its expression can lead to significant pathophysiological consequences.

The human *GHR* gene spans ~300kb on chromosome 5 (Barton et al. 1989b; Godowski et al. 1989). The coding region is defined by exons 2-10 where exon 2 contains the translation start site (**Figure IV-1a**). The *GHR* gene has a complex promoter region that produces fourteen different *GHR* mRNAs with unique 5'UTRs derived from thirteen different first exons. These variant mRNAs all splice into the same site in exon 2, 11bp upstream from the ATG translation start site, and thus encode the same protein (Goodyer et al. 2001c; Orlovskii et al. 2004; Pekhletsky et al. 1992; Wei et al. 2006). Four of the variant mRNAs are tightly regulated and are expressed only in normal postnatal liver; while the other transcripts are present in all fetal and postnatal tissues (Goodyer et al. 2001c; Wei et al. 2006). The four exons responsible for postnatal liver-specific *GHR* mRNAs (V1, V4, V7, V8) are present in a 2kb domain (Module B) located ~18kb upstream of exon 2, while the three major exons transcribing ubiquitous *GHR* mRNAs (V2, V3, V9) are located within a 1.6kb region ~100kb further upstream (Module A) **(Figure IV-1a**).

Our lab has previously reported the presence of a GT repeat microsatellite polymorphism in the V9 promoter region, with a range of 15-37 (median 26) repeats (**Figure IV-1b**) (Dias et al. 2017; Hadjiyannakis et al. 2001). In addition, we have described several functional transcription factor binding sites (DBP [D site binding protein], C/EBPß [CCAAT enhancer-binding protein β], HIF1α [hypoxia-inducible factor-1α]) flanking the GT repeat sequence (Erman et al. 2011b; Kenth et al. 2008). More recently, we have shown that this polymorphic microsatellite is associated with two different complex traits, short stature (Dias et al. 2017) and obesity (**Chapter III**), with marked sex specificity differences.

The aim of the present study has been to better understand the functional relevance of the GT repeat polymorphism in Module A of the *GHR* gene. Polymorphic dinucleotide repeats are common throughout the human genome and widely used as genetic markers. Population studies have found that high mutability due to slippage can lead to complex polymorphic characteristics (Bhargava and Fuentes 2010). The length of the GT repeat in promoter regions has also been shown to modulate the effects of flanking transcription factor response elements in several genes (Gonzalez et al. 2007; Itokawa et al. 2003a; Searle and Blackwell 1999). Therefore, we examined the potential effects of the Module A GT polymorphism in modulating *GHR* transcriptional activity using three different approaches: **(1)** luciferase reporter assays, to compare the ability of the S (short), M (medium) and L (long) repeats to modulate transcriptional activity under both basal and transcription factor-stimulated conditions; **(2)** droplet digital PCR, to determine the effects of the native GT polymorphism on allelic differential expression *in vivo* using lymphoblastoid cell lines (LCLs); and **(3)** quantitative PCR to measure expression of total *GHR* and four axis-related genes (*GH*, *IGF1*, *BCL2* and *SOCS2*), to determine if there is an association with a specific GT genotype in LCLs. Our study characterizes potential molecular mechanisms whereby this microsatellite may affect *GHR* expression.

3. MATERIALS AND METHODS

A. Cell cultures

HEK293 cells (American Type Culture Collection, Bethesda, MD, USA) were maintained in Dulbecco's modified Eagle's medium (DMEM) (Invitrogen, Burlington, ON, Canada)

supplemented with 10% heat-inactivated fetal bovine serum and 25 mM HEPES. Forty-one Epstein-Barr virus (EBV)-transformed lymphoblastoid cell lines (LCLs) with specific GT genotypes were obtained from the Coriell Cell Repositories (Camden, NJ, USA); they were derived from individuals originating from France, Venezuela and the state of Utah in the United States (collected by Fondation Jean Dausset-CEPH, Centre d'Etude du Polymorphisme Humain) (**Table IV-S1**). The cell lines were grown in suspension in RPMI 1640 medium (Invitrogen, Burlington, ON, Canada) supplemented with 15% (vol/vol) non-heat-inactivated fetal bovine serum, 100 IU/ ml penicillin G and 1.6 mg/ ml gentamycin. All cells were incubated at 37°C in 5% CO_2 in air.

B. GT repeat constructs

Three pA3Luc-reporter constructs containing different lengths of the GT repeat were generated by amplifying ~1.2 kb of the *GHR* Module A region (from the last 91 bp of V2 to 14 bp past the end of V3) (**Figure IV-2a**), using genomic DNA of individuals previously genotyped: a small (S) GT19, a medium (M) GT26 and a long (L) GT33 reporter construct, following previously described GT allele classifications: S = <24 GT, M = 24-28 GT, L = >28 GT (Dias et al. 2017). In addition, a GT-deleted reporter construct (GTdel) was generated using specific forward and reverse primers (**Table IV-S2**) and the QuickChange Lightning Site-directed Mutagenesis kit (Stratagene, La Jolla, CA, USA). The cloned inserts and the GT-deletion were confirmed by sequencing (Génome-Québec, Montreal, QC, Canada).

C. Transient transfection and luciferase reporter gene assays

HEK293 cells (0.5×10^5) were seeded in 24-well plates and grown to 40–60% confluency. Transfections were performed using Polyfect reagent (Qiagen, Germantown, MD, USA). Each well was transfected with 0.25 μg of *GHR* pA3Luc reporter vector (GT19, GT26, GT33 or GTdel) or empty pA3Luc vector (Tremblay et al. 2000), along with 5 ng of pRSV-β-galactosidase (Promega Corporation, Madison, WI, USA) for normalization; the total DNA per well was made up to 0.4 μg with Sp64. For the stimulation experiments, cells were co-transfected with expression vectors for DBP (pcDNA3-DBP), HIF-1α (pcDNA3-HA-HIF-1α) or C/EBPβ (pcDNA3-C/EBPβ) at two doses (100/200ng, 50/100ng and 2/4ng, respectively). Forty-eight hours after transfection, the cells were washed with cold phosphate-buffered saline and harvested in 200 μl of lysis buffer (0.5% NP-40, 0.01 M DTT and 0.1 M Tris [pH 8]) for 15 min

at room temperature. For the β-galactosidase assay, 10 µl of the lysates were added to each well of a 96-well microtiter plate and incubated for 1 h at room temperature in the dark along with 100 µl of the β-galactosidase solution (0.1 mM β-galactosidase substrate in 100 mM sodium phosphate [pH 7.5], 1 mm $MgCl_2$ and 5% Sapphire II Enhancer [Tropix Galactonstar, Bedford, MA, USA]). For the luciferase assay, 80 µl of the lysates were dispensed into a 96-well microtiter plate and luminescence activity was assayed with 1x luciferin solution (0.1 mM coenzyme A, 2.5 mM ATP, 1x luciferin, 5 mM $MgCl_2$ and 500 mM Tris-HCl [pH 7.9]). Measurements were taken using a bioluminometer (GloMax; Promega Corporation). Samples were analyzed in triplicate and experiments were performed a minimum of three times. Data are expressed as a ratio of luciferase activity over β-galactosidase activity and normalized to the empty pA3Luc vector data. Variance calculations and significance were calculated with the one-way ANOVA test.

D. Quantitative PCR

LCLs were harvested at a density of $0.8x10^6$ to $1.2x10^6$ cells/mL every 3-5 days of growth and lysed in Qiazol reagent following manufacturer's instructions (Qiagen). Four to six extracts (corresponding to successive passages) were obtained for each cell line. Total RNA was treated with 2 U of TurboDNAse (Ambion/Thermo Fisher Scientific, Waltham, MA, USA) followed by acid phenol-chloroform precipitation. 1 µg of purified RNA was used for reverse transcription using the QuantiTect Reverse Transcription kit (Qiagen). Approximately 10 ng of total cDNA were used in qPCR assays with the QuantiTect SYBR Green PCR kit (Qiagen). Primers for *GH*, *GHR*, *IGF1*, *BCL2*, *SOCS2* and *B2M* were designed using the Primer 3 software (Primer3 v. 0.4.0; http://bioinfo.ut.ee/primer3-0.4.0/) (**Table IV-S2)**. Normalization and quantification were performed using the comparative Ct method. Data are expressed as fold change ($2^{-\Delta Ct}$) relative to β2M normalizer.

E. Allelic Imbalance Measurements

Our LCL set was genotyped using the Sequenom technology (Genotyping Platform, Génome Québec, Montreal, QC) at rs6180 (A/C) at chromosome 5:42719239 with a MAF=0.42 (C) in the CEU (Utah residents with Northern and Western European ancestry) population (Human assembly GRCh37) (**Table IV-S2**). Only heterozygous cell lines were used for the measurement of allelic expression and homozygous ones were used to test the specificity of each probe. A

series of blanks (n=12) were run to assess the sensitivity and positivity threshold of each probe. Primers and probes suitable for digital PCR were custom-designed for allelic discrimination (IDT, Skokie, IL, USA) (**see Table IV-S2 for the sequences**). For each LCL, cDNA and gDNA were assayed using the BioRad system (example of raw results in **Figure IV-S1**). Each 20 µl reaction contained 1 × ddPCR Supermix for Probes (Bio-Rad), 900 nM specific forward and reverse primers flanking rs6180, 250 nM of specific locked nucleic acid (LNA) probe for each allele at rs6180 and 20 ng of gDNA or cDNA. Each reaction was mixed with 70 µl Droplet Generation Oil (Bio-Rad), partitioned into 12,000–18,000 droplets in the QX100 Droplet Generator (Bio-Rad), transferred to a 96-well plate and sealed. The PCRs were performed in a T100 Thermal Cycler (Bio-Rad) with the following cycling conditions: 1 × (95 °C for 10 min), 45 × (95 °C for 30 s, 60C for 60 s, 72 °C for 1 min) and 1 × (98 °C for 10 min). Following end-point amplification, the fluorescence intensity of individual droplets was measured with the QX100 Droplet Reader (Bio-Rad) and data analysis was performed with QuantaSoft software (Bio-Rad). Positive and negative droplet populations were detected manually by setting a threshold at 3000 (fluorescence intensity of FAM or HEX probes). Technical and biological (different passages) replicates were run and the ratio of the average copies/µl of each allele (cDNA) was normalized to the ratio of the gDNA for each cell line. The normalized ratio was then presented as a percentage of each allele (C:A).

F. Statistical analysis

Statistical significance was assessed using Chi-square tests (**Figure IV-3**), unpaired two-tailed t-tests (**Figure IV-4**), one- or two-way ANOVA followed by Tukey, Dunnett or linear trend (linear regression) post-hoc tests (**Figures IV-4-5, Table IV-1**), using GraphPad Prism v7.0 software (La Jolla, CA, USA). A $p<0.05$ was considered significant. For the student t-tests and Chi-square tests, we applied a correction when the three classes of alleles were tested: the threshold used to assess significance was $p_{corr}<0.025$ (1df) and $p_{corr\text{-}sex}<0.0125$ after stratification by sex; a $p<0.05$ was considered nominally significant.

4. RESULTS

A. Luciferase reporter vector experiments

To investigate the biological significance of the GT repeat polymorphism, we subcloned ~1.2 kb of the Module A region flanking 19, 26 and 33 GT repeats into luciferase reporter vectors to test the ability of S vs. M vs. L repeats to modulate *GHR* transcriptional activity; in addition, we deleted the entire GT repeat sequence to create a GT-deleted reporter vector (**Figure IV-2a**). For the following analyses, we used the GT26(M) vector as the reference construct. We observed that all the constructs had a stimulatory effect under basal conditions; however, there were no significant differences in reporter expression among the different GT lengths in HEK293 cells (**Figure IV-2b**). We next studied the effects of co-transfecting three transcription factors that had response elements flanking the GT repeat: while all three transcription factors increased the expression of reporters containing S, M and L GT repeats, we observed three different response patterns.

First, we overexpressed the DBP transcription factor: a binding site is located ~300bp upstream of the GT repeat and has been previously validated in our lab using ChIP assays (**Figure IV-2a and Figure IV-S1**). At levels of 100ng or 200ng, DBP stimulated all three GT vectors as well as the GT-deleted vector 2 to 3 fold. The overall increase was significant ($F(1, 32) = 15.82$, $p=0.0004$) (**Figure IV-2c**); however, there were no significant differences between the GT lengths at either of the two doses of DBP and no interactions between DBP and the GT genotype.

Next, we tested the effects of overexpressing HIF-1α; two functional response elements had previously been reported by our lab, one in each of the V9 and V3 exons (Erman et al 2011). Following exposure to HIF-1α at 100 and 200ng, there was a dose-related significant stimulatory effect on the different GT constructs ($F(1, 24) = 12.92$, $p=0.0015$) (**Figure IV-2d**). The GT genotype also had an overall significant effect on the differences observed ($F(3, 24) = 4.209$, $p=0.0158$). More specifically, the GT19 construct showed a significant decrease in luciferase transcriptional activation compared to GT26 at the 200ng HIF-1α dose ($p=0.03$).

In 2001 we characterized a functional CCAAT box located immediately 5' to the GT repeat (Goodyer et al 2001) (**Figure IV-1b**). In the present study, treatment with C/EBPβ resulted in a significant dose-related increase in luciferase transactivation for all three GT lengths tested (F (1,

20) = 85.71, p<0.0001) (**Figure IV-2e**). Moreover, the GT genotype had a significant effect on the response pattern observed (F (3, 20) = 9.056, p=0.0005): specifically, the GT-deleted vector showed a significantly higher level of luciferase stimulation than the GT26, indicating a repressor role for the microsatellite in modulating C/EBPß effects (p=0.011 at 4ng). When the GT-deleted vector was used as a reference, it showed significantly higher luciferase induction than the GT19 vector at 2ng C/EBPβ dose (p=0.049) and, at 4ng of C/EBPβ dose, it was significantly higher than all of the GT vectors (GTdel *vs.* GT19 p=0.0009; GTdel *vs.* GT26 p=0.012; GTdel *vs.* GT33 p=0.0095).

Our results showed that the GT allelic length had a small but significant *cis*-regulatory effect and that the GT repeat element could act as a repressor under specific test conditions.

B. Association of differential *GHR* allelic expression with the GT repeat polymorphism

In order to characterize the GT repeat polymorphism using more *in vivo* test conditions, we genotyped more than 400 LCLs in order to obtain a minimum of four for each of the six GT genotypes (total n=41, range: 4-13 LCLs/genotype). LCLs have been extensively used for expression quantitative trait analyses (Pastinen 2010) and constitute an ideal model in which to investigate whether, under basal conditions, the GT repeat microsatellite is a causal variant associated with the expression of each individual *GHR* genotype or allele.

We first assessed the degree of differential *GHR* allelic expression (or allelic imbalance [AI]) using the digital droplet PCR technology. Using the transcribed SNP rs6180 located in exon10, we evaluated the relative amount of each allele in heterozygous individuals (n=21 LCLs); we then normalized the ratio obtained in the cDNA to the ratio observed in the genomic DNA. Allelic differential expression is observed if the ratio of the two alleles differs from the theoretical 50:50. However, even if there is no allelic imbalance, there is variability inherent to the technique. This experimental variability was calculated by measuring the differences between duplicates (maximum 3%) which corresponds to a ratio threshold of 53:47 or 47:53 (dotted blue lines in **Figure IV-3**). A ratio above and below these thresholds indicates a true differential expression between the two copies of the *GHR* gene. We observed significant allelic imbalance at rs6180 for ~86% of the LCLs in our study (**Figure IV-3**). Moreover, among the cell lines showing significant allelic imbalance, a larger number of cell lines derived from males showed a

high degree of AI (above 80:20 and below 20:80) (blue symbols; 57%; p=0.0045). No females (orange symbols) reached this level of AI, even though there was a bias towards a larger number of females in our LCL set (n=13 females *vs.* n=8 males).

It is important to note that the detection SNP (rs6180) was not in linkage disequilibrium with the GT repeat; the microsatellite is located ~300kb upstream, beyond the limit of the rs6180 haplotype block. We, therefore, could not phase the genotypes and, thus, could not directly correlate the allelic class of the GT repeat with a specific allele of the detection SNP. Nevertheless, even without being able to infer a specific haplotype, we could perform association tests (chi-squared) for the individual cell lines classified by GT genotype or by allelic classes presenting with or without the allelic imbalance phenotype directed towards the C or A allele. We did not find a significant association between a specific GT genotype and the allelic differential expression phenotype, likely due to the low number of heterozygous samples/genotype (**Figure IV-3a**). However, when classified by alleles, LCLs derived from male individuals with at least one L allele showed a deviation towards the A allele compared to the non-L cell lines (p=0.047) (**Figure IV-3d**). There was no significant association of the LCLs presenting with either the S or M alleles with a differential allelic expression phenotype (**Figure IV-3b-c**). Interestingly, the profiles for the three alleles suggest changes in the clustering of the LCLs, from a higher proportion of C alleles in **Figure IV-3b** to a higher proportion of A alleles in **Figure IV-3d**.

These endogenous data are in line with the *in vitro* luciferase reporter assay results: they both suggest a *cis*-effect of the GT repeat polymorphism on *GHR* allelic differential expression.

C. GT genotype association with the expression of *GHR* and *GH*/*IGF-1* axis genes in LCLs

We next evaluated whether the polymorphic GT repeat could affect total *GHR* gene expression and, by extension, expression of different gene members of the GH-IGF1 axis. We examined mRNA levels under basal conditions in our LCLs classified according to their GT genotypes (**Figures IV-4 and IV-5**). *GHR* gene expression levels were not associated with any specific GT genotype. However, when organized in an ascending order of *GHR* expression levels, there was a trend for L/M LCLs to have the lowest *GHR* expression compared to the other genotypes and L/S to have the highest mean levels (ANOVA post-test for linear trend: p=0.025, r^2=0.84)

(**Figure IV-4a, Table IV-1**). Interestingly, this trend was driven by LCLs derived from female individuals and not by males (p=0.015, r^2=0.95), as shown by the increased effect size from 13% in both sexes compared to 31% in females (**Table IV-1**). Furthermore, classifying the cell lines according to their GT allelic class (S, M or L) revealed that LCLs presenting with at least one M allele expressed significantly less *GHR* mRNA than the non-M cell lines (p=0.031); and this signal was again driven by the females (p=0.029) (**Figure IV-4c**). No significant differences in *GHR* levels were observed for LCLs presenting with the S or L alleles (**Figures IV-4b** and **4d**). We also tested for correlations between the degree of allelic imbalance in the individual LCLs and the level of *GHR* expression but found no significant effects, suggesting that some compensation mechanism (e.g. *trans*-acting feedback, which would not be measured by AI assays) may be controlling the total *GHR* transcript levels (Pastinen 2010; Sladek and Hudson 2006).

In order to understand if the *GHR* GT repeat genotype could also be associated with expression of different up- and down-stream members of the GH-IGF1 axis, the same analysis was performed for *GH*, *IGF1*, *BCL2* and *SOCS2* (**Figure IV-5**). Initial classification of the genotypes in the same order as for the *GHR* gene analysis did not reveal a linear trend of association of the GT genotypes for any of the four genes. However, when the genotypes were classified in an ascending level of *IGF1* gene expression, there was a highly significant linear trend of association of the GT genotypes with *IGF1* expression (p=0.0005, r^2=0.96). Cells with the L/M genotype had the lowest *IGF1* expression, similar to what was observed with *GHR*, and that association was driven specifically by the males (p=0.0093, r^2=0.71) (**Figure IV-5b**, **Table IV-1**). The L/M cell lines also expressed significantly less *IGF1* than the L/L lines (p=0.023) (**Figure IV-5b**). Keeping the same GT genotype order as for *IGF1* for the other genes resulted in a significant linear trend for *BCL2* (p=0.0007, r^2=0.79) (**Figure IV-5c, Table IV-1**) but not *GH* or *SOCS2* (**Figure IV-5a and IV-5d**). The L/M cell lines showed significantly less expression of the *BCL2* gene compared to the L/L lines (p=0.0047) and this was significantly driven by the females (p=0.012) (**Figure IV-5c**).

These data suggest both a *cis*-effect of the GT repeat polymorphism on *GHR* total expression and a *trans*-effect on *GH-IGF1* gene members, with significant sex-specific differences.

5. DISCUSSION

DNA microsatellites are a class of tandem repeats found throughout the human genome, with the CA/GT motif the most common form amongst the dinucleotide motifs (Ellegren 2004; Sawaya et al. 2013). They've been widely used in genetic mapping and forensics because of their frequent polymorphic nature that is mostly due to slippage during DNA replication (Ellegren 2004). Recently, DNA microsatellites have been shown to accumulate at a high density in promoter regions where they co-localize with *cis* regulatory elements and can act as potential regulators of gene expression (Sawaya et al. 2013). Several studies have shown the association of gene promoter GT repeat polymorphisms with disease risk (e.g. *ER-α*, *NRAMP1*) (Cai et al. 2003; Searle and Blackwell 1999) or disease severity (e.g. *GRIN2A*, *β-ENaC*) (Itokawa et al. 2003a), (Gonzalez et al. 2007), and many of them have shown a modulation of gene expression with the GT length (Gonzalez et al. 2007; Itokawa et al. 2003a; Searle and Blackwell 1999). The HO-1 gene is a well-documented example: its promoter GT microsatellite has been associated with a wide range of clinical disorders, including emphysema, rheumatoid arthritis, type 2 diabetes mellitus, cardiovascular diseases and certain cancers (Bao et al. 2010; Daenen et al. 2016; Kikuchi et al. 2005; Rueda et al. 2007; Tang et al. 2016; Yamada et al. 2000); several of these studies have also reported a modulatory effect of the GT repeat length on HO-1 transcriptional activity (Chen et al. 2016b; Hirai et al. 2003; Yamada et al. 2000).

However, the question of how variability in the length of a promoter microsatellite can modulate gene expression is not often addressed. Several of the mechanisms proposed have been reviewed recently (Bagshaw 2017) and include: a modification of the distance between functional elements located outside of the short tandem repeat (Chen et al. 2016a) with a reorganization of the local chromatin conformation mediated by a DNA looping process (Chen et al. 2016b); a modulation of transcription factor binding (Bayele et al. 2007; Taka et al. 2013); and a potential role in genome organization via long range chromatin interactions (Nikumbh and Pfeifer 2017). These effects may also be complicated by the fact that there is a high propensity of the CA/GT microsatellites to form Z-DNA, the left handed form of B-DNA. This transient conformation occurs mostly during replication or transcription: in the latter case, the unwinding of the DNA behind the RNA polymerase II (RNApolII) generates negative supercoiling and stabilizes the Z-DNA conformation (Rich and Zhang 2003).

In previous studies, we showed that a GT dinucleotide repeat, located ~80bp upstream of the transcription start site of the V9 5'UTR exon in Module A of the *GHR* gene, was polymorphic in the general population (Hadjiyannakis et al. 2001). Subsequent case-control analyses of cohorts of idiopathic short stature children showed an association of specific GT genotypes with the short stature phenotype, specifically in boys (Dias et al. 2017). Most recently, we found that obesity indices, like BMI (body mass index) and FMI (fat mass index), have confounding effects, strengthening the association of the GT repeat polymorphism with adult severe short stature, only in women (**Chapter III**). The aim of the present study was to characterize the potential mechanisms by which this GT microsatellite may be modulating GHR gene transcription. Initially, we used luciferase reporter assays to identify the modulatory potential of the GT repeat polymorphism. The Module A region of *GHR* contains three exons that ubiquitously express the most abundant *GHR* mRNA variants in human tissues (Goodyer et al. 2001c). Multiple *cis* factor response elements in their promoter regions have been investigated in our lab, revealing their influence on *GHR* transcriptional activity (Erman et al. 2011b; Kenth et al. 2011; Wei et al. 2009a). The regions immediately flanking the GT repeat in our luciferase reporter constructs contain response elements for three different cues relevant to GHR function: diurnal rhythms (DBP), adipose hypoxic environment (HIF-1α) and adipogenic and chondrogenic processes (C/EBPβ).

Under basal test conditions, all of the constructs exerted a stimulatory effect on *luciferase* transcription but there were no differences amongst the GT19, GT26, GT33 and GTdel groups. This result led us to hypothesize that the GT repeat microsatellite may regulate promoter activity only in response to specific signals. First we tested DBP, a circadian rhythm transcription factor, as *GHR* mRNA has been shown to have a circadian rhythmicity in murine calvarial bone, liver and skeletal muscle (Itoh et al. 2004; Zvonic et al. 2007). Our results showed that DBP was, indeed, an inducer of *luciferase* transcriptional activity in our test system. However, the effect was not regulated by the GT polymorphism, possibly since the DBP binding site is located ~300bp upstream from the GT repeat element. Next we tested HIF-1α. This transcription factor is present in adipose tissue during obesity (Ye 2009) and *GHR* mRNA levels in mature human adipose cells have been shown to increase in response to HIF-1α in a dose-related fashion (Erman et al. 2011b). In the present study, we found that, while all of the constructs were stimulated by HIF1-α, the shortest allele showed a significantly lower response. The closest response element to the microsatellite is a

CCAAT box immediately 5'. C/EBPβ has been shown to be involved in several processes that also involve the GH/GHR/IGF1 axis, including adipogenesis and chondrogenesis (Tsukada et al. 2011). Our data show that the C/EBPβ response was not affected by GT allele length; however, a reporter with no GT repeats (GTdel) had a significantly greater response to C/EBPβ, suggesting a repressor role for the microsatellite. Thus, we demonstrated that our GT polymorphic microsatellite could act as a modulator of *GHR* promoter activity in response to specific cues. Our data also suggest that the positioning of the response elements relative to the microsatellite may be partly responsible for the differential effects observed.

Reporter construct studies are limited due to their use of a restricted region of the promoter that may not represent the folded state of native chromatin. Therefore, in order to understand if the GT repeat exerts *cis*-regulatory effects within its original genomic context, we also evaluated its activity using a cell model. We chose to study LCLs from the CEPH collection to ensure finding as many different GT genotypes as possible, including the rare S/S and L/L ones, in order to evaluate if the GT repeat polymorphism could be an expression quantitative trait locus (eQTL).

Allelic specific expression is a method to evaluate the effect of a potential *cis* regulator on the relative expression of two alleles within the same sample. The strength of the technique relies on the fact that two alleles measured at the same *locus* become a control for each other, removing the need for a reference gene and neutralizing the effect of confounding factors from the cell context (*trans*-acting) (Lo et al. 2003; Pastinen and Hudson 2004; Serre et al. 2008; Yan et al. 2002). Several studies have reported that allelic differential expression is a common phenomenon in the human genome (Ge et al. 2009; Lo et al. 2003; Pastinen 2010): the expression of 30% of the genes in LCLs is associated with common *cis* variants (Ge et al. 2009), and 40-60% of these *cis*-regulatory variants are shared between cell types (Adoue et al. 2014). Genes showing allelic differential expression are implicated in many biological pathways (Palacios et al. 2009) and have been associated with a predisposition to certain diseases (de la Chapelle 2009; Wang et al. 2016). A number of techniques have been developed to measure allelic differential expression, including qPCR (Sun et al. 2010), quantitative sequencing (Ge et al. 2005), single-base extension methods like SNUpe or SNApSHOT, pyrosequencing (Wang and Elbein 2007), RNA sequencing (Pastinen 2010), mass-spectrometry (Wang et al. 2016) and array-based methods (Ge et al. 2009; Ronald et al. 2005; Serre et al. 2008); most of these techniques are able to detect changes in allelic expression

of more than 1.5 fold (60:40 ratio).

In this report, we developed a new application for the droplet digital PCR assay to measure allele specific expression. Combining the sensitivity of the droplet-based PCR and highly specific LNA probes, we were able to discriminate allelic *GHR* expression in LCLs with high sensitivity. We observed a high prevalence of allelic imbalance in *GHR* at the rs6180 locus (~86%); even after taking a more common threshold for AI at 60:40, the prevalence remained elevated (>70%). Interestingly, only the males in our set of LCLs showed a high degree of AI (>80:20) and, notably, two male individuals showed extreme allelic imbalance (>95:5). The high degree of AI in LCLs has already been reported although sex differences were not examined (Ge et al. 2009; Light et al. 2014; Serre et al. 2008). Differential allelic expression in *GHR* has only been assessed in one previous study. In 2010, Verlaan et al. performed high-throughput allelic expression measurements in LCLs and osteoblast cell lines (HObs) using rs6180 and intronic detection SNPs to map *cis*-regulatory variants more efficiently in the human genome; the *GHR* gene was part of an 81 gene set presenting with significant AI in both cell types (Verlaan et al. 2009).

Although our cell-based approach gave only nominally significant evidence of a *cis* regulatory effect of the GT repeat on *GHR* allelic expression, it specifically showed that male individuals carrying a long version of the GT repeat allele in their genotype had an effect on the AI phenotype and potentially on the degree of the imbalances observed. In contrast, the examination of total *GHR* mRNA expression levels for the full set of LCLs showed a significant linear trend, with the L/M specific GT genotype correlating with the lowest levels of *GHR*. Although no significant differences between the GT genotype categories were observed using ANOVA and post-hoc tests, classifying the cell lines by alleles showed that those carrying at least one M allele had significantly lower average expression levels of *GHR* mRNA than non-M cell lines, suggesting a *cis* regulatory effect of the GT repeat genotype on *GHR* expression.

To understand if the GT repeat polymorphism has potential *trans*-acting effects within the GH-IGF1 axis, we tested its association with mRNA levels of *GH* and three GH-target genes. We observed a significant association with *IGF1* and *BCL2*, but not *GH* or *SOCS2*, gene expression; for both *IGF1* and *BCL2*, there was evidence for significant sexual dimorphism as we observed for *GHR*. Lymphocytes in healthy individuals express both GH and GHR, with B cells having the highest levels; there is also considerable individual variability (Hattori et al. 2001). GH has been reported

to exert a sex-specific secretory pattern in humans and in rodents (Lichanska and Waters 2008a). Recent studies in the rodent liver described sets of 'male-biased' and 'female-biased' genes that had chromatin modifications in response to pulsatile (male) *vs.* continuous (female) serum GH levels (Connerney et al. 2017; Lau-Corona et al. 2017). Although our LCLs were not treated with exogenous GH, basal levels of GH produced endogenously by B cells (Weigent and Blalock 1989) may be sufficient for an autocrine stimulation leading to the sex-specific transcriptional differences that we observed for *GHR*, *IGF1* and *BCL2*.

In addition, LCLs carrying the L/M genotype consistently showed the lowest *GHR*, *IGF1* and *BCL2* expression, suggesting a continuum of the effect of this genotype downstream of GHR. *GHR/IGF1/BCL2* could, thus, constitute a transcriptional unit co-regulated by the L/M GT genotype regulatory variant. This *trans* effect may be partly explained by the property of this specific dinucleotide GT repeat to form Z-DNA; *in silico* analyses using three different prediction algorithms, Z-Hunt (Ho et al. 1986), SIBZ (Zhabinskaya and Benham 2011) and the nonB-DB database (Cer et al. 2013), provide evidence for this (**Figure IV-S3**). Several genes have been shown to be regulated by a Z-DNA forming dinucleotide repeat in their promoter regions, including HO-1, ADAM12, SLC11A1 and c-myc ((Bayele et al. 2007; Ray et al. 2011; Wittig et al. 1992; Zhang et al. 2006). Z-DNA binding proteins (ZBPs), like ADAR-1, can specifically recognize Z-DNA conformations (Herbert et al. 1995) and cause them to act as enhancer elements (Oh et al. 2002). Interestingly, LCLs have the highest levels of ADAR-1 when compared to 53 different tissues in the GTEx database (www.gtexportal.org). It is possible that the amount of ZBP binding could be proportional to the length of the GT allele and that increased Z-DNA stability may have effects both (a**)** in *cis*, on *GHR* expression levels by regulating the entrance of RNApolII and altering the formation of core transcriptional units; and (b) in *trans*, as Z-DNA formation excludes the replacement of the nucleosomes, implying a local chromatin reorganization (Garner and Felsenfeld 1987; Liu et al. 2006). The ZBPs may also constitute bridges to potential Z-forming elements located in the promoter regions of other genes, such as the polymorphic CA repeat in the IGF1 gene promoter. Recent studies using genome-wide Chromatin Interaction Analysis with Paired-End-Tag sequencing (ChIA-PET) found that mapping long range interactions with RNApolII uncovered a set of genes that cooperated in their transcriptional activity through promoter-promoter interactions (Li et al. 2012). Moreover, short tandem repeats, specifically with GT motifs, have been shown to be implicated in the 3D organization of the human genome with

potential long range DNA-DNA interactions (Nikumbh and Pfeifer 2017).

Certain limitations of the present study should be noted. Our observation of an association between the male LCL carriers of at least one GT L allele with a deviation towards the A variant of rs6180 is intriguing but only of nominal significance. This is likely due to both the small number of cell lines carrying the L allele and the need to work only with cell lines heterozygous at the detection SNP. We had chosen rs6180 because of its high MAF (>40%) in order to have as large a group of LCLs as possible to test but were still limited in the number of rarer GT genotypes available. The other limitation is the absence of LD between the GT repeat and rs6180, making it impossible to infer a regulatory haplotype. Taking a SNP in the V2, V9 or V3 first exons would have been ideal: unfortunately, these regions contain a very small number of SNPs and are rich in CpG islands which meant that we could not design proper probes. Another option would have been to use intronic detection SNPs (Verlaan et al. 2009); however, heteronuclear non-spliced RNA is present at very low concentrations in the nucleus and many of the LCLs had low total *GHR* mRNA levels making this approach unfeasible.

To conclude, we have demonstrated for the first time the functional relevance of a GT repeat polymorphism in a core promoter region of the *GHR* gene. As with many other promoter microsatellites, this GT repeat has properties of a fine-tuning *cis* modulator of *GHR* expression in a context-specific manner. Its potential *trans* effects need further experimentation to better understand the underlying mechanisms that may shed new light on the complex regulation of the *GHR* promoter and its effects on transcriptional targets of GH signalling.

ACKNOWLEDGEMENTS: We thank Dr. Gurvinder Kenth for the DBP ChIP data, Dr. Aimee Ryan for her review of the manuscript, Dr. Michael Ohh (University of Toronto) for providing the HIF-1α expression vector, Dr. Ronald Hines (Medical College of Wisconsin) for the DBP expression vector and Dr. Elly Holthuizen for the C/EPBβ expression vector.

FUNDING: This project was funded by the Canadian Institutes for Health Research (CGG MOP119313). CD received doctoral studentships from the Montreal Children's Hospital Research Institute and the Fonds de recherche du Québec-Santé (FRQ-S). This research was

undertaken at the Research Institute of the McGill University Health Centre which is supported in part by the FRQ-S.

DISCLOSURE: No conflicts of interest, financial or otherwise, are declared by the authors.

6. REFERENCES

Adoue V, Schiavi A, Light N, Almlof JC, Lundmark P, Ge B, Kwan T, Caron M, Ronnblom L, Wang C et al. . 2014. Allelic expression mapping across cellular lineages to establish impact of non-coding SNPs. Mol Syst Biol 10:754.

Bagshaw ATM. 2017. Functional Mechanisms of Microsatellite DNA in Eukaryotic Genomes. Genome Biol Evol 9(9):2428-2443.

Bao W, Song F, Li X, Rong S, Yang W, Wang D, Xu J, Fu J, Zhao Y, and Liu L. 2010. Association between heme oxygenase-1 gene promoter polymorphisms and type 2 diabetes mellitus: a HuGE review and meta-analysis. Am J Epidemiol 172(6):631-636.

Barton DE, Foellmer BE, Wood WI, and Francke U. 1989. Chromosome mapping of the growth hormone receptor gene in man and mouse. Cytogenet Cell Genet 50(2-3):137-141.

Bayele HK, Peyssonnaux C, Giatromanolaki A, Arrais-Silva WW, Mohamed HS, Collins H, Giorgio S, Koukourakis M, Johnson RS, Blackwell JM et al. . 2007. HIF-1 regulates heritable variation and allele expression phenotypes of the macrophage immune response gene SLC11A1 from a Z-DNA forming microsatellite. Blood 110(8):3039-3048.

Ben-Shlomo A, and Melmed S. 2008. Acromegaly. Endocrinol Metab Clin North Am 37(1):101-122, viii.

Bhargava A, and Fuentes FF. 2010. Mutational dynamics of microsatellites. Mol Biotechnol 44(3):250-266.

Brooks AJ, and Waters MJ. 2010. The growth hormone receptor: mechanism of activation and clinical implications. Nat Rev Endocrinol 6(9):515-525.

Cai Q, Gao YT, Wen W, Shu XO, Jin F, Smith JR, and Zheng W. 2003. Association of breast cancer risk with a GT dinucleotide repeat polymorphism upstream of the estrogen receptor-alpha gene. Cancer Res 63(18):5727-5730.

Cer RZ, Donohue DE, Mudunuri US, Temiz NA, Loss MA, Starner NJ, Halusa GN, Volfovsky N, Yi M, Luke BT et al. . 2013. Non-B DB v2.0: a database of predicted non-B DNA-forming motifs and its associated tools. Nucleic Acids Res 41(Database issue):D94-D100.

Chen HY, Ma SL, Huang W, Ji L, Leung VH, Jiang H, Yao X, and Tang NL. 2016a. The mechanism of transactivation regulation due to polymorphic short tandem repeats (STRs) using IGF1 promoter as a model. Sci Rep 6:38225.

Chen WJ, Chen YH, Lai YJ, Hsu YJ, Yeh YH, Tsai CS, and Lin CY. 2016b. GT-repeat length polymorphism in heme oxygenase-1 promoter determines the effect of cilostazol on vascular smooth muscle cells. Int J Cardiol 222:407-415.

Connerney J, Lau-Corona D, Rampersaud A, and Waxman DJ. 2017. Activation of Male Liver Chromatin Accessibility and STAT5-Dependent Gene Transcription by Plasma Growth Hormone Pulses. Endocrinology 158(5):1386-1405.

Daenen KE, Martens P, and Bammens B. 2016. Association of HO-1 (GT)n Promoter Polymorphism and Cardiovascular Disease: A Reanalysis of the Literature. Can J Cardiol 32(2):160-168.

de la Chapelle A. 2009. Genetic predisposition to human disease: allele-specific expression and low-penetrance regulatory loci. Oncogene 28(38):3345-3348.

Dias C, Giordano M, Frechette R, Bellone S, Polychronakos C, Legault L, Deal CL, and Goodyer CG. 2017. Genetic variations at the human growth hormone receptor (GHR) gene locus are associated with idiopathic short stature. J Cell Mol Med 21(11):2985-2999.

Ellegren H. 2004. Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5(6):435-445.

Erman A, Wabitsch M, and Goodyer CG. 2011. Human growth hormone receptor (GHR) expression in obesity: II. Regulation of the human GHR gene by obesity-related factors. Int J Obes (Lond) 35(12):1520-1529.

Garner MM, and Felsenfeld G. 1987. Effect of Z-DNA on nucleosome placement. J Mol Biol 196(3):581-590.

Ge B, Gurd S, Gaudin T, Dore C, Lepage P, Harmsen E, Hudson TJ, and Pastinen T. 2005. Survey of allelic expression using EST mining. Genome Res 15(11):1584-1591.

Ge B, Pokholok DK, Kwan T, Grundberg E, Morcos L, Verlaan DJ, Le J, Koka V, Lam KC, Gagne V et al. . 2009. Global patterns of cis variation in human cells revealed by high-density allelic expression analysis. Nat Genet 41(11):1216-1222.

Godowski PJ, Leung DW, Meacham LR, Galgani JP, Hellmiss R, Keret R, Rotwein PS, Parks JS, Laron Z, and Wood WI. 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(20):8083-8087.

Gonzalez AA, Carvajal CA, Riquelme E, Krall PM, Munoz CR, Mosso LM, Kalergis AM, and Fardella CE. 2007. A polymorphic GT short tandem repeat affecting beta-ENaC mRNA expression is associated with low renin essential hypertension. Am J Hypertens 20(7):800-806.

Goodyer CG, Zogopoulos G, Schwartzbauer G, Zheng H, Hendy GN, and Menon RK. 2001. Organization and evolution of the human growth hormone receptor gene 5'-flanking region. Endocrinology 142(5):1923-1934.

Hadjiyannakis S, Zheng H, Hendy GN, and Goodyer CG. 2001. GT repeat polymorphism in the 5' flanking region of the human growth hormone receptor gene. Mol Cell Probes 15(4):239-242.

Hattori N, Saito T, Yagyu T, Jiang BH, Kitagawa K, and Inagaki C. 2001. GH, GH receptor, GH secretagogue receptor, and ghrelin expression in human T cells, B cells, and neutrophils. J Clin Endocrinol Metab 86(9):4284-4291.

Herbert A, Lowenhaupt K, Spitzner J, and Rich A. 1995. Chicken double-stranded RNA adenosine deaminase has apparent specificity for Z-DNA. Proc Natl Acad Sci U S A 92(16):7550-7554.

Hirai H, Kubo H, Yamaya M, Nakayama K, Numasaki M, Kobayashi S, Suzuki S, Shibahara S, and Sasaki H. 2003. Microsatellite polymorphism in heme oxygenase-1 gene promoter is associated with susceptibility to oxidant-induced apoptosis in lymphoblastoid cell lines. Blood 102(5):1619-1621.

Ho PS, Ellison MJ, Quigley GJ, and Rich A. 1986. A computer aided thermodynamic approach for predicting the formation of Z-DNA in naturally occurring sequences. EMBO J 5(10):2737-2744.

Itoh E, Iida K, del Rincon JP, Kim DS, and Thorner MO. 2004. Diurnal variation in growth hormone receptor messenger ribonucleic acid in liver and skeletal muscle of lit/+ and lit/lit mice. Endocr J 51(6):529-535.

Itokawa M, Yamada K, Iwayama-Shigeno Y, Ishitsuka Y, Detera-Wadleigh S, and Yoshikawa T. 2003. Genetic analysis of a functional GRIN2A promoter (GT)n repeat in bipolar disorder pedigrees in humans. Neurosci Lett 345(1):53-56.

Kenth G, Mergelas JA, and Goodyer CG. 2008. Developmental changes in the human GH receptor and its signal transduction pathways. J Endocrinol 198(1):71-82.

Kenth G, Puzhko S, and Goodyer CG. 2011. Human growth hormone receptor gene expression is regulated by Gfi-1/1b and GAGA cis-elements. Mol Cell Endocrinol 335(2):135-147.

Kikuchi A, Yamaya M, Suzuki S, Yasuda H, Kubo H, Nakayama K, Handa M, Sasaki T, Shibahara S, Sekizawa K et al. . 2005. Association of susceptibility to the development of lung adenocarcinoma with the heme oxygenase-1 gene promoter polymorphism. Hum Genet 116(5):354-360.

Lau-Corona D, Suvorov A, and Waxman DJ. 2017. Feminization of male mouse liver by persistent growth hormone stimulation: Activation of sex-biased transcriptional networks and dynamic changes in chromatin states. Mol Cell Biol.

Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, Poh HM, Goh Y, Lim J, Zhang J et al. . 2012. Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. Cell 148(1-2):84-98.

Lichanska AM, and Waters MJ. 2008. How growth hormone controls growth, obesity and sexual dimorphism. Trends Genet 24(1):41-47.

Light N, Adoue V, Ge B, Chen SH, Kwan T, and Pastinen T. 2014. Interrogation of allelic chromatin states in human cells by high-density ChIP-genotyping. Epigenetics 9(9):1238-1251.

Liu H, Mulholland N, Fu H, and Zhao K. 2006. Cooperative activity of BRG1 and Z-DNA formation in chromatin remodeling. Mol Cell Biol 26(7):2550-2559.

Lo HS, Wang Z, Hu Y, Yang HH, Gere S, Buetow KH, and Lee MP. 2003. Allelic variation in gene expression is common in the human genome. Genome Res 13(8):1855-1862.

Nikumbh S, and Pfeifer N. 2017. Genetic sequence-based prediction of long-range chromatin interactions suggests a potential role of short tandem repeat sequences in genome organization. BMC Bioinformatics 18(1):218.

Oh DB, Kim YG, and Rich A. 2002. Z-DNA-binding proteins can act as potent effectors of gene expression in vivo. Proc Natl Acad Sci U S A 99(26):16666-16671.

Orlovskii IV, Sverdlova PS, and Rubtsov PM. 2004. [Fine structure, expression and polymorphism of the human growth hormone receptor gene]. Mol Biol (Mosk) 38(1):29-39.

Palacios R, Gazave E, Goni J, Piedrafita G, Fernando O, Navarro A, and Villoslada P. 2009. Allele-specific gene expression is widespread across the genome and biological processes. PLoS One 4(1):e4150.

Pastinen T. 2010. Genome-wide allele-specific analysis: insights into regulatory variation. Nat Rev Genet 11(8):533-538.

Pastinen T, and Hudson TJ. 2004. Cis-acting regulatory variation in the human genome. Science 306(5696):647-650.

Pekhletsky RI, Chernov BK, and Rubtsov PM. 1992. Variants of the 5'-untranslated sequence of human growth hormone receptor mRNA. Mol Cell Endocrinol 90(1):103-109.

Ray BK, Dhar S, Shakya A, and Ray A. 2011. Z-DNA-forming silencer in the first exon regulates human ADAM-12 gene expression. Proc Natl Acad Sci U S A 108(1):103-108.

Rich A, and Zhang S. 2003. Timeline: Z-DNA: the long road to biological function. Nat Rev Genet 4(7):566-572.

Ronald J, Akey JM, Whittle J, Smith EN, Yvert G, and Kruglyak L. 2005. Simultaneous genotyping, gene-expression measurement, and detection of allele-specific expression with oligonucleotide arrays. Genome Res 15(2):284-291.

Rowlinson SW, Yoshizato H, Barclay JL, Brooks AJ, Behncken SN, Kerr LM, Millard K, Palethorpe K, Nielsen K, Clyde-Smith J et al. . 2008. An agonist-induced conformational change in the growth hormone receptor determines the choice of signalling pathway. Nat Cell Biol 10(6):740-747.

Rueda B, Oliver J, Robledo G, Lopez-Nevot MA, Balsa A, Pascual-Salcedo D, Gonzalez-Gay MA, Gonzalez-Escribano MF, and Martin J. 2007. HO-1 promoter polymorphism associated with rheumatoid arthritis. Arthritis Rheum 56(12):3953-3958.

Savage MO, Attie KM, David A, Metherell LA, Clark AJ, and Camacho-Hubner C. 2006. Endocrine assessment, molecular characterization and treatment of growth hormone insensitivity disorders. Nat Clin Pract Endocrinol Metab 2(7):395-407.

Sawaya S, Bagshaw A, Buschiazzo E, Kumar P, Chowdhury S, Black MA, and Gemmell N. 2013. Microsatellite tandem repeats are abundant in human promoters and are associated with regulatory elements. PLoS One 8(2):e54710.

Searle S, and Blackwell JM. 1999. Evidence for a functional repeat polymorphism in the promoter of the human NRAMP1 gene that correlates with autoimmune versus infectious disease susceptibility. J Med Genet 36(4):295-299.

Serre D, Gurd S, Ge B, Sladek R, Sinnett D, Harmsen E, Bibikova M, Chudin E, Barker DL, Dickinson T et al. . 2008. Differential allelic expression in the human genome: a robust approach to identify genetic and epigenetic cis-acting mechanisms regulating gene expression. PLoS Genet 4(2):e1000006.

Sladek R, and Hudson TJ. 2006. Elucidating cis- and trans-regulatory variation using genetical genomics. Trends Genet 22(5):245-250.

Sun C, Southard C, Witonsky DB, Olopade OI, and Di Rienzo A. 2010. Allelic imbalance (AI) identifies novel tissue-specific cis-regulatory variation for human UGT2B15. Hum Mutat 31(1):99-107.

Taka S, Gazouli M, Politis PK, Pappa KI, and Anagnou NP. 2013. Transcription factor ATF-3 regulates allele variation phenotypes of the human SLC11A1 gene. Mol Biol Rep 40(3):2263-2271.

Tang D, Tang WJ, Shi XL, Li WP, Zhou H, Lu LM, and Tao L. 2016. Association of the microsatellite (GT)n repeat polymorphisms of the HO-1 gene promoter and corresponding serum levels with the risk of laryngeal squamous cell carcinoma. Acta Otolaryngol 136(8):806-811.

Tremblay JJ, Goodyer CG, and Drouin J. 2000. Transcriptional properties of Ptx1 and Ptx2 isoforms. Neuroendocrinology 71(5):277-286.

Tsukada J, Yoshida Y, Kominato Y, and Auron PE. 2011. The CCAAT/enhancer (C/EBP) family of basic-leucine zipper (bZIP) transcription factors is a multifaceted highly-regulated system for gene regulation. Cytokine 54(1):6-19.

Verlaan DJ, Ge B, Grundberg E, Hoberman R, Lam KC, Koka V, Dias J, Gurd S, Martin NW, Mallmin H et al. . 2009. Targeted screening of cis-regulatory variation in human haplotypes. Genome Res 19(1):118-127.

Wang H, and Elbein SC. 2007. Detection of allelic imbalance in gene expression using pyrosequencing. Methods Mol Biol 373:157-176.

Wang Y, Cao Y, Huang X, Yu T, Wei Z, McGrath J, Xu F, Bi Y, Li X, Yang F et al. . 2016. Allele-specific expression of mutated in colorectal cancer (MCC) gene and alternative susceptibility to colorectal cancer in schizophrenia. Sci Rep 6:26688.

Wei Y, Puzhko S, Wabitsch M, and Goodyer CG. 2009. Structure and activity of the human growth hormone receptor (hGHR) gene V2 promoter. Mol Endocrinol 23(3):360-372.

Wei Y, Rhani Z, and Goodyer CG. 2006. Characterization of growth hormone receptor messenger ribonucleic acid variants in human adipocytes. J Clin Endocrinol Metab 91(5):1901-1908.

Weigent DA, and Blalock JE. 1989. Expression of growth hormone by lymphocytes. Int Rev Immunol 4(3):193-211.

Wittig B, Wolfl S, Dorbic T, Vahrson W, and Rich A. 1992. Transcription of human c-myc in permeabilized nuclei is associated with formation of Z-DNA in three discrete regions of the gene. EMBO J 11(12):4653-4663.

Yamada N, Yamaya M, Okinaga S, Nakayama K, Sekizawa K, Shibahara S, and Sasaki H. 2000. Microsatellite polymorphism in the heme oxygenase-1 gene promoter is associated with susceptibility to emphysema. Am J Hum Genet 66(1):187-195.

Yan H, Yuan W, Velculescu VE, Vogelstein B, and Kinzler KW. 2002. Allelic variation in human gene expression. Science 297(5584):1143.

Ye J. 2009. Emerging role of adipose tissue hypoxia in obesity and insulin resistance. Int J Obes (Lond) 33(1):54-66.

Zhabinskaya D, and Benham CJ. 2011. Theoretical analysis of the stress induced B-Z transition in superhelical DNA. PLoS Comput Biol 7(1):e1001051.

Zhang J, Ohta T, Maruyama A, Hosoya T, Nishikawa K, Maher JM, Shibahara S, Itoh K, and Yamamoto M. 2006. BRG1 interacts with Nrf2 to selectively mediate HO-1 induction in response to oxidative stress. Mol Cell Biol 26(21):7942-7952.

Zvonic S, Ptitsyn AA, Kilroy G, Wu X, Conrad SA, Scott LK, Guilak F, Pelled G, Gazit D, and Gimble JM. 2007. Circadian oscillation of gene expression in murine calvarial bone. J Bone Miner Res 22(3):357-365.

Table IV-1: GT repeat genotype associations with *GH, GHR, IGF1, BCL2* and *SOCS2* mRNA expression stratified by sex.

Gene ID	Sex	ANOVA	Linear Trend		
		P-value	P-value	r^2	Effect size
GH	B	0.14			
	M	0.97			
	F	0.072			
GHR	B	0.281	0.025	0.84	0.13
	M	0.89			
	F	0.23	0.015	0.95	0.31
IGF-1	B	0.023	0.0005	0.96	0.29
	M	0.077	0.0093	0.71	0.36
	F	0.47			
BCL-2	B	0.0047	0.0007	0.79	0.26
	M	0.39	0.043	0.88	0.27
	F	0.012	0.0008	0.87	0.48
SOCS-2	B	0.67			
	M	0.47			
	F	0.91			

Statistical analyses were conducted using the one-way ANOVA test followed by Tukey and linear trend (linear regression) post-hoc tests. B = both sexes, M = males, F = females.

FIGURE LEGENDS

Figure IV-1: Schematic of the human *GHR* gene: (a) The *GHR* gene is located on the short arm of chromosome 5. Exons 2-10 code for the protein. Thirteen non-coding exons have been reported within the 150kb upstream of exon 2 in the 5' *UTR.* Seven of the non-coding exons are clustered in 2 small regions defined as Module A (~1.6kb) and Module B (~2kb). VA-VD and V3a/b/VE exons are found between the two Modules. V5 is located adjacent to the first coding exon, exon 2. (Goodyer et al. 2001a; Orlovskii et al. 2004; Pekhletsky et al. 1992; Wei et al. 2006) **(b)** Promoter regions of the three major ubiquitously expressing human *GHR* 5′UTR exons in Module A with putative response elements for DBP (D site binding protein), HIF-1α (hypoxia-inducible factor-1α) and C/EBPβ (CCAAT enhancer-binding protein β) transcription factors.

Figure IV-2: Effects of the GT repeat polymorphism on *GHR* gene promoter activity. HEK293 cells were co-transfected with a β-galactosidase vector, one of four *GHR*-luciferase promoter reporter vectors containing different lengths of the GT repeat (19(S), 26(M), 33(L) or GTdel), and different transcription factor expression vectors. **(a)** A schematic of the 1.2 kb GHR promoter reporter construct shows the location of response elements for transcription factors characterized in this study. Luciferase and β-galactosidase activities were assayed 48 h after transfection. Luciferase activity was normalized using pA3Luc backbone data and corrected for transfection efficiency with β-galactosidase data. Effects of the different GT lengths on luciferase activity were measured **(b)** with no stimulation (n=22) or following stimulation by **(c)** DBP, at doses of 100 (n=5) and 200 ng (n=11), **(d)** HIF-1α, at 100 or 200 ng (n=4), or **(e)** C/EBPβ, at 2 (n=4) or 4 ng (n=3). All data are presented as mean ± SEM. Significant differences were assessed by one- or two-way ANOVA followed by Tukey multiple comparisons tests with GT26 as the reference (*$p<0.05$, p** $p<0.01$, ***$p<0.001$, **** $p<0.0001$).

Figure IV-3: Differential allelic expression association with GT repeat length at the *GHR* gene locus in LCLs. Twenty-one LCL cell lines heterozygous at rs6180 (located in exon 10 of *GHR*) have been analyzed using digital PCR with specific probes to discriminate between the C and A alleles. Average allelic ratios (of two consecutive passages) normalized by the gDNA

allelic ratios are displayed for each cell line as a percentage (orange symbols = females, blue = males). The allelic imbalance threshold was estimated at 53:47 or 47:53 (dotted lines) based on experimental variability. A high degree of allelic imbalance (80:20) was observed more significantly in cell lines derived from males than females (p=0.0008) independent of the GT genotype or allelic class. Upper panel **(a)** showed no significant association between the GT genotypes and the allelic ratios. However, classification of the cell lines according to their GT allelic classes revealed a significant difference for **(d)** the L allele but not for the S and M alleles **(b-c)**: there were significantly more L cell lines derived from males expressing lower levels of the rs6180 C allele compared to non-L ones **(d)** (p=0.047). Statistical analyses were carried out using the chi^2 test.

Figure IV-4: GT genotype and allelic associations with *GHR* mRNA expression in LCLs. Forty one LCLs classified **(a)** according to their GT genotype or **(b-d)** GT allelic class were grown for 3-5 days to a concentration of $0.7\text{-}1.2x10^6$ cells/mL for six consecutive passages. Total *GHR* gene expression was assessed by quantitative PCR for a minimum of 3 passages for each cell line. Results are presented as mean ± SEM; orange symbols = females, blue = males. Statistics were carried out using one-way ANOVA followed by Tukey's post-hoc test for multiple comparisons or an unpaired 2-tailed t-test (* $p<0.038$).

Figure IV-5: GT genotype and allelic associations with *GH, IGF1, BCL2 and SOCS2* mRNA expression in LCLs. Forty-one LCLs classified according to their GT genotype were grown for 3-5 days to a concentration of $0.7\text{-}1.2x10^6$ cells/mL and RNA was harvested for six consecutive passages. Total gene expression of **(a)** *GH*, **(b)** *IGF1*, **(c)** *BCL2* and **(d)** *SOCS2* gene expression was assessed by quantitative PCR for a minimum of 3 passages for each cell line. Results are presented as mean ± SEM; orange symbols = females, blue = males. Statistics were carried out using one-way ANOVA followed by Tukey's post-hoc multiple comparisons tests (* $p<0.02$, ** $p<0.0047$).

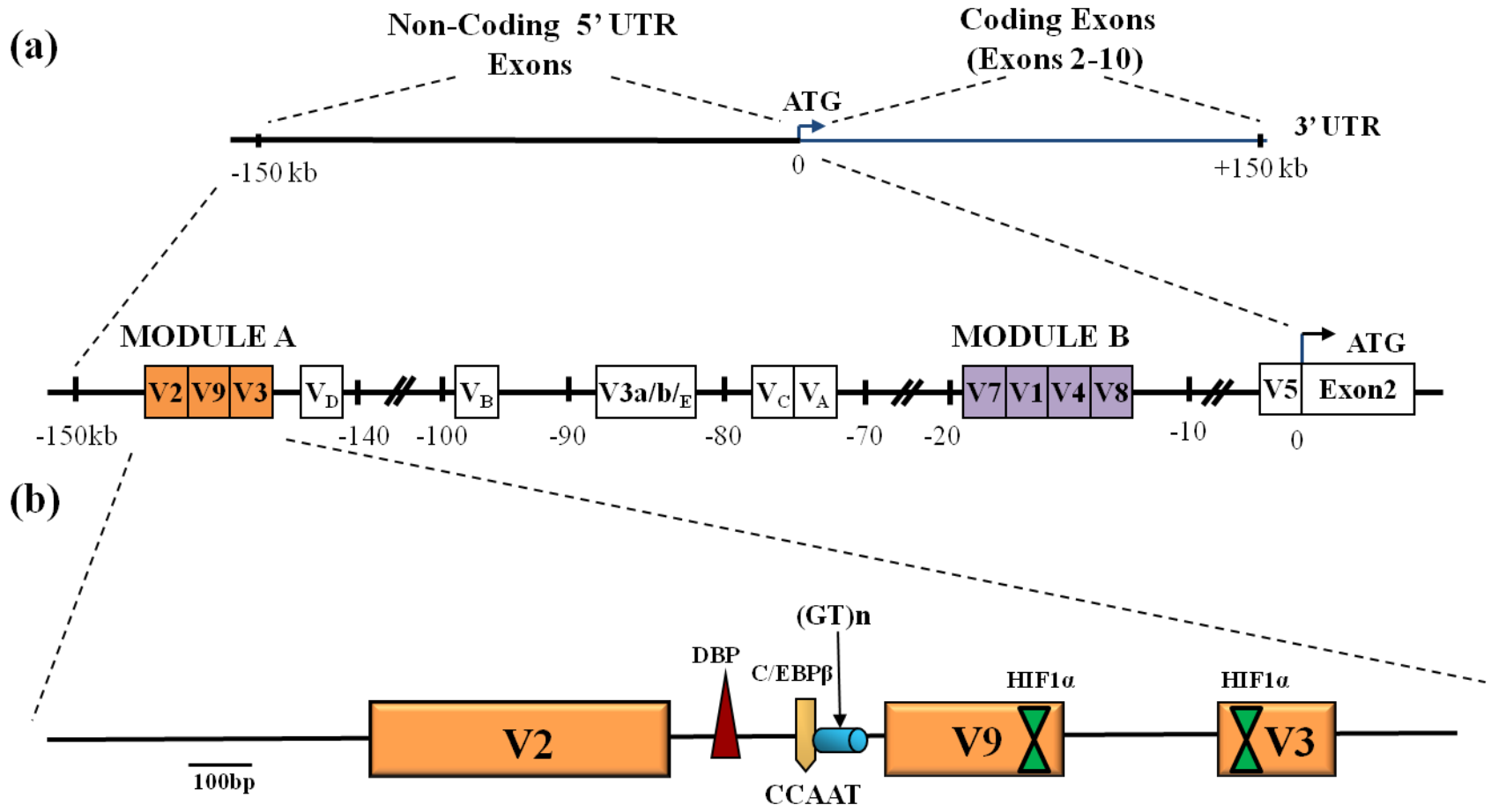


Figure IV-1: Schematic of the human *GHR* gene

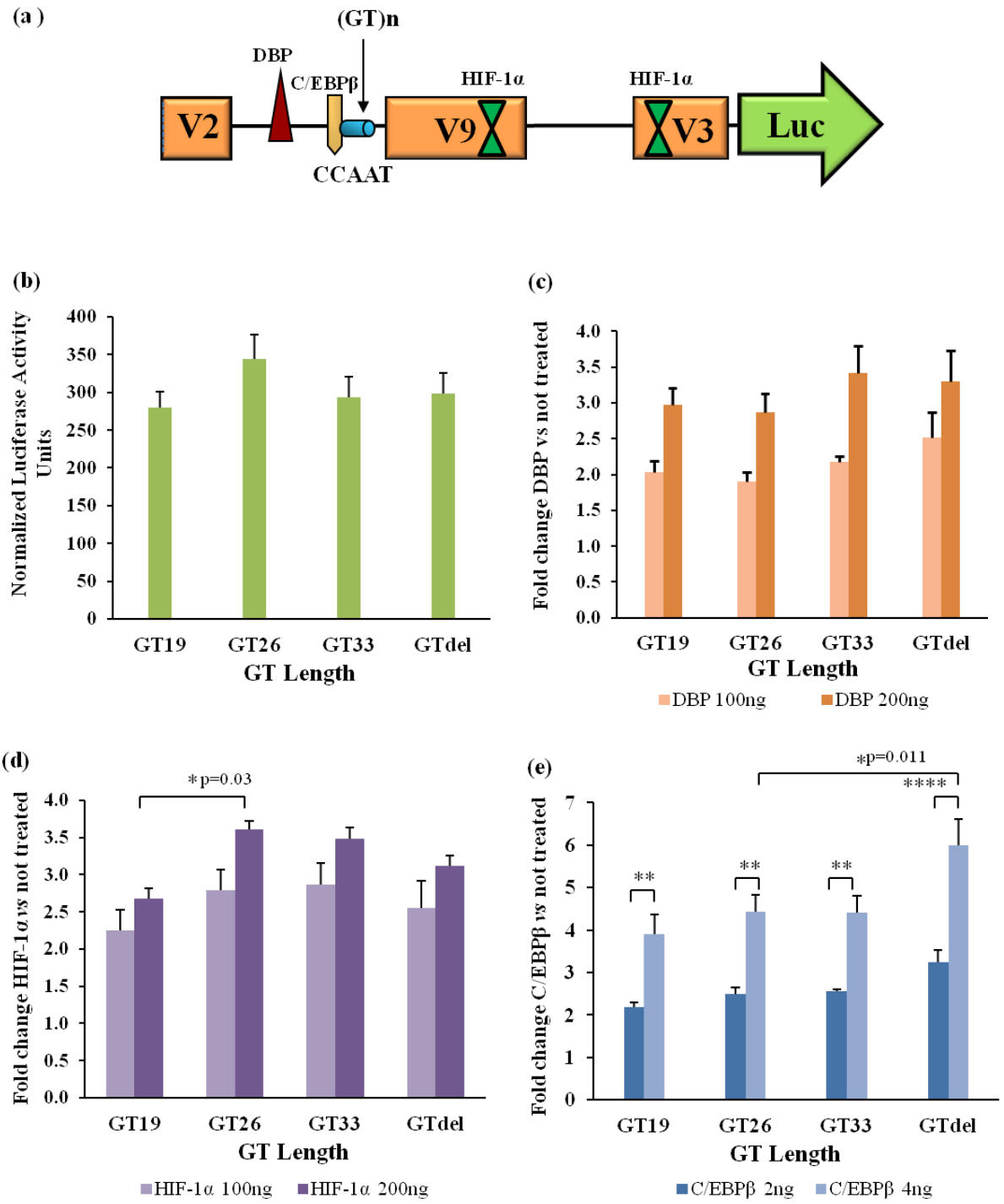


Figure IV-2: Effects of the GT repeat polymorphism on *GHR* gene promoter activity.

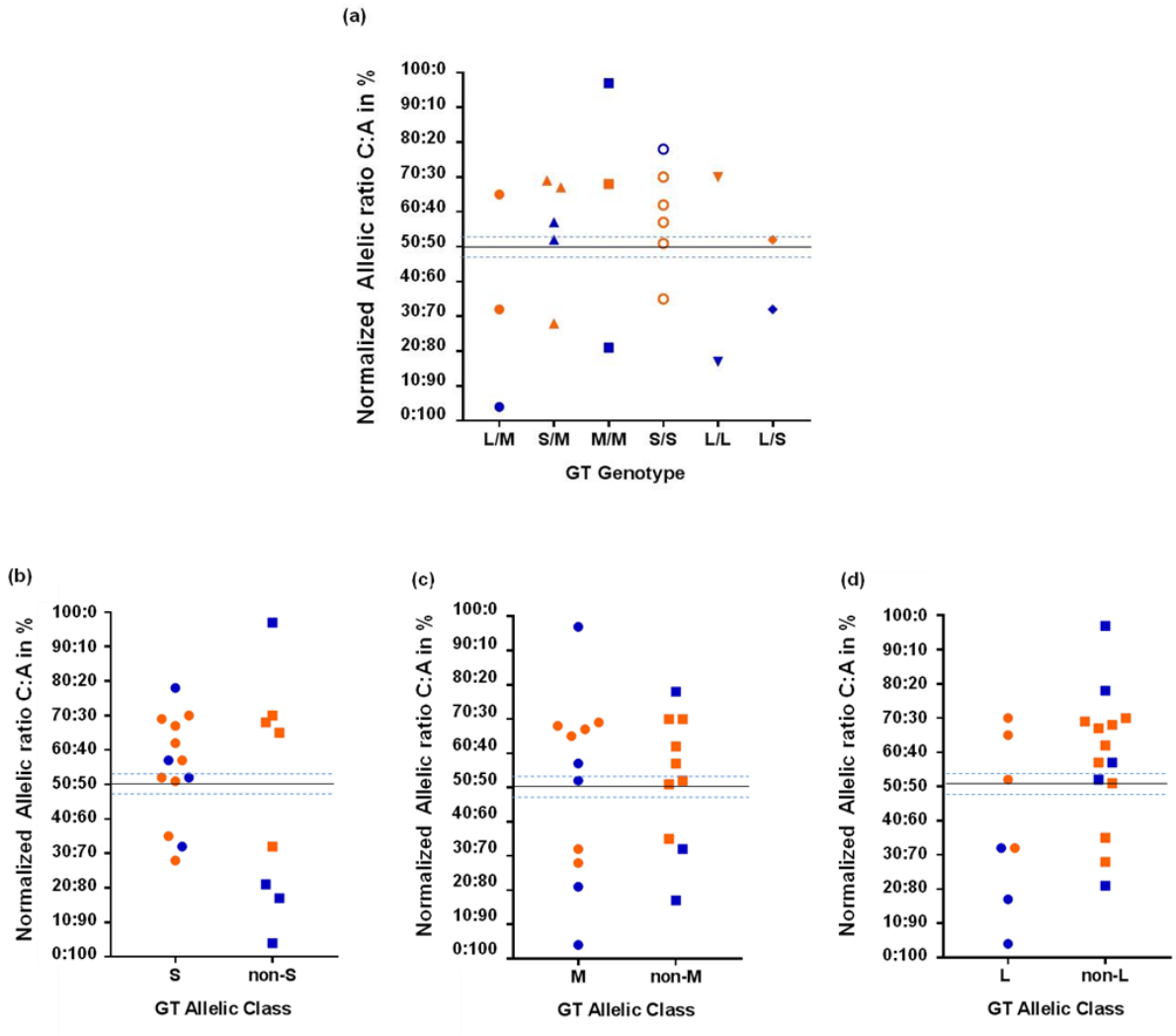


Figure IV-3: Differential allelic expression association with GT repeat length at the *GHR* gene locus in LCLs.

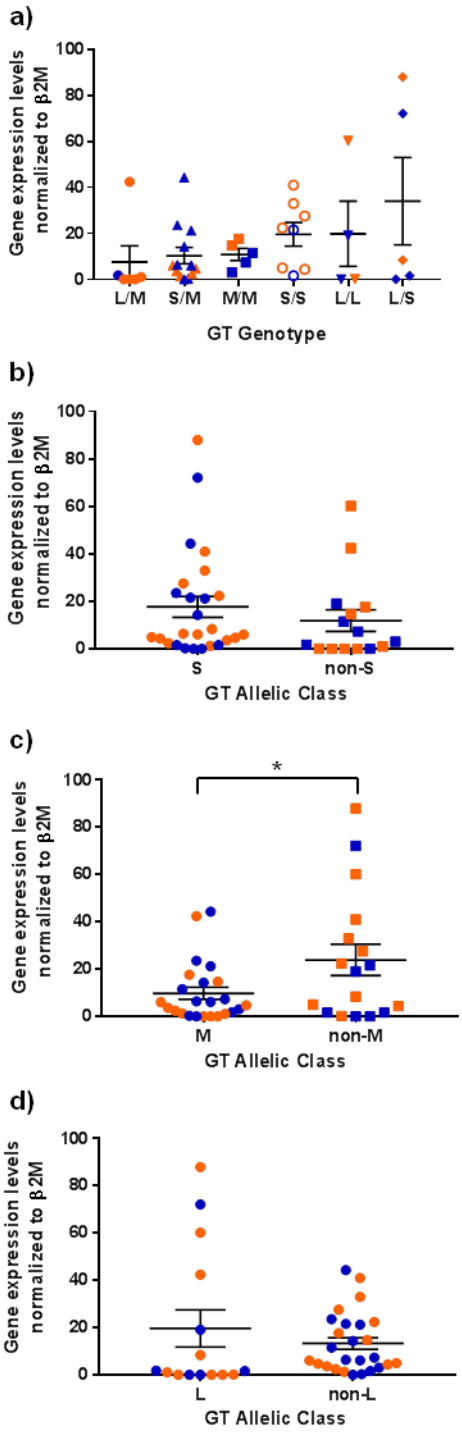


Figure IV-4: GT genotype and allelic associations with *GHR* mRNA expression in LCLs

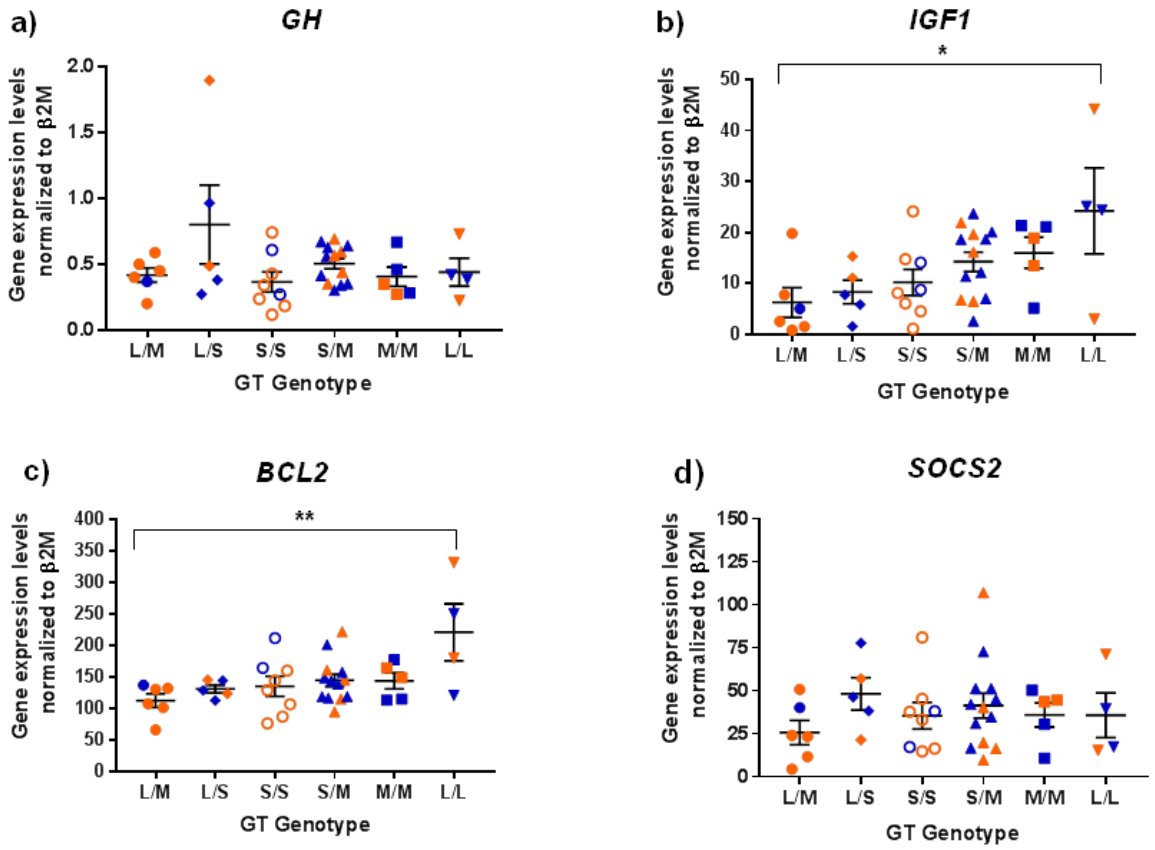


Figure IV-5: GT genotype and allelic associations with *GH, IGF1, BCL2 and SOCS2* mRNA expression in LCLs

Table IV-S1: Characteristics of the LCLs used in our study

LCL ID	Pedigree	Sex	GT Allele 1	GT Allele 2	Genotype
GM12003	CEPH/UTAH PEDIGREE 1420	Male	18	22	S/S
GM12910	CEPH/UTAH PEDIGREE 1477	Female	19	19	S/S
GM11830	CEPH/UTAH PEDIGREE 1350	Female	19	22	S/S
GM12567	CEPH/FRENCH PEDIGREE 12	Male	19	20	S/S
GM10843	CEPH/UTAH PEDIGREE 1423	Female	22	23	S/S
GM12004	CEPH/UTAH PEDIGREE 1420	Female	23	23	S/S
GM12001	CEPH/UTAH PEDIGREE 1420	Female	23	23	S/S
GM12743	CEPH/UTAH PEDIGREE 1444	Female	22	23	S/S
GM7017	CEPH/UTAH PEDIGREE 1333	Male	22	30	L/S
GM11919	CEPH/UTAH PEDIGREE 1423	Male	23	29	L/S
GM11039	CEPH/VENEZUELAN PEDIGREE 104	Female	21	31	L/S
GM12815	CEPH/UTAH PEDIGREE 1454	Female	19	31	L/S
GM10845	CEPH/UTAH PEDIGREE 1424	Male	19	30	L/S
GM11997	CEPH/UTAH PEDIGREE 1420	Female	18	27	S/M
GM11917	CEPH/UTAH PEDIGREE 1423	Male	19	25	S/M
GM12903	CEPH/UTAH PEDIGREE 1477	Male	19	27	S/M
GM12251	CEPH/UTAH PEDIGREE 1416	Female	19	24	S/M
GM12740	CEPH/UTAH PEDIGREE 1444	Female	22	28	S/M
GM12812	CEPH/UTAH PEDIGREE 1454	Male	19	26	S/M
GM11992	CEPH/UTAH PEDIGREE 1362	Male	23	25	S/M
GM12762	CEPH/UTAH PEDIGREE 1447	Male	23	24	S/M
GM11916	CEPH/UTAH PEDIGREE 1423	Male	23	25	S/M
GM10844	CEPH/UTAH PEDIGREE 1424	Female	23	24	S/M
GM12763	CEPH/UTAH PEDIGREE 1447	Female	23	28	S/M
GM12144	CEPH/UTAH PEDIGREE 1334	Male	23	28	S/M
GM07357	CEPH/UTAH PEDIGREE 1345	Male	23	28	S/M
GM12240	CEPH/UTAH PEDIGREE 1416	Male	25	26	M/M
GM12874	CEPH/UTAH PEDIGREE 1459	Male	26	26	M/M
GM12253	CEPH/UTAH PEDIGREE 1416	Female	24	28	M/M
GM12750	CEPH/UTAH PEDIGREE 1444	Male	25	28	M/M
GM12761	CEPH/UTAH PEDIGREE 1447	Female	24	28	M/M
GM12813	CEPH/UTAH PEDIGREE 1454	Female	29	31	L/L
GM11930	CEPH/UTAH PEDIGREE 1424	Male	30	30	L/L
GM12249	CEPH/UTAH PEDIGREE 1416	Female	30	30	L/L
GM12682	CEPH/FRENCH PEDIGREE 17	Male	30	30	L/L
GM11840	CEPH/UTAH PEDIGREE 1349	Female	27	29	L/M
GM11832	CEPH/UTAH PEDIGREE 1350	Female	26	30	L/M
GM12237	CEPH/UTAH PEDIGREE 1332	Female	26	29	L/M
GM11923	CEPH/UTAH PEDIGREE 1424	Male	24	29	L/M
GM07341	CEPH/UTAH PEDIGREE 1333	Female	26	31	L/M
GM12892	CEPH/UTAH PEDIGREE 1463	Female	27	32	L/M

Table IV-S2: List of primers and LNA probes used in quantitative PCR and droplet digital PCR

qPCR primers	
GHR_F	TGGATGGTTCTGGAGTATGAAC
GHR_R	AGGAAGTGTTACATAGAGCACCT
IGF-1_F	TGTGGAGACAGGGGCTTTTA
IGF-1_R	ATCCACGATGCCTGTCTGAG
B2M_F	AGATGAGTATGCCTGCCGTGT
B2M_R	GCTTACATGTCTCGATCCCACTTA
SOCS2_F	GAGCTCGGTCAGACAGGATG
SOCS2_R	TTCCGTCTTGGTATTCGATTC
BCL2_F	ATGTGTGTGGAGAGCGTCAA
BCL2_R	TTCAGAGACAGCCAGGAGAAA
GH1_F	CCTAGAGGAAGGCATCCAAA
GH1_R	GCAGCCCGTAGTTCTTGAGTAG
ddPCR primers /LNA probes	
rs6180_F	TGGACAATGCCTACTTCTGTG
rs6180_R	CTTGGTTTAAGCTTGGCTGTATG
rs6180_A_allele	5'-/5HEX/AG+T+G+C+A+TC+C CT/3IABkFQ
rs6180_C_allele	5'-/56-FAM/AG+T+G+C+C+TCC C/3IABkFQ

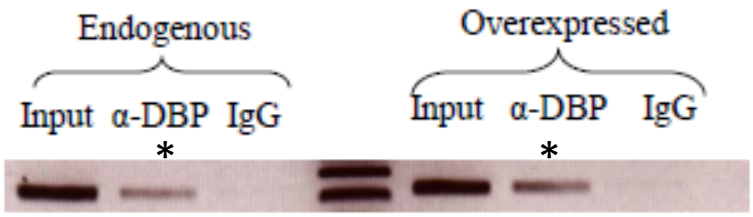


Figure IV-S1: **DBP binds its response element in V9 promoter.** The ChIP assay was performed as previously described (Goodyer et al. 2008). Briefly, HEK293 cells with endogenous or overexpressed DBP were cross-linked and lysates immunoprecipitated with an antibody against DBP (Dr. Christopher Mueller, Queen's University, Kingston, ON, Canada) or with IgG. After reversal of the cross-linking, the DNA was purified and PCR assays were performed across the binding site (forward primer: 5' GCT TTA TTT TCC TCC TGT TGT GCC 3'; reverse primer: 5' GCC AGA GCA GAC GCC AGA GTG 3'). Products were resolved on a 2% agarose gel. * indicates binding of the DBP within the region studied.

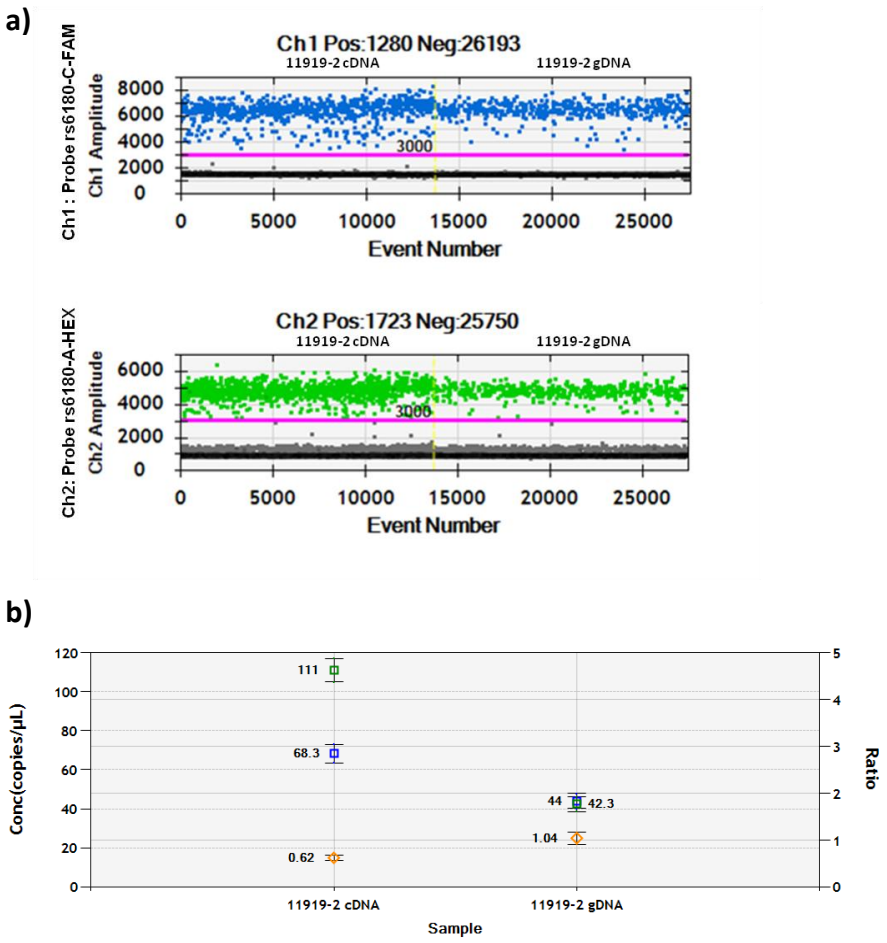
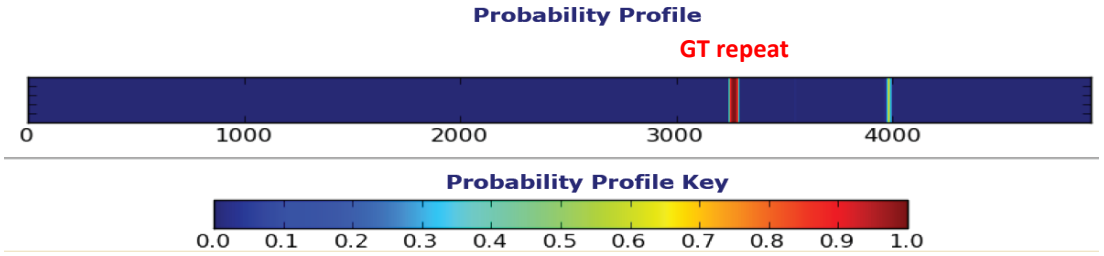


Figure IV-S2: Sample result of allelic imbalance measurement at rs6180 by droplet digital PCR in the LCL GM11919. a) The FAM or HEX fluorescent signals (y-axis) of each droplet is plotted against the droplet cumulative count (x-axis). The pink line indicates the positive fluorescence threshold set manually at 3000. The yellow line separates the two wells with LCL GM11919 starting cDNA on the left or gDNA on the right from the same passage. The blue or green dots represent individual droplets containing at least one copy of the allele (for rs6180: allele C or A, respectively). The QuantaSoft Analysis Software measures the positive (above the threshold) and negative (below the threshold) droplets for each fluorophore (FAM or HEX). Analysis of the fraction of positive droplets is then fitted to a Poisson distribution to determine the absolute initial copy number of the target DNA molecule in the input reaction mixture in units of copies/µl. **b**) Concentration results, presented as copies/µl in the sample, are displayed for each sample studied along with the ratio of the concentrations obtained of each allele (C/A) (orange).Ch, channel.

a)

Position (GRCh37/hg19)	Gene ID	Length	Z-Score	Sequence	Location
chr5:42424274-42424323	GHR	49	11495.27	(GT)24	87 bp upstream of V9 GHR TSS

b)



c)

Feature	Chromosome	Chrom Start	Chrom Stop	Strand
Z DNA Motif	chr5	42424274	42424321	+

Figure IV-S3: In silico analysis of the GT repeat propensity to form a Z-DNA helix. a) The Z-Hunt algorithm has been originally developed in 1986 by Ho and colleagues to measure the Z-DNA forming potential of a sequence (probability score) referred as the Z-score. The cut-off is 700.(Ho et al. 1986) **b)** The SIBZ algorithm (calculates the propensity of regions within the DNA molecule to transition from B to Z conformations under negative superhelical stresses (Zhabinskaya and Benham 2011) **c)** The non-B DB database predicted our GT repeat located in the V9 GHR promoter to form a Z-DNA motif (nonb-abcc.ncifcrf.gov).

CHAPTER V

GENERAL DISCUSSION AND FUTURE DIRECTIONS

1. The short stature phenotype: a complex trait

A. The ISS phenotype is a heterogeneous population

a) The GT repeat as a potential new marker associated with the ISS phenotype

The hallmark endocrine features of individuals with GHI are normal or increased GH levels with variable deficiencies of IGF-1; these have been shown primarily to correlate with the type of *GHR* mutations and the degree of short stature observed (David et al. 2011). The ISS population is composed of heterogeneous individuals presenting with varying degrees of GHI. The etiology of the ISS phenotype is not fully known, partly because of the different clinical subgroups it comprises (Wit et al. 2008). Although classified as GHI and, thus, showing a certain degree of GH resistance at the receptor level, there is little evidence of associations of *GHR* mutations with the ISS phenotype nor a *continuum* of the effects of these mutations on IGF-1 levels. Heterozygous mutations in the *GHR* coding regions have been detected in exonic regions in 2 to 5% of ISS individuals but with little functional evidence to support their relevance (Bonioli et al. 2005; Hujeirat et al. 2006).

Mutations in some genes outside the GH/IGF-1 axis have also been shown to be associated with ISS. The most common are mutations in the short stature homeobox-containing (*SHOX*) gene, located in the short arm pseudoautosomal region of the sex chromosomes and encoding a transcription factor that plays a role in chondrocyte proliferation and differentiation. Defects in the coding as well as cis-regulatory regions of the gene (point mutations, deletions, duplications) resulting in absence or haploinsufficiency of function are found in patients with Leri-Weill, Langer mesomelic dysplasia and Turner syndromes; these are characterized by short stature and dysmorphic features. In addition, clinical studies have shown that 1 to 17% of ISS individuals have SHOX mutations, with slightly affected body proportions or within the normal range (Binder 2011; Fukami et al. 2016; Marstrand-Joergensen et al. 2017; Rosilio

et al. 2012). It should be noted that our Montreal cohort was not tested for *SHOX* mutations, only ISS children without *SHOX* mutations were included in the Novara cohort and we have no information for the CARTaGENE cohorts, although it is likely that a certain percentage of the short stature adults may be carriers of some defects in the *SHOX* gene. Two other genes, aggrecan (*ACAN*) and the natriuretic peptide receptor 2 gene (*NRP2*) have been associated with the ISS phenotype in 1.5 to 3% of the cases (Kang 2017). These two genes were not examined in any of our cohorts.

The results from **Chapter II** showed that specific GT repeat genotypes within the *GHR* are associated specifically with the pediatric ISS phenotype. In order to understand what type of subgroup these genotypes were associated with, it was important to have as many clinical data as possible for these children. For the Montreal cohort, we obtained two height measurements (one at diagnosis and one at recruitment) so that we could extrapolate height velocities on growth curves. These data enabled us to conclude that the L/S carriers were the shortest of our group at diagnosis and showed the least catch-up growth with age. This led us to hypothesize a potential functional role for the extreme allelic lengths of this repeat. Unfortunately, we were unable to retrieve final height data for the majority of these children and, thus, could not conclude which GT subgroup(s) had children with catch-up growth *vs* those who remained under the 3rd percentile.

In the larger Italian Novara cohort, the ISS boys showed an increased prevalence of the S/M genotype but, because we did not have any height data other than at the time of diagnosis, we were unable to assess a more specific association with a subgroup. The ideal goal for future studies, even though this will take several years, would be to continue the collaboration with the Italian pediatricians until they have been able to obtain the final heights of their ISS and control children. Interestingly, the boys presenting with the S/M genotype did not show any difference in height relative to the non S/M carriers, so we cannot rule out that this genotype could be a marker of children with catch-up growth at puberty. In this case, we could have a marker for a subgroup of children that could avoid unnecessary GH treatments.

The pooled ISS cohort analysis confirmed, at least partly, the results: we again observed that the S/M genotype was more prevalent in the male ISS children. However, we failed to reproduce the L/S genotype even though we observed more than twice more ISS females with

this genotype; the small size of the cohorts was a problem as the prevalence of L/S was only ~5% in the different control cohorts. Although, in the end, a common etiology for the ISS phenotype could not be obtained in our study, the data strongly suggest that the GT repeat microsatellite within *GHR* constitutes an interesting potential marker of ISS subgroups.

b) The missing heritability and the genetic complexity

Height is a highly heritable polygenic trait and a good model to study the genetic architecture of a complex trait (Lettre 2011). The first two genes shown to be associated with variation in height in the human population, HMGA2 and GDF5, accounted for less than 1% of the variance in height (Sanna et al. 2008; Weedon et al. 2007). *HMGA2* is one of the high-mobility group (HMG) proteins and encodes for an architectural factor that binds AT-rich DNA sequences in the minor groove. It plays a role in DNA packaging as well as in the assembly of transcriptional complexes to regulate specific genes (Bianchi and Agresti 2005). Interestingly, HMGA2 has been shown to have a role in general growth control, osteogenesis and adipogenesis. HMGA2 also has been shown to be overexpressed in several cancers, correlating with a poor overall survival score (Huang et al. 2018). $Hmga2^{-/-}$ mice have a pygmy phenotype with decreased fat levels (Zhou et al. 1995) and microdeletions at the HMGA2 locus have been associated with osteopoikilosis, mental retardation and proportional short stature (Menten et al. 2007). Variants in the GDF5-UQCC locus have been associated with variation in height (Sanna et al. 2008). Rs143383, associated with height and located in the 5'UTR of GDF5, has been shown to be also associated with risk of osteoarthritis in European and Asian populations (Chapman et al. 2008; Miyamoto et al. 2007) and with GDF5 differential allelic expression. The SNP we chose in our analysis in **Chapter III** is in complete LD with Rs143383. GDF5 is an important growth factor, a member of the TGFβ superfamily, which has major roles in skeletogenesis, particularly at the early stages of chondrogenesis (Francis-West et al. 1999). Mutations of this gene result in brachydactyly and chondrodysplasias in humans characterized by severe shortening of skeletal elements (Polinkovsky et al. 1997; Thomas et al. 1996)

Ten years after the identification of these two genes, the latest height-related GWAS carried out by the GIANT (Genetic Investigation of Anthropometric Traits) consortium reported

almost 700 loci associated with variation in height in the human population (Weedon et al. 2007; Wood et al. 2014). Surprisingly, these ~700 loci explain only ~20% of the phenotype variability. The identified loci are common SNP variants (with a MAF ≥ 5%) and have mostly small effects so that the impact on height comes from a combination of small effects (Lettre 2011). Interestingly, these loci show an enrichment for genes involved in growth, including genes contributing to syndromes of abnormal skeletal growth, but also for genes and pathways not previously linked with skeletal growth (e.g. mTOR, osteoglycin and binding of hyaluronic acid). A more recent study, using an Exome ChiP on ~700 000 individuals, uncovered 83 new rare or low frequency variants (MAF < 5%) with moderate to large effect on height (up to 2 cm/allele), showing that rare sequence variants can also influence height variation (Marouli et al. 2017).

For our **Chapter II** SNP analysis, we were limited to the choice of common SNPs due to the small size of our cohorts. However, to find if there could be rare variants associated with the ISS phenotype in the regulatory regions of *GHR*, I initially sequenced the Module A region (~2kb) in the Montreal ISS cohort along with ~70 controls but did not find novel mutations. The absence of mutations in the Module A region coupled with our small sample size and the fact that, as mentioned earlier, only 2-5% of the ISS children have been found with heterozygous mutations in their coding exons, led us to decide not to go forward with sequencing the exonic regions of the Montreal ISS children. We instead chose a genotyping strategy with a selected set of known common SNPs encompassing the *GHR* locus.

SNPs are very practical as markers as they are common, stable and easy to use in genotyping platforms. But they may not capture all the variability at certain loci. Some authors, seeing the limitation of the SNP-based GWAS approach to explain phenotypic variability, have brought forward the concept of 'missing heritability' and suggested the need to look at other types of genetic variations "to fill the gap" (Hannan 2010). Microsatellites are frequent and have a high heterozygosity rate which makes them interesting as they provide a wide variable range of possible genotypes compared to the binary nature of SNPs (Hannan 2010). Moreover, it has been shown that microsatellites are poorly tagged by nearby SNPs so that both SNPs and microsatellites would add up to constitute a multi-marker platform to map quantitative trait loci.

In **Chapter II,** we found that rs4292454 (located in intron 2) associated with the ISS phenotype as did significant 3' risk haplotypes spanning exon 2 to 120kb downstream of the *GHR* gene. Classifying the ISS children and their controls according to the rs4292454 genotypes alone did not show significant differences in height nor a correlation with the L/S or the S/M GT genotypes; and the same lack of correlation was observed with the 3' risk haplotypes alone. Our study suggests that the three types of markers identified (GT repeat genotype, SNP and haplotype) may be associated with a specific subgroup of ISS, but the present lack of final adult height data for our cohorts has so far prevented us from pursuing this hypothesis.

Another possible future route to take, to overcome these obstacles, would be to try again to collaborate with the one other group in Europe with large ISS and control cohorts (~430 ISS and ~1200 controls). There is a network of European pediatric endocrinologists (NESTEGG) which focuses on the genomics of growth, including children diagnosed with small for gestational age (SGA) and ISS. They have collected clinical diagnostic data and DNA samples but whether they have been able to follow all of the ISS and control children to obtain final adult heights is an important question (Johnston et al. 2009). We approached them several years ago, when we were just beginning the Montreal ISS studies, but they were not open to a collaboration at the time. It may be that, now we have published our initial data, they may be more agreeable. If they are, and they have collected final adult heights, I would propose to carry out the same analyses as in **Chapter II** in order to perform association analysis with these three markers and generate a genetic risk score using conditional regression analysis. If we were able to get Novara final adult height data and pool results of the three cohorts, the cumulative n~600 ISS children should provide sufficient power to allow for assessing the "three marker" hypothesis. The categorical output would entail dividing the ISS group into children that retrieve a normal height following puberty and those who stayed under the 3rd percentile. This is important because, at present, there is no diagnostic marker for the pediatric endocrinologist to decide whether GH treatment of an ISS child is appropriate or unnecessary.

A CA repeat in the promoter of the *IGF-1* gene has also been extensively studied, looking for its potential association with IGF-1 serum concentration, susceptibility to certain diseases and

risk in certain types of cancer. In 2011, Chen et al. reported a specific haplotype association combining Tag SNPs in LD and the promoter CA dinucleotide repeat with serum IGF-1 concentrations in a cohort of 450 premenopausal Chinese women (Chen et al. 2011). In this study, they used a similar approach to ours with the Haploview software bioinformatic tool to define the haploblock structure of the *IGF-1* gene and a Tag SNP approach to capture IGF-1 variability. They also used an alternative software program to infer haplotypes, PHASE, that is based on a different algorithm; while Haploview uses the expectation-maximisation (EM) algorithm (Barrett et al. 2005), PHASE uses Bayesian methods (Stevens, 2002) that also infer the diplotypes for each individual. As an additional investigation of our ISS cohorts, I would propose using the same strategy as Chen et al. With the final height data of the ISS, we could perform two kinds of analysis: one with the ISS *vs* control height but with a previous filtering of the ISS, keeping only the ones that stayed under the 3rd percentile, and one with a dichotimization of the ISS group itself, with an ISS subgroup that showed catch-up growth and a subgroup that did not. I hypothesize that this approach will allow us to find haplotypes and diplotypes composed of the GT repeat and SNPs in LD in the 5' regulatory region of *GHR* that would constitute sequence variants underlying shortness or catch-up growth more precisely than the GT repeat alone.

B. Short stature and obesity

a) Short stature as a risk factor for obesity

The **Chapter III** study showed a significantly higher prevalence of obesity (BMI>30kg.m^2) in the short stature compared to normal height women; at the genetic level this result translated into a set of variants specifically located in the *GHR* promoter of short stature women with both risk or protective odds after adjustment by BMI, BAI or FMI. In order to further characterize these effects, I would need to perform further stratification, to determine which category/threshold (e.g. normal *vs* obese *vs* morbidly obese) of each adiposity index drives the changes. With our present cohort, this resulted in a significant loss in statistical power. To overcome this limitation, a larger cohort of short stature adults will be needed. When we started our project, CARTaGENE had recruited ~20,000 individuals; today they have reached ~43,000 making it possible to at least double our cohort size in any future studies.

b) Sex-specificity across the studies

From the control case analyses in **Chapters II and III** to the investigation of a functional role for the GT repeat in **Chapter IV**, there was evidence of sex specificity effects. In **Chapter II**, the significantly higher prevalence of the S/M genotype of the *GHR* GT repeat was only observed for the ISS boys. In **Chapter III**, short stature women, but not men, presented with both increased adiposity (measured by different indices) compared to normal height controls; this was combined with a strong genetic component that included a cluster of nominally significant variants in the *GHR* promoter and the L/M GT genotype whose association with height is affected by different adiposity indices. Finally, in **Chapter IV**, sexual dimorphism was observed in CEPH cell lines at the transcriptional level of *IGF-1* in males and *BCL-2* in females: in both cases their lowest expression levels were associated with the L/M genotype. In addition, post-transcriptionally high levels of AI in *GHR* were measured in the males, potentially in association with the L allele. This *continuum* of sex specificity at the genotype-phenotype level and at the level of gene expression regulation not only in *GHR* but also two downstream gene targets (*IGF-1*, *BCL2*) strongly suggests sex-specific mechanisms are involved.

c) Functional characterization of sex-specificity in the GH/IGF-1 axis

Sexual dimorphism in gene expression is common in mammalian somatic tissues and may underlie the differential susceptibility to certain diseases between men and women (Rinn and Snyder 2005). Sexually dimorphic patterns of GH secretion observed in humans and in rodents have been linked to sex differences in downstream signaling pathways in target tissues with consequences on gene expression and body composition as well as metabolism. It has been shown that 20 to 30% of rodent hepatic genes have a sex-specific expression pattern; these are genes involved in metabolic functions, including drug and lipid metabolism, but also transcription factors or signaling proteins (Liu et al. 2016). GH is the main regulator of sexually dimorphic gene expression in the rodent liver through the pulsatile activation of STAT-5b (Liu et al. 2016). Recently, Lau-Corona et al. showed that the loss of pulsatile STAT-5 activity, through continuous infusion of GH in the male mouse, led to a down-regulation of 'male-biased' genes that were dependent on a pulsatile STAT-5 stimulation

(Lau-Corona et al. 2017). The absence of a total reversal to a feminine gene expression pattern indicated a resistance to continuous GH stimulation, suggesting some programming probably at the epigenetic level of the highly 'male-specific' genes (Lau-Corona et al. 2017).

It is still not known today how exactly GH drives and maintains the sex-specific gene expression patterns observed at the molecular level in the liver and in different somatic tissues as well. However, in the rodent liver, these mechanisms are thought to include sex-specific chromatin states, determined by DNAse hypersensitivity mapping and high throughput sequencing, and local chromatin marks, characterized by specific histone modifications (Ling et al. 2010; Sugathan and Waxman 2013). Sexually dimorphic physiological patterns are also likely the result of complex interactions between the GH/IGF-1 axis and gonadal steroids, as best illustrated in puberty: the increase of both GH and estrogen triggers the growth spurt (Leung et al. 2004). The sexually dimorphic nature of the GH secretory pattern itself has been shown to be the result of a neonatal programming of the hypothalamus (GHRH and somatostatin producing neurons) *via* exposure to sex steroids (androgens and estrogens), with subsequent effects during puberty likely affecting expression of the *GH* and *GHR* genes (Fernandez-Perez et al. 2016; Raisman and Field 1973).

At the *GHR* gene level, there are few studies looking for a molecular mechanism to explain direct regulation of *GHR* gene expression by sex steroids in humans. Previous animal studies have shown a positive regulation of hepatic *GHR* mRNA expression by estrogens in rodents, with higher levels of the liver-specific *GHR1* variant in female rats (Gabrielsson et al. 1995). In extra-hepatic tissues, the effect of estrogens has been shown to be tissue-specific, with up-regulation in osteoblasts and down-regulation in the brain (Bennett et al. 1996). In humans, one study has shown a biphasic effect of estrogen on *GHR* mRNA levels in osteoblasts (Slootweg et al. 1997).

To further investigate the sexual dimorphism effects observed in **Chapter IV** at the level of *GHR*, I would propose a two-part study: I would hypothesize, first, that the *GHR* GT repeat polymorphism can regulate the response to estrogen and, second, that there are specific epigenetic marks in LCLs at the level of the *GHR* GT repeat locus that could be specific in males *vs* females ('programmed' to answer to GH in a sex-specific manner) and may be linked to a specific GT genotype. The *GHR* promoter has several ERα response elements

(EREs) predicted by *in silico* analysis (PROMO database). Those within the proximal promoter region for Module A would be tested for functionality by Luc reporter promoter constructs followed by site-directed mutagenesis of the EREs or by ChIP directly in HEK293 cells transfected with an expression vector for ERα.

The potential response modulation by the GT polymorphism would be tested in our set of LCLs, with stimulation of estrogen alone or in combination with GH, as a synergistic action of the two hormones has been demonstrated in primary mouse hepatocytes in upregulating *GHR* expression (Contreras and Talamantes 1999). Using ChIP with specific antibodies for histone post-translational modifications followed by sequencing, we can investigate differential epigenetic marks at the GT locus region in males *vs* females and in relation to their GT genotype. Interestingly, most of the binding sites for ER (~95%) are located in distal parts of gene promoters and often associate with enhancer elements. To characterize a potential transcriptionally active ER, we could verify the presence of associated co-factors, such as AIB1, p300 and CBP as well as the pioneer factor FOXA1 (Carroll 2016), and histone marks usually associated with active enhancers (e.g. H3K4me1, H3K4me2, H3K27ac and H2A.Z). Moreover, it has also been demonstrated that long-range chromatin interactions occur between ERα bound regions and their target promoters; the same authors also showed that ERα, through chromatin looping, could bring genes together for a coordinated transcriptional regulation (Fullwood et al. 2009). In line with these properties of ERα, and if we validate a functional ERE in the distal promoters of *GHR*, *IGF-1* and *BCL-2*, it would be of interest to explore the potential ERα long range regulation of these three genes through chromosomal conformation capture (3C) techniques.

Differential methylation of CpG is another epigenetic mark that could be interesting to investigate as Module A is a GC rich region and CpG islands overlap V2, V9 and V3 exons. Promoter CpG islands at promoter sites are generally hypomethylated while CpGs outside these regions are mainly methylated, however, little is known about how methylation modification could play a role in the case of gender-specific gene expression. The LCLs derived from males showed the highest degree of allelic imbalance. I hypothesize that differential allelic methylation at CpG islands and CpG sites distal to and/or within the *GHR* Module A locus may play a role in two ways: in the sex specificity and in the level of allelic

imbalance observed. To test this hypothesis, I would use pyrosequencing technology after bisulfite treatment of genomic DNA for an accurate assessment of the levels and positions of methylated CpGs and design an assay to detect allele-specific methylation profiles in the set of LCLs.

2. Towards a molecular mechanism for the GT repeat in *GHR*

The results of my studies have led me to propose potential mechanisms for the *GHR* GT microsatellite (**Figure V-1**). This model involves the formation of Z-DNA by the GT repeat as a driver for the differential effects observed at the transcriptional level. The discovery of proteins able to bind Z-DNA forming elements has been crucial in elucidating the potential role of this conformation *in vivo*. Since the isolation of the double-stranded RNA deaminase (ADAR-1) from chicken blood using a Z-DNA affinity column (Herbert and Rich 1993), three more Z-DNA binding proteins (ZBPs) have been identified: ZBP1/DLM-1 in mammals (Schwartz et al. 2001), PKZ in zebrafish (Rothenburg et al. 2005) and the poxvirus virulence factor, E3L (Kim et al. 2003). All of these have been implicated in the interferon response pathway (Herbert et al. 1995). ADAR-1 is particularly highly expressed in LCLs (www.gtexportal.org) and has the ability to bind the GT repeat in its Z-DNA conformation through its Z-DNA binding domain. Lafer et al. have demonstrated that, in contrast to B-DNA, Z-DNA is a strong immunogen and highly specific polyclonal as well as monoclonal antibodies have been produced (Lafer et al. 1981). Z-DNA antibodies are also produced during certain human and murine auto-immune diseases, including systemic lupus erythematosus (Lafer et al. 1981; Lafer et al. 1983).

I hypothesize that the amount of ADAR-1 binding to each GT allele is proportional to its length; that imbalance would translate into a differential transcriptional activity (so a differential loading of RNApol II) between the two alleles which would explain, at least in part, the AI measured in the GHR gene. To prove this differential enrichment, I would use Z-DNA and ADAR-1 specific antibodies to perform ChIP on LCLs with homozygous GT repeat lengths (S/S *vs* M/M *vs* L/L) and compare the relative enrichment between the cell lines by quantitative PCR. To verify if this differential enrichment of ADAR-1 is linked to a differential transcriptional activity I would also use an antibody for RNApol II

(phosphorylated at serine 5) to analyze the loading level of RNApol II at the transcription start site of the GHR V9 variant (located 80bp downstream of the GT repeat).

It has been demonstrated, through fluorescence *in situ* hybridization (FISH) experiments, that chromosome organization is not random in the nucleus; instead, each chromosome occupies a specific territory and provides a spatial organization for mainly intra-chromosomal interactions (Bolzer et al. 2005). However, there is also recent evidence for inter-chromosomal interactions: RNApol II and H3K9me3 marks accumulate in distinct foci dispersed throughout the nucleus, marking actively transcribed gene transcriptional factories as well as inter-chromosomal interaction clusters (Belyaeva et al. 2017). Recently, cluster formation and transcriptional activity of two genes, *Gfap* and *Osmr*, have been shown to be regulated by *Brahma-related gene 1* (*BRG1*), a chromatin remodeler, and STAT3 in astrocytes (Ito et al. 2018). In addition, *IGF2*, an imprinted gene, was reported to be located at the edge of a chromosomal territory where it is involved in inter-chromosomal associations (physically or co-localized) with two other imprinted genes located in different chromosomes (Lahbib-Mansais et al. 2016). In line with these recent advances, and because short tandem repeats, specifically with GT motifs, have been implicated in the 3D organization of the human genome (Nikumbh and Pfeifer 2017), I would test the possibility of inter-chromosomal interactions of *GHR* with *IGF-1* and *BCL-2* driven by the GT repeat. Using FISH techniques, I would assess first if the three genes tend to co-localize in LCLs under basal *vs* stimulated (GH, estrogen) conditions, followed by 3C technique using antibodies for RNApol II or Z-DNA or Z-DNA binding proteins (ChIP-loop) to know which regions (Z-DNA forming or not) are potentially interacting with our *GHR* GT repeat. Finding these 'trans' activities would open up an important new field of study for the GT repeat/Z-DNA motifs.

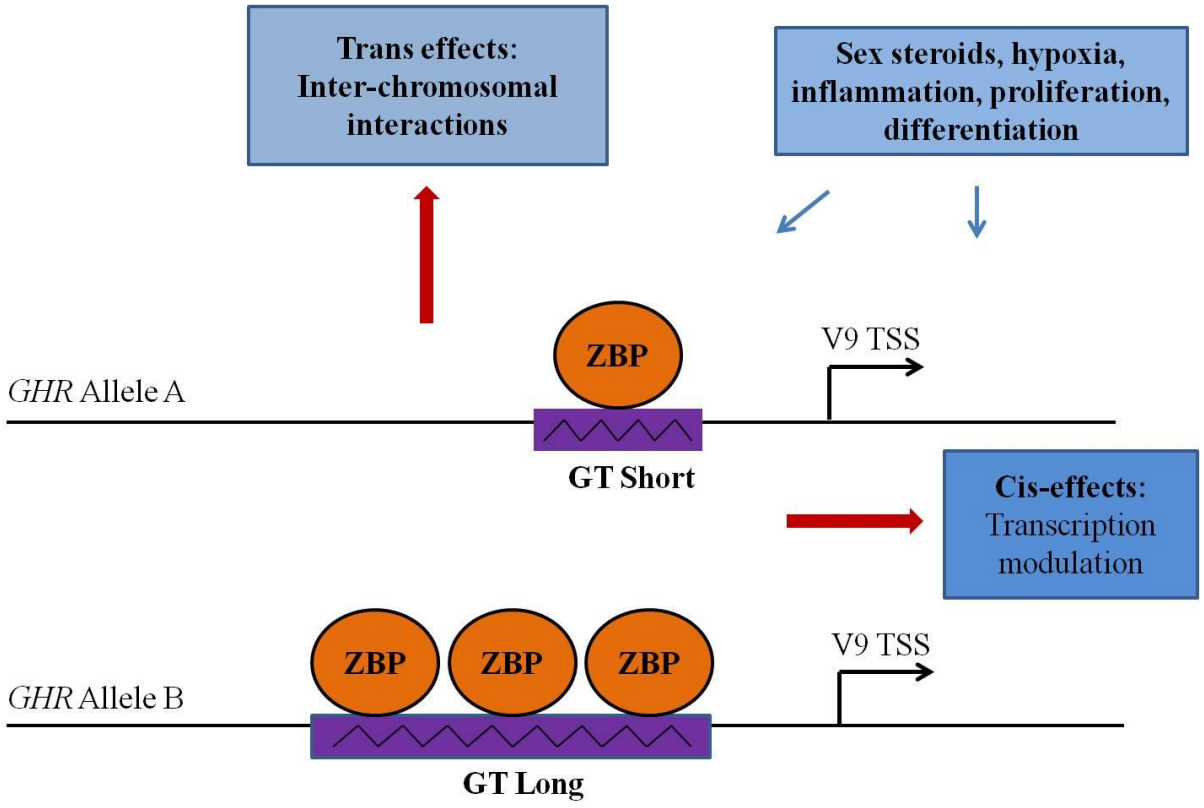


Figure V-1: Hypothetical model of the regulatory effects by the *GHR* promoter GT microsatellite.

CONCLUSION

My thesis work examined the roles of the *GHR* gene in its association with two complex traits, short stature and obesity. My discovery of the association of specific genotypes (SNPs, haplotypes, GT) with pediatric idiopathic short stature is promising and, combined with future clinical data, has the potential to formulate diagnostic markers for specific ISS subgroups. Given that 1-2% of children worldwide are considered to have ISS, the ability to provide families with information about the DNA variants underlying shortness or predict who will have catch-up growth would help to limit unnecessary GH treatments. My finding of increased risk of obesity in short stature women is the first study to show that several growth-related genes (*GHR, HMGA2, GDF5*) share associations with both severe short stature and obesity-linked anthropometric traits in a sex-specific manner. Finally, the functional characterization of the GT repeat has shed a new light on its biological significance as a fine-tuning modulator of *GHR* expression in a context- and sex-specific fashion. Ideally, future functional studies should be carried out in lymphoblasts or fibroblasts derived from ISS children, something that has only been reported once to date (Ocaranza et al. 2012).

REFERENCES

Agarwal AK, Giacchetti G, Lavery G, Nikkila H, Palermo M, Ricketts M, McTernan C, Bianchi G, Manunta P, Strazzullo P et al. . 2000. CA-Repeat polymorphism in intron 1 of HSD11B2 : effects on gene expression and salt sensitivity. Hypertension 36(2):187-194.

Allard JB, and Duan C. 2018. IGF-Binding Proteins: Why Do They Exist and Why Are There So Many? Front Endocrinol (Lausanne) 9:117.

Allen DB, Backeljauw P, Bidlingmaier M, Biller BM, Boguszewski M, Burman P, Butler G, Chihara K, Christiansen J, Cianfarani S et al. . 2016. GH safety workshop position paper: a critical appraisal of recombinant human GH therapy in children and adults. Eur J Endocrinol 174(2):P1-9.

Ambler GR, Fairchild J, and Wilkinson DJ. 2013. Debate: idiopathic short stature should be treated with growth hormone. J Paediatr Child Health 49(3):165-169.

Amit T, Youdim MB, and Hochberg Z. 2000. Clinical review 112: Does serum growth hormone (GH) binding protein reflect human GH receptor function? J Clin Endocrinol Metab 85(3):927-932.

Argetsinger LS, Campbell GS, Yang X, Witthuhn BA, Silvennoinen O, Ihle JN, and Carter-Su C. 1993. Identification of JAK2 as a growth hormone receptor-associated tyrosine kinase. Cell 74(2):237-244.

Argetsinger LS, and Carter-Su C. 1996. Mechanism of signaling by growth hormone receptor. Physiol Rev 76(4):1089-1107.

Ariyasu H, Takaya K, Tagami T, Ogawa Y, Hosoda K, Akamizu T, Suda M, Koh T, Natsui K, Toyooka S et al. . 2001. Stomach is a major source of circulating ghrelin, and feeding state determines plasma ghrelin-like immunoreactivity levels in humans. J Clin Endocrinol Metab 86(10):4753-4758.

Attie KM, Carlsson LM, Rundle AC, and Sherman BM. 1995. Evidence for partial growth hormone insensitivity among patients with idiopathic short stature. The National Cooperative Growth Study. The Journal of pediatrics 127(2):244-250.

Ayling RM, Ross R, Towner P, Von Laue S, Finidori J, Moutoussamy S, Buchanan CR, Clayton PE, and Norman MR. 1997. A dominant-negative mutation of the growth hormone receptor causes familial short stature. Nature genetics 16(1):13-14.

Ayuk J, and Sheppard MC. 2006. Growth hormone and its disorders. Postgrad Med J 82(963):24-30.

Bach LA. 2018. What Happened to the IGF Binding Proteins? Endocrinology 159(2):570-578.

Bagshaw ATM. 2017. Functional Mechanisms of Microsatellite DNA in Eukaryotic Genomes. Genome Biol Evol 9(9):2428-2443.

Ballesteros M, Leung KC, Ross RJ, Iismaa TP, and Ho KK. 2000a. Distribution and abundance of messenger ribonucleic acid for growth hormone receptor isoforms in human tissues. J Clin Endocrinol Metab 85(8):2865-2871.

Ballesteros M, Leung KC, Ross RJ, Iismaa TP, and Ho KK. 2000b. Distribution and abundance of messenger ribonucleic acid for growth hormone receptor isoforms in human tissues. J Clin Endocrinol Metab 85(8):2865-2871.

Barnard R, Haynes KM, Werther GA, and Waters MJ. 1988. The ontogeny of growth hormone receptors in the rabbit tibia. Endocrinology 122(6):2562-2569.

Barrett JC, Fry B, Maller J, and Daly MJ. 2005. Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21(2):263-265.

Barton DE, Foellmer BE, Wood WI, and Francke U. 1989. Chromosome mapping of the growth hormone receptor gene in man and mouse. Cytogenetics and cell genetics 50(2-3):137-141.

Bass SH, Mulkerrin MG, and Wells JA. 1991. A systematic mutational analysis of hormone-binding determinants in the human growth hormone receptor. Proc Natl Acad Sci U S A 88(10):4498-4502.

Baumann G. 1991. Growth hormone heterogeneity: genes, isohormones, variants, and binding proteins. Endocr Rev 12(4):424-449.

Baumann G. 1994. Growth hormone-binding proteins: state of the art. J Endocrinol 141(1):1-6.

Baumann G. 2001. Growth hormone binding protein 2001. J Pediatr Endocrinol Metab 14(4):355-375.

Baumann GP. 2009. Growth hormone isoforms. Growth Horm IGF Res 19(4):333-340.

Baumbach WR, Horner DL, and Logan JS. 1989. The growth hormone-binding protein in rat serum is an alternatively spliced form of the rat growth hormone receptor. Genes Dev 3(8):1199-1205.

Baumgartner JW, Wells CA, Chen CM, and Waters MJ. 1994. The role of the WSXWS equivalent motif in growth hormone receptor function. J Biol Chem 269(46):29094-29101.

Bayele HK, Peyssonnaux C, Giatromanolaki A, Arrais-Silva WW, Mohamed HS, Collins H, Giorgio S, Koukourakis M, Johnson RS, Blackwell JM et al. . 2007. HIF-1 regulates heritable variation and allele expression phenotypes of the macrophage immune response gene SLC11A1 from a Z-DNA forming microsatellite. Blood 110(8):3039-3048.

Belyaeva A, Venkatachalapathy S, Nagarajan M, Shivashankar GV, and Uhler C. 2017. Network analysis identifies chromosome intermingling regions as regulatory hotspots for transcription. Proc Natl Acad Sci U S A 114(52):13714-13719.

Bennett PA, Levy A, Carmignac DF, Robinson IC, and Lightman SL. 1996. Differential regulation of the growth hormone receptor gene: effects of dexamethasone and estradiol. Endocrinology 137(9):3891-3896.

Bergan-Roller HE, and Sheridan MA. 2018. The growth hormone signaling system: Insights into coordinating the anabolic and catabolic actions of growth hormone. Gen Comp Endocrinol 258:119-133.

Bernstein BE, Meissner A, and Lander ES. 2007. The mammalian epigenome. Cell 128(4):669-681.

Bhargava A, and Fuentes FF. 2010. Mutational dynamics of microsatellites. Mol Biotechnol 44(3):250-266.

Bianchi ME, and Agresti A. 2005. HMG proteins: dynamic players in gene regulation and differentiation. Curr Opin Genet Dev 15(5):496-506.

Bidosee M, Karry R, Weiss-Messer E, and Barkey RJ. 2009. Regulation of growth hormone receptors in human prostate cancer cell lines. Mol Cell Endocrinol 309(1-2):82-92.

Binder G. 2011. Short stature due to SHOX deficiency: genotype, phenotype, and therapy. Horm Res Paediatr 75(2):81-89.

Binder G, Baur F, Schweizer R, and Ranke MB. 2006. The d3-growth hormone (GH) receptor polymorphism is associated with increased responsiveness to GH in Turner syndrome and short small-for-gestational-age children. J Clin Endocrinol Metab 91(2):659-664.

Blum WF, Machinis K, Shavrikova EP, Keller A, Stobbe H, Pfaeffle RW, and Amselem S. 2006. The growth response to growth hormone (GH) treatment in children with isolated GH deficiency is independent of the presence of the exon 3-minus isoform of the GH receptor. J Clin Endocrinol Metab 91(10):4171-4174.

Bolzer A, Kreth G, Solovei I, Koehler D, Saracoglu K, Fauth C, Muller S, Eils R, Cremer C, Speicher MR et al. . 2005. Three-dimensional maps of all chromosomes in human male fibroblast nuclei and prometaphase rosettes. PLoS Biol 3(5):e157.

Bonioli E, Taro M, Rosa CL, Citana A, Bertorelli R, Morcaldi G, Gastaldi R, and Coviello DA. 2005. Heterozygous mutations of growth hormone receptor gene in children with idiopathic short stature. Growth Horm IGF Res 15(6):405-410.

Bourque G. 2009. Transposable elements in gene regulation and in the evolution of vertebrate genomes. Curr Opin Genet Dev 19(6):607-612.

Brazeau P, Vale W, Burgus R, Ling N, Butcher M, Rivier J, and Guillemin R. 1973. Hypothalamic polypeptide that inhibits the secretion of immunoreactive pituitary growth hormone. Science 179(4068):77-79.

Britten RJ, and Kohne DE. 1968. Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. Science 161(3841):529-540.

Brooks AJ, and Waters MJ. 2010a. The growth hormone receptor: mechanism of activation and clinical implications. Nat Rev Endocrinol 6(9):515-525.

Brooks AJ, and Waters MJ. 2010b. The growth hormone receptor: mechanism of activation and clinical implications. Nat Rev Endocrinol 6(9):515-525.

Bruchim I, Attias Z, and Werner H. 2009. Targeting the IGF1 axis in cancer proliferation. Expert Opin Ther Targets 13(10):1179-1192.

Butler AA, and Le Roith D. 2001. Control of growth by the somatropic axis: growth hormone and the insulin-like growth factors have related and independent roles. Annu Rev Physiol 63:141-164.

Capatina C, and Wass JA. 2015. 60 YEARS OF NEUROENDOCRINOLOGY: Acromegaly. J Endocrinol 226(2):T141-160.

Carlsson LM, Attie KM, Compton PG, Vitangcol RV, and Merimee TJ. 1994. Reduced concentration of serum growth hormone-binding protein in children with idiopathic short stature. National Cooperative Growth Study. J Clin Endocrinol Metab 78(6):1325-1330.

Carrascosa A, Esteban C, Espadero R, Fernandez-Cancio M, Andaluz P, Clemente M, Audi L, Wollmann H, Fryklund L, and Parodi L. 2006. The d3/fl-growth hormone (GH) receptor polymorphism does not influence the effect of GH treatment (66 microg/kg per day) or the spontaneous growth in short non-GH-deficient small-for-gestational-age children: results from a two-year controlled prospective study in 170 Spanish patients. J Clin Endocrinol Metab 91(9):3281-3286.

Carroll JS. 2016. Mechanisms of oestrogen receptor (ER) gene regulation in breast cancer. Eur J Endocrinol 175(1):R41-49.

Carter-Su C, Schwartz J, and Argetsinger LS. 2016. Growth hormone signaling pathways. Growth Horm IGF Res 28:11-15.

Cataldi M, Magnan E, Guillaume V, Dutour A, Conte-Devolx B, Lombardi G, and Oliver C. 1994. Relationship between hypophyseal portal GHRH and somatostatin and peripheral GH levels in the conscious sheep. J Endocrinol Invest 17(9):717-722.

Cattini PA, Yang X, Jin Y, and Detillieux KA. 2006. Regulation of the human growth hormone gene family: possible role for Pit-1 in early stages of pituitary-specific expression and repression. Neuroendocrinology 83(3-4):145-153.

Chanoine JP, De Waele K, and Walia P. 2009. Ghrelin and the growth hormone secretagogue receptor in growth and development. Int J Obes (Lond) 33 Suppl 1:S48-52.

Chapman K, Takahashi A, Meulenbelt I, Watson C, Rodriguez-Lopez J, Egli R, Tsezou A, Malizos KN, Kloppenburg M, Shi D et al. . 2008. A meta-analysis of European and Asian cohorts reveals a global role of a functional SNP in the 5' UTR of GDF5 with osteoarthritis susceptibility. Hum Mol Genet 17(10):1497-1504.

Chaves VE, Junior FM, and Bertolini GL. 2013. The metabolic effects of growth hormone in adipose tissue. Endocrine 44(2):293-302.

Chen EY, Liao YC, Smith DH, Barrera-Saldana HA, Gelinas RE, and Seeburg PH. 1989. The human growth hormone locus: nucleotide sequence, biology, and evolution. Genomics 4(4):479-497.

Chen HY, Chan IH, Sham AL, Leung VH, Ma SL, Ho SC, and Tang NL. 2011. Haplotype effect in the IGF1 promoter accounts for the association between microsatellite and serum IGF1 concentration. Clin Endocrinol (Oxf) 74(4):520-527.

Chen HY, Ma SL, Huang W, Ji L, Leung VH, Jiang H, Yao X, and Tang NL. 2016. The mechanism of transactivation regulation due to polymorphic short tandem repeats (STRs) using IGF1 promoter as a model. Sci Rep 6:38225.

Chhabra Y, Wong HY, Nikolajsen LF, Steinocher H, Papadopulos A, Tunny KA, Meunier FA, Smith AG, Kragelund BB, Brooks AJ et al. . 2018. A growth hormone receptor SNP promotes lung cancer by impairment of SOCS2-mediated degradation. Oncogene 37(4):489-501.

Chia DJ, Varco-Merth B, and Rotwein P. 2010. Dispersed Chromosomal Stat5b-binding elements mediate growth hormone-activated insulin-like growth factor-I gene transcription. J Biol Chem 285(23):17636-17647.

Chikani V, and Ho KK. 2014. Action of GH on skeletal muscle function: molecular and metabolic mechanisms. J Mol Endocrinol 52(1):R107-123.

Chin E, Zhou J, Dai J, Baxter RC, and Bondy CA. 1994. Cellular localization and regulation of gene expression for components of the insulin-like growth factor ternary binding protein complex. Endocrinology 134(6):2498-2504.

Christiansen JS, Backeljauw PF, Bidlingmaier M, Biller BM, Boguszewski MC, Casanueva FF, Chanson P, Chatelain P, Choong CS, Clemmons DR et al. . 2016. Growth Hormone Research Society perspective on the development of long-acting growth hormone preparations. Eur J Endocrinol 174(6):C1-8.

Christoforidis A, Maniadaki I, and Stanhope R. 2005. Growth hormone / insulin-like growth factor-1 axis during puberty. Pediatr Endocrinol Rev 3(1):5-10.

Colosi P, Wong K, Leong SR, and Wood WI. 1993. Mutational analysis of the intracellular domain of the human growth hormone receptor. J Biol Chem 268(17):12617-12623.

Contreras B, and Talamantes F. 1999. Growth hormone (GH) and 17beta-estradiol regulation of the expression of mouse GH receptor and GH-binding protein in cultured mouse hepatocytes. Endocrinology 140(10):4725-4731.

Dahlgren J. 2011. Metabolic benefits of growth hormone therapy in idiopathic short stature. Horm Res Paediatr 76 Suppl 3:56-58.

Dalla Libera L, Ravara B, Volterrani M, Gobbo V, Della Barbera M, Angelini A, Danieli Betto D, Germinario E, and Vescovo G. 2004. Beneficial effects of GH/IGF-1 on skeletal muscle atrophy and function in experimental heart failure. Am J Physiol Cell Physiol 286(1):C138-144.

Dastot F, Sobrier ML, Duquesnoy P, Duriez B, Goossens M, and Amselem S. 1996. Alternatively spliced forms in the cytoplasmic domain of the human growth hormone (GH) receptor regulate its ability to generate a soluble GH-binding protein. Proceedings of the National Academy of Sciences of the United States of America 93(20):10723-10728.

David A, Hwa V, Metherell LA, Netchine I, Camacho-Hubner C, Clark AJ, Rosenfeld RG, and Savage MO. 2011. Evidence for a continuum of genetic, phenotypic, and biochemical abnormalities in children with growth hormone insensitivity. Endocr Rev 32(4):472-497.

de Vos AM, Ultsch M, and Kossiakoff AA. 1992. Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. Science 255(5042):306-312.

Deaton AM, and Bird A. 2011. CpG islands and the regulation of transcription. Genes Dev 25(10):1010-1022.

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

Deininger P. 2011. Alu elements: know the SINEs. Genome Biol 12(12):236.

Derr MA, Fang P, Sinha SK, Ten S, Hwa V, and Rosenfeld RG. 2011. A novel Y332C missense mutation in the intracellular domain of the human growth hormone receptor does not alter STAT5b signaling: redundancy of GHR intracellular tyrosines involved in STAT5b signaling. Horm Res Paediatr 75(3):187-199.

Dietz J, and Schwartz J. 1991. Growth hormone alters lipolysis and hormone-sensitive lipase activity in 3T3-F442A adipocytes. Metabolism 40(8):800-806.

Dineen R, Stewart PM, and Sherlock M. 2017. Acromegaly. QJM 110(7):411-420.

Domene HM, and Fierro-Carrion G. 2018. Genetic disorders of GH action pathway. Growth Horm IGF Res 38:19-23.

Domene HM, Scaglia PA, Martinez AS, Keselman AC, Karabatas LM, Pipman VR, Bengolea SV, Guida MC, Ropelato MG, Ballerini MG et al. . 2013. Heterozygous IGFALS gene variants in idiopathic short stature and normal children: impact on height and the IGF system. Horm Res Paediatr 80(6):413-423.

Dorr HG, Bettendorf M, Hauffa BP, Mehls O, Rohrer T, Stahnke N, Pfaffle R, and Ranke MB. 2011. Different relationships between the first 2 years on growth hormone treatment and the d3-growth hormone receptor polymorphism in short small-for-gestational-age (SGA) children. Clin Endocrinol (Oxf) 75(5):656-660.

Dos Santos C, Essioux L, Teinturier C, Tauber M, Goffin V, and Bougneres P. 2004. A common polymorphism of the growth hormone receptor is associated with increased responsiveness to growth hormone. Nat Genet 36(7):720-724.

Ellegren H. 2004. Microsatellites: simple sequences with complex evolution. Nat Rev Genet 5(6):435-445.

Elzein S, and Goodyer CG. 2014. Regulation of human growth hormone receptor expression by microRNAs. Mol Endocrinol 28(9):1448-1459.

Erman A, Veilleux A, Tchernof A, and Goodyer CG. 2011a. Human growth hormone receptor (GHR) expression in obesity: I. GHR mRNA expression in omental and subcutaneous adipose tissues of obese women. Int J Obes (Lond) 35(12):1511-1519.

Erman A, Wabitsch M, and Goodyer CG. 2011b. Human growth hormone receptor (GHR) expression in obesity: II. Regulation of the human GHR gene by obesity-related factors. Int J Obes (Lond) 35(12):1520-1529.

Eshet R, Laron Z, Pertzelan A, Arnon R, and Dintzman M. 1984. Defect of human growth hormone receptors in the liver of two patients with Laron-type dwarfism. Isr J Med Sci 20(1):8-11.

Fan Y, Menon RK, Cohen P, Hwang D, Clemens T, DiGirolamo DJ, Kopchick JJ, Le Roith D, Trucco M, and Sperling MA. 2009. Liver-specific deletion of the growth hormone

receptor reveals essential role of growth hormone signaling in hepatic lipid metabolism. J Biol Chem 284(30):19937-19944.

Fernandez-Perez L, de Mirecki-Garrido M, Guerra B, Diaz M, and Diaz-Chico JC. 2016. Sex steroids and growth hormone interactions. Endocrinol Nutr 63(4):171-180.

Feschotte C. 2008. Transposable elements and the evolution of regulatory networks. Nat Rev Genet 9(5):397-405.

Feschotte C, Keswani U, Ranganathan N, Guibotsy ML, and Levine D. 2009. Exploring repetitive DNA landscapes using REPCLASS, a tool that automates the classification of transposable elements in eukaryotic genomes. Genome Biol Evol 1:205-220.

Fofanova-Gambetti OV, Hwa V, Wit JM, Domene HM, Argente J, Bang P, Hogler W, Kirsch S, Pihoker C, Chiu HK et al. . 2010. Impact of heterozygosity for acid-labile subunit (IGFALS) gene mutations on stature: results from the international acid-labile subunit consortium. J Clin Endocrinol Metab 95(9):4184-4191.

Francis-West PH, Abdelfattah A, Chen P, Allen C, Parish J, Ladher R, Allen S, MacPherson S, Luyten FP, and Archer CW. 1999. Mechanisms of GDF-5 action during skeletal development. Development 126(6):1305-1315.

Frankenne F, Rentier-Delrue F, Scippo ML, Martial J, and Hennen G. 1987. Expression of the growth hormone variant gene in human placenta. J Clin Endocrinol Metab 64(3):635-637.

Friedman RC, Farh KK, Burge CB, and Bartel DP. 2009. Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 19(1):92-105.

Fukami M, Seki A, and Ogata T. 2016. SHOX Haploinsufficiency as a Cause of Syndromic and Nonsyndromic Short Stature. Mol Syndromol 7(1):3-11.

Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH et al. . 2009. An oestrogen-receptor-alpha-bound human chromatin interactome. Nature 462(7269):58-64.

Gabrielsson BG, Carmignac DF, Flavell DM, and Robinson IC. 1995. Steroid regulation of growth hormone (GH) receptor and GH-binding protein messenger ribonucleic acids in the rat. Endocrinology 136(1):209-217.

Gadelha MR, Kasuki L, and Korbonits M. 2017. The genetic background of acromegaly. Pituitary 20(1):10-21.

Gebre-Medhin M, Kindblom LG, Wennbo H, Tornell J, and Meis-Kindblom JM. 2001. Growth hormone receptor is expressed in human breast cancer. Am J Pathol 158(4):1217-1222.

Gemayel R, Cho J, Boeynaems S, and Verstrepen KJ. 2012. Beyond junk-variable tandem repeats as facilitators of rapid evolution of regulatory and coding sequences. Genes (Basel) 3(3):461-480.

Godowski PJ, Leung DW, Meacham LR, Galgani JP, Hellmiss R, Keret R, Rotwein PS, Parks JS, Laron Z, and Wood WI. 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(20):8083-8087.

Goldenberg N, and Barkan A. 2007. Factors regulating growth hormone secretion in humans. Endocrinol Metab Clin North Am 36(1):37-55.

Goodyer CG, Figueiredo RM, Krackovitch S, De Souza Li L, Manalo JA, and Zogopoulos G. 2001a. Characterization of the growth hormone receptor in human dermal fibroblasts and liver during development. Am J Physiol Endocrinol Metab 281(6):E1213-1220.

Goodyer CG, Rhani Z, and Zheng H. 2008. Expression of the hepatic specific V1 messenger ribonucleic acid of the human growth hormone receptor gene is regulated by hepatic nuclear factor (HNF)-4alpha2 and HNF-4alpha8. Mol Endocrinol 22(2):485-500.

Goodyer CG, Zheng H, and Hendy GN. 2001b. Alu elements in human growth hormone receptor gene 5' untranslated region exons. J Mol Endocrinol 27(3):357-366.

Goodyer CG, Zogopoulos G, Schwartzbauer G, Zheng H, Hendy GN, and Menon RK. 2001c. Organization and evolution of the human growth hormone receptor gene 5'-flanking region. Endocrinology 142(5):1923-1934.

Gorin E, and Goodman HM. 1985. Turnover of growth hormone receptors in rat adipocytes. Endocrinology 116(5):1796-1805.

Green H, Morikawa M, and Nixon T. 1985. A dual effector theory of growth-hormone action. Differentiation 29(3):195-198.

Greenhalgh CJ, Rico-Bautista E, Lorentzon M, Thaus AL, Morgan PO, Willson TA, Zervoudakis P, Metcalf D, Street I, Nicola NA et al. . 2005. SOCS2 negatively regulates growth hormone action in vitro and in vivo. J Clin Invest 115(2):397-406.

Grumbach MM. 2000. Estrogen, bone, growth and sex: a sea change in conventional wisdom. J Pediatr Endocrinol Metab 13 Suppl 6:1439-1455.

Guevara-Aguirre J, Balasubramanian P, Guevara-Aguirre M, Wei M, Madia F, Cheng CW, 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(70):70ra13.

Guevara-Aguirre J, Rosenbloom AL, Vasconez O, Martinez V, Gargosky SE, Allen L, and Rosenfeld RG. 1997. Two-year treatment of growth hormone (GH) receptor deficiency with recombinant insulin-like growth factor I in 22 children: comparison of two dosage levels and to GH-treated GH deficiency. J Clin Endocrinol Metab 82(2):629-633.

Guillemin R, Brazeau P, Bohlen P, Esch F, Ling N, and Wehrenberg WB. 1982. Growth hormone-releasing factor from a human pancreatic tumor that caused acromegaly. Science 218(4572):585-587.

Gunnell D, Okasha M, Smith GD, Oliver SE, Sandhu J, and Holly JM. 2001. Height, leg length, and cancer risk: a systematic review. Epidemiol Rev 23(2):313-342.

Gymrek M, Willems T, Guilmatre A, Zeng H, Markus B, Georgiev S, Daly MJ, Price AL, Pritchard JK, Sharp AJ et al. . 2016. Abundant contribution of short tandem repeats to gene expression variation in humans. Nat Genet 48(1):22-29.

Hadjiyannakis S, Zheng H, Hendy GN, and Goodyer CG. 2001. GT repeat polymorphism in the 5' flanking region of the human growth hormone receptor gene. Mol Cell Probes 15(4):239-242.

Han VK, Lund PK, Lee DC, and D'Ercole AJ. 1988. Expression of somatomedin/insulin-like growth factor messenger ribonucleic acids in the human fetus: identification, characterization, and tissue distribution. J Clin Endocrinol Metab 66(2):422-429.

Hannan AJ. 2010. Tandem repeat polymorphisms: modulators of disease susceptibility and candidates for 'missing heritability'. Trends Genet 26(2):59-65.

Hansen LH, Wang X, Kopchick JJ, Bouchelouche P, Nielsen JH, Galsgaard ED, and Billestrup N. 1996. Identification of tyrosine residues in the intracellular domain of the growth hormone receptor required for transcriptional signaling and Stat5 activation. J Biol Chem 271(21):12669-12673.

Harding PA, Wang XZ, Kelder B, Souza S, Okada S, and Kopchick JJ. 1994. In vitro mutagenesis of growth hormone receptor Asn-linked glycosylation sites. Mol Cell Endocrinol 106(1-2):171-180.

Heap D, Lucy MC, Collier RJ, Boyd CK, and Warren WC. 1995. Rapid communication: nucleotide sequence of the promoter and first exon of the somatotropin receptor gene in cattle. J Anim Sci 73(5):1529.

Herbert A, Lowenhaupt K, Spitzner J, and Rich A. 1995. Chicken double-stranded RNA adenosine deaminase has apparent specificity for Z-DNA. Proc Natl Acad Sci U S A 92(16):7550-7554.

Herbert AG, and Rich A. 1993. A method to identify and characterize Z-DNA binding proteins using a linear oligodeoxynucleotide. Nucleic Acids Res 21(11):2669-2672.

Hill DJ, Riley SC, Bassett NS, and Waters MJ. 1992. Localization of the growth hormone receptor, identified by immunocytochemistry, in second trimester human fetal tissues and in placenta throughout gestation. J Clin Endocrinol Metab 75(2):646-650.

Ho KY, Evans WS, Blizzard RM, Veldhuis JD, Merriam GR, Samojlik E, Furlanetto R, Rogol AD, Kaiser DL, and Thorner MO. 1987. Effects of sex and age on the 24-hour profile of growth hormone secretion in man: importance of endogenous estradiol concentrations. J Clin Endocrinol Metab 64(1):51-58.

Ho Y, Liebhaber SA, and Cooke NE. 2004. Activation of the human GH gene cluster: roles for targeted chromatin modification. Trends Endocrinol Metab 15(1):40-45.

Hoshina M, Boothby M, and Boime I. 1982. Cytological localization of chorionic gonadotropin alpha and placental lactogen mRNAs during development of the human placenta. J Cell Biol 93(1):190-198.

Howard AD, Feighner SD, Cully DF, Arena JP, Liberator PA, Rosenblum CI, Hamelin M, Hreniuk DL, Palyha OC, Anderson J et al. . 1996. A receptor in pituitary and hypothalamus that functions in growth hormone release. Science 273(5277):974-977.

Huang B, Yang J, Cheng Q, Xu P, Wang J, Zhang Z, Fan W, Wang P, and Yu M. 2018. Prognostic Value of HMGA2 in Human Cancers: A Meta-Analysis Based on Literatures and TCGA Datasets. Front Physiol 9:776.

Hujeirat Y, Hess O, Shalev S, and Tenenbaum-Rakover Y. 2006. Growth hormone receptor sequence changes do not play a role in determining height in children with idiopathic short stature. Horm Res 65(4):210-216.

Hwa V. 2016. STAT5B deficiency: Impacts on human growth and immunity. Growth Horm IGF Res 28:16-20.

Iida K, Takahashi Y, Kaji H, Nose O, Okimura Y, Abe H, and Chihara K. 1998. Growth hormone (GH) insensitivity syndrome with high serum GH-binding protein levels caused by a heterozygous splice site mutation of the GH receptor gene producing a lack of intracellular domain. J Clin Endocrinol Metab 83(2):531-537.

Iida K, Takahashi Y, Kaji H, Takahashi MO, Okimura Y, Nose O, Abe H, and Chihara K. 1999. Functional characterization of truncated growth hormone (GH) receptor-(1-277) causing partial GH insensitivity syndrome with high GH-binding protein. J Clin Endocrinol Metab 84(3):1011-1016.

Ikeno Y, Hubbard GB, Lee S, Cortez LA, Lew CM, Webb CR, Berryman DE, List EO, Kopchick JJ, and Bartke A. 2009. Reduced incidence and delayed occurrence of fatal neoplastic diseases in growth hormone receptor/binding protein knockout mice. J Gerontol A Biol Sci Med Sci 64(5):522-529.

Illingworth RS, and Bird AP. 2009. CpG islands--'a rough guide'. FEBS Lett 583(11):1713-1720.

Isaksson OG, Jansson JO, and Gause IA. 1982. Growth hormone stimulates longitudinal bone growth directly. Science 216(4551):1237-1239.

Ito K, Noguchi A, Uosaki Y, Taga T, Arakawa H, and Takizawa T. 2018. Gfap and Osmr regulation by BRG1 and STAT3 via interchromosomal gene clustering in astrocytes. Mol Biol Cell 29(2):209-219.

Itoh E, Iida K, del Rincon JP, Kim DS, and Thorner MO. 2004. Diurnal variation in growth hormone receptor messenger ribonucleic acid in liver and skeletal muscle of lit/+ and lit/lit mice. Endocr J 51(6):529-535.

Itokawa M, Yamada K, Yoshitsugu K, Toyota T, Suga T, Ohba H, Watanabe A, Hattori E, Shimizu H, Kumakura T et al. . 2003. A microsatellite repeat in the promoter of the N-methyl-D-aspartate receptor 2A subunit (GRIN2A) gene suppresses transcriptional

activity and correlates with chronic outcome in schizophrenia. Pharmacogenetics 13(5):271-278.

Jaffe CA, Ocampo-Lim B, Guo W, Krueger K, Sugahara I, DeMott-Friberg R, Bermann M, and Barkan AL. 1998. Regulatory mechanisms of growth hormone secretion are sexually dimorphic. J Clin Invest 102(1):153-164.

Jiang H, Okamura CS, and Lucy MC. 1999. Isolation and characterization of a novel promoter for the bovine growth hormone receptor gene. J Biol Chem 274(12):7893-7900.

Jin Y, Vakili H, Liu S, Menticoglou S, Bock ME, and Cattini PA. 2018. Chromosomal architecture and placental expression of the human growth hormone gene family are targeted by pre-pregnancy maternal obesity. Am J Physiol Endocrinol Metab.

Johnston LB, Ester W, Caliebe J, Molinas C, Wollmann H, Fryklund L, Clark AJ, Ranke MB, Tauber M, Hokken Koelega A et al. . 2009. Network of European studies of genes in growth (NESTEGG). Horm Res 71 Suppl 2:48-54.

Jones BK, Monks BR, Liebhaber SA, and Cooke NE. 1995. The human growth hormone gene is regulated by a multicomponent locus control region. Mol Cell Biol 15(12):7010-7021.

Jorge AA, Marchisotti FG, Montenegro LR, Carvalho LR, Mendonca BB, and Arnhold IJ. 2006. Growth hormone (GH) pharmacogenetics: influence of GH receptor exon 3 retention or deletion on first-year growth response and final height in patients with severe GH deficiency. J Clin Endocrinol Metab 91(3):1076-1080.

Kadakia R, and Josefson J. 2016. The Relationship of Insulin-Like Growth Factor 2 to Fetal Growth and Adiposity. Horm Res Paediatr 85(2):75-82.

Kang MJ. 2017. Novel genetic cause of idiopathic short stature. Ann Pediatr Endocrinol Metab 22(3):153-157.

Kasukawa Y, Miyakoshi N, and Mohan S. 2004. The anabolic effects of GH/IGF system on bone. Curr Pharm Des 10(21):2577-2592.

Katti MV, Ranjekar PK, and Gupta VS. 2001. Differential distribution of simple sequence repeats in eukaryotic genome sequences. Mol Biol Evol 18(7):1161-1167.

Kenth G, Puzhko S, and Goodyer CG. 2011. Human growth hormone receptor gene expression is regulated by Gfi-1/1b and GAGA cis-elements. Mol Cell Endocrinol 335(2):135-147.

Kenth G, Shao Z, Cole DE, and Goodyer CG. 2007. Relationship of the human growth hormone receptor exon 3 genotype with final adult height and bone mineral density. J Clin Endocrinol Metab 92(2):725-728.

Kershaw NJ, Murphy JM, Liau NP, Varghese LN, Laktyushin A, Whitlock EL, Lucet IS, Nicola NA, and Babon JJ. 2013. SOCS3 binds specific receptor-JAK complexes to control cytokine signaling by direct kinase inhibition. Nat Struct Mol Biol 20(4):469-476.

Kim YG, Muralinath M, Brandt T, Pearcy M, Hauns K, Lowenhaupt K, Jacobs BL, and Rich A. 2003. A role for Z-DNA binding in vaccinia virus pathogenesis. Proc Natl Acad Sci U S A 100(12):6974-6979.

Lafer EM, Moller A, Nordheim A, Stollar BD, and Rich A. 1981. Antibodies specific for left-handed Z-DNA. Proc Natl Acad Sci U S A 78(6):3546-3550.

Lafer EM, Valle RP, Moller A, Nordheim A, Schur PH, Rich A, and Stollar BD. 1983. Z-DNA-specific antibodies in human systemic lupus erythematosus. J Clin Invest 71(2):314-321.

Lahbib-Mansais Y, Barasc H, Marti-Marimon M, Mompart F, Iannuccelli E, Robelin D, Riquet J, and Yerle-Bouissou M. 2016. Expressed alleles of imprinted IGF2, DLK1 and MEG3 colocalize in 3D-preserved nuclei of porcine fetal cells. BMC Cell Biol 17(1):35.

Lander ES, Linton LM, Birren B, Nusbaum C, Zody MC, Baldwin J, Devon K, Dewar K, Doyle M, FitzHugh W et al. . 2001. Initial sequencing and analysis of the human genome. Nature 409(6822):860-921.

Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F, Willer CJ, Jackson AU, Vedantam S, Raychaudhuri S et al. . 2010. Hundreds of variants clustered in genomic loci and biological pathways affect human height. Nature 467(7317):832-838.

Lanktree MB, Guo Y, Murtaza M, Glessner JT, Bailey SD, Onland-Moret NC, Lettre G, Ongen H, Rajagopalan R, Johnson T et al. . 2011. Meta-analysis of Dense Genecentric Association Studies Reveals Common and Uncommon Variants Associated with Height. Am J Hum Genet 88(1):6-18.

Lanning NJ, and Carter-Su C. 2006. Recent advances in growth hormone signaling. Reviews in endocrine & metabolic disorders 7(4):225-235.

Laron Z. 2004. Laron syndrome (primary growth hormone resistance or insensitivity): the personal experience 1958-2003. J Clin Endocrinol Metab 89(3):1031-1044.

Laron Z, Pertzelan A, and Mannheimer S. 1966. Genetic pituitary dwarfism with high serum concentation of growth hormone--a new inborn error of metabolism? Isr J Med Sci 2(2):152-155.

Lau-Corona D, Suvorov A, and Waxman DJ. 2017. Feminization of male mouse liver by persistent growth hormone stimulation: Activation of sex-biased transcriptional networks and dynamic changes in chromatin states. Mol Cell Biol.

Le Roith D, Bondy C, Yakar S, Liu JL, and Butler A. 2001. The somatomedin hypothesis: 2001. Endocr Rev 22(1):53-74.

Lefterova MI, Zhang Y, Steger DJ, Schupp M, Schug J, Cristancho A, Feng D, Zhuo D, Stoeckert CJ, Jr., Liu XS et al. . 2008. PPARgamma and C/EBP factors orchestrate adipocyte biology via adjacent binding on a genome-wide scale. Genes Dev 22(21):2941-2952.

Legendre M, Pochet N, Pak T, and Verstrepen KJ. 2007. Sequence-based estimation of minisatellite and microsatellite repeat variability. Genome Res 17(12):1787-1796.

Legraverend C, Simar-Blanchet AE, Paul C, Sotiropoulos A, Finidori J, and Cam AL. 1996. A novel growth hormone response element unrelated to STAT (signal transducer and activator of transcription)-binding sites is a bifunctional enhancer. Mol Endocrinol 10(12):1507-1518.

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(3):302-310.

Lettre G. 2011. Recent progress in the study of the genetics of height. Hum Genet 129(5):465-472.

Leung DW, Spencer SA, Cachianes G, Hammonds RG, Collins C, Henzel WJ, Barnard R, Waters MJ, and Wood WI. 1987. Growth hormone receptor and serum binding protein: purification, cloning and expression. Nature 330(6148):537-543.

Leung KC, Johannsson G, Leong GM, and Ho KK. 2004. Estrogen regulation of growth hormone action. Endocr Rev 25(5):693-721.

Li L, He S, Sun JM, and Davie JR. 2004. Gene regulation by Sp1 and Sp3. Biochem Cell Biol 82(4):460-471.

Lichanska AM, and Waters MJ. 2008a. How growth hormone controls growth, obesity and sexual dimorphism. Trends Genet 24(1):41-47.

Lichanska AM, and Waters MJ. 2008b. New insights into growth hormone receptor function and clinical implications. Horm Res 69(3):138-145.

Lindsey RC, and Mohan S. 2016. Skeletal effects of growth hormone and insulin-like growth factor-I therapy. Mol Cell Endocrinol 432:44-55.

Ling G, Sugathan A, Mazor T, Fraenkel E, and Waxman DJ. 2010. Unbiased, genome-wide in vivo mapping of transcriptional regulatory elements reveals sex differences in chromatin structure associated with sex-specific liver gene expression. Mol Cell Biol 30(23):5531-5544.

Linossi EM, Babon JJ, Hilton DJ, and Nicholson SE. 2013. Suppression of cytokine signaling: the SOCS perspective. Cytokine Growth Factor Rev 24(3):241-248.

List EO, Sackmann-Sala L, Berryman DE, Funk K, Kelder B, Gosney ES, Okada S, Ding J, Cruz-Topete D, and Kopchick JJ. 2011. Endocrine parameters and phenotypes of the growth hormone receptor gene disrupted (GHR-/-) mouse. Endocr Rev 32(3):356-386.

Liu H, Mulholland N, Fu H, and Zhao K. 2006. Cooperative activity of BRG1 and Z-DNA formation in chromatin remodeling. Mol Cell Biol 26(7):2550-2559.

Liu JL, and LeRoith D. 1999. Insulin-like growth factor I is essential for postnatal growth in response to growth hormone. Endocrinology 140(11):5178-5184.

Liu R, Liu H, Chen X, Kirby M, Brown PO, and Zhao K. 2001. Regulation of CSF1 promoter by the SWI/SNF-like BAF complex. Cell 106(3):309-318.

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

MacLeod JN, Pampori NA, and Shapiro BH. 1991. Sex differences in the ultradian pattern of plasma growth hormone concentrations in mice. J Endocrinol 131(3):395-399.

Maiter D, Underwood LE, Maes M, Davenport ML, and Ketelslegers JM. 1988. Different effects of intermittent and continuous growth hormone (GH) administration on serum somatomedin-C/insulin-like growth factor I and liver GH receptors in hypophysectomized rats. Endocrinology 123(2):1053-1059.

Manolagas SC. 2000. Birth and death of bone cells: basic regulatory mechanisms and implications for the pathogenesis and treatment of osteoporosis. Endocr Rev 21(2):115-137.

Marouli E, Graff M, Medina-Gomez C, Lo KS, Wood AR, Kjaer TR, Fine RS, Lu Y, Schurmann C, Highland HM et al. . 2017. Rare and low-frequency coding variants alter human adult height. Nature 542(7640):186-190.

Marstrand-Joergensen MR, Jensen RB, Aksglaede L, Duno M, and Juul A. 2017. Prevalence of SHOX haploinsufficiency among short statured children. Pediatr Res 81(2):335-341.

Maruyama A, Mimura J, Harada N, and Itoh K. 2013. Nrf2 activation is associated with Z-DNA formation in the human HO-1 promoter. Nucleic Acids Res 41(10):5223-5234.

McKay JD, Kaaks R, Johansson M, Biessy C, Wiklund F, Balter K, Adami HO, Boillot C, Gioia-Patricola L, Canzian F et al. . 2007. Haplotype-based analysis of common variation in the growth hormone receptor gene and prostate cancer risk. Cancer Epidemiol Biomarkers Prev 16(1):169-173.

McNally R, Toms AV, and Eck MJ. 2016. Crystal Structure of the FERM-SH2 Module of Human Jak2. PLoS One 11(5):e0156218.

Menon RK, Shaufl A, Yu JH, Stephan DA, and Friday RP. 2001. Identification and characterization of a novel transcript of the murine growth hormone receptor gene exhibiting development- and tissue-specific expression. Mol Cell Endocrinol 172(1-2):135-146.

Menten B, Buysse K, Zahir F, Hellemans J, Hamilton SJ, Costa T, Fagerstrom C, Anadiotis G, Kingsbury D, McGillivray BC et al. . 2007. Osteopoikilosis, short stature and mental retardation as key features of a new microdeletion syndrome on 12q14. J Med Genet 44(4):264-268.

Metcalf D, Greenhalgh CJ, Viney E, Willson TA, Starr R, Nicola NA, Hilton DJ, and Alexander WS. 2000. Gigantism in mice lacking suppressor of cytokine signalling-2. Nature 405(6790):1069-1073.

Metherell LA, Akker SA, Munroe PB, Rose SJ, Caulfield M, Savage MO, Chew SL, and Clark AJ. 2001. Pseudoexon activation as a novel mechanism for disease resulting in atypical growth-hormone insensitivity. Am J Hum Genet 69(3):641-646.

Mirkin SM. 2007. Expandable DNA repeats and human disease. Nature 447(7147):932-940.

Miyamoto Y, Mabuchi A, Shi D, Kubo T, Takatori Y, Saito S, Fujioka M, Sudo A, Uchida A, Yamamoto S et al. . 2007. A functional polymorphism in the 5' UTR of GDF5 is associated with susceptibility to osteoarthritis. Nat Genet 39(4):529-533.

Moffat JG, Dao H, and Talamantes F. 2000. Alternative 5'-untranslated regions of mouse GH receptor/binding protein messenger RNA are derived from sequences adjacent to the major L2 promoter. J Endocrinol 167(1):145-152.

Moffat JG, Edens A, and Talamantes F. 1999. Structure and expression of the mouse growth hormone receptor/growth hormone binding protein gene. J Mol Endocrinol 23(1):33-44.

Mohseni-Zadeh S, and Binoux M. 1997. Insulin-like growth factor (IGF) binding protein-3 interacts with the type 1 IGF receptor, reducing the affinity of the receptor for its ligand: an alternative mechanism in the regulation of IGF action. Endocrinology 138(12):5645-5648.

Moller N, and Jorgensen JO. 2009. Effects of growth hormone on glucose, lipid, and protein metabolism in human subjects. Endocr Rev 30(2):152-177.

Montefusco L, Filopanti M, Ronchi CL, Olgiati L, La-Porta C, Losa M, Epaminonda P, Coletti F, Beck-Peccoz P, Spada A et al. . 2010. d3-Growth hormone receptor polymorphism in acromegaly: effects on metabolic phenotype. Clin Endocrinol (Oxf) 72(5):661-667.

Murphy KG, and Bloom SR. 2006. Gut hormones and the regulation of energy homeostasis. Nature 444(7121):854-859.

Murray PG, and Clayton PE. 2013. Endocrine control of growth. Am J Med Genet C Semin Med Genet 163C(2):76-85.

Nakazato M, Murakami N, Date Y, Kojima M, Matsuo H, Kangawa K, and Matsukura S. 2001. A role for ghrelin in the central regulation of feeding. Nature 409(6817):194-198.

Newbern D, and Freemark M. 2011. Placental hormones and the control of maternal metabolism and fetal growth. Curr Opin Endocrinol Diabetes Obes 18(6):409-416.

Ng TK, Lam CY, Lam DS, Chiang SW, Tam PO, Wang DY, Fan BJ, Yam GH, Fan DS, and Pang CP. 2009. AC and AG dinucleotide repeats in the PAX6 P1 promoter are associated with high myopia. Mol Vis 15:2239-2248.

Nikumbh S, and Pfeifer N. 2017. Genetic sequence-based prediction of long-range chromatin interactions suggests a potential role of short tandem repeat sequences in genome organization. BMC Bioinformatics 18(1):218.

O'Mahoney JV, Brandon MR, and Adams TE. 1994. Identification of a liver-specific promoter for the ovine growth hormone receptor. Mol Cell Endocrinol 101(1-2):129-139.

Oberbauer AM, and Peng R. 1995. Fractionation of growth plate chondrocytes: differential expression of IGF-I and growth hormone and IGF-I receptor mRNA in purified populations. Connect Tissue Res 31(3):179-187.

Ocaranza P, Morales F, Roman R, Iniguez G, and Fernando C. 2012. Expression of SOCS1, SOCS2, and SOCS3 in growth hormone-stimulated skin fibroblasts from children with idiopathic short stature. J Pediatr Endocrinol Metab 25(3-4):273-278.

Ohlsson C, Nilsson A, Isaksson O, and Lindahl A. 1992. Growth hormone induces multiplication of the slowly cycling germinal cells of the rat tibial growth plate. Proc Natl Acad Sci U S A 89(20):9826-9830.

Orgel LE, and Crick FH. 1980. Selfish DNA: the ultimate parasite. Nature 284(5757):604-607.

Orlovskii IV, Sverdlova PS, and Rubtsov PM. 2004. [Fine structure, expression and polymorphism of the human growth hormone receptor gene]. Mol Biol (Mosk) 38(1):29-39.

Ottosson M, Vikman-Adolfsson K, Enerback S, Elander A, Bjorntorp P, and Eden S. 1995. Growth hormone inhibits lipoprotein lipase activity in human adipose tissue. J Clin Endocrinol Metab 80(3):936-941.

Pantel J, Grulich-Henn J, Bettendorf M, Strasburger CJ, Heinrich U, and Amselem S. 2003. Heterozygous nonsense mutation in exon 3 of the growth hormone receptor (GHR) in severe GH insensitivity (Laron syndrome) and the issue of the origin and function of the GHRd3 isoform. J Clin Endocrinol Metab 88(4):1705-1710.

Pantel J, Legendre M, Cabrol S, Hilal L, Hajaji Y, Morisset S, Nivot S, Vie-Luton MP, Grouselle D, de Kerdanet M et al. . 2006. Loss of constitutive activity of the growth hormone secretagogue receptor in familial short stature. J Clin Invest 116(3):760-768.

Pantel J, Machinis K, Sobrier ML, Duquesnoy P, Goossens M, and Amselem S. 2000. Species-specific alternative splice mimicry at the growth hormone receptor locus revealed by the lineage of retroelements during primate evolution. J Biol Chem 275(25):18664-18669.

Pasquali C, Curchod ML, Walchli S, Espanel X, Guerrier M, Arigoni F, Strous G, and Hooft van Huijsduijnen R. 2003. Identification of protein tyrosine phosphatases with specificity for the ligand-activated growth hormone receptor. Mol Endocrinol 17(11):2228-2239.

Paulson H. 2018. Repeat expansion diseases. Handb Clin Neurol 147:105-123.

Pekhletsky RI, Chernov BK, and Rubtsov PM. 1992. Variants of the 5'-untranslated sequence of human growth hormone receptor mRNA. Mol Cell Endocrinol 90(1):103-109.

Perry JK, Emerald BS, Mertani HC, and Lobie PE. 2006. The oncogenic potential of growth hormone. Growth Horm IGF Res 16(5-6):277-289.

Pilecka I, Patrignani C, Pescini R, Curchod ML, Perrin D, Xue Y, Yasenchak J, Clark A, Magnone MC, Zaratin P et al. . 2007. Protein-tyrosine phosphatase H1 controls growth hormone receptor signaling and systemic growth. J Biol Chem 282(48):35405-35415.

Pilotta A, Mella P, Filisetti M, Felappi B, Prandi E, Parrinello G, Notarangelo LD, and Buzi F. 2006. Common polymorphisms of the growth hormone (GH) receptor do not correlate with the growth response to exogenous recombinant human GH in GH-deficient children. J Clin Endocrinol Metab 91(3):1178-1180.

Piwien Pilipuk G, Galigniana MD, and Schwartz J. 2003. Subnuclear localization of C/EBP beta is regulated by growth hormone and dependent on MAPK. J Biol Chem 278(37):35668-35677.

Poger D, and Mark AE. 2010. Turning the growth hormone receptor on: evidence that hormone binding induces subunit rotation. Proteins 78(5):1163-1174.

Polinkovsky A, Robin NH, Thomas JT, Irons M, Lynn A, Goodman FR, Reardon W, Kant SG, Brunner HG, van der Burgt I et al. . 1997. Mutations in CDMP1 cause autosomal dominant brachydactyly type C. Nat Genet 17(1):18-19.

Pugliese-Pires PN, Fortin JP, Arthur T, Latronico AC, Mendonca BB, Villares SM, Arnhold IJ, Kopin AS, and Jorge AA. 2011. Novel inactivating mutations in the GH secretagogue receptor gene in patients with constitutional delay of growth and puberty. Eur J Endocrinol 165(2):233-241.

Quilez J, Guilmatre A, Garg P, Highnam G, Gymrek M, Erlich Y, Joshi RS, Mittelman D, and Sharp AJ. 2016. Polymorphic tandem repeats within gene promoters act as modifiers of gene expression and DNA methylation in humans. Nucleic Acids Res 44(8):3750-3762.

Raisman G, and Field PM. 1973. Sexual dimorphism in the neuropil of the preoptic area of the rat and its dependence on neonatal androgen. Brain Res 54:1-29.

Ray BK, Dhar S, Shakya A, and Ray A. 2011. Z-DNA-forming silencer in the first exon regulates human ADAM-12 gene expression. Proc Natl Acad Sci U S A 108(1):103-108.

Reh CS, and Geffner ME. 2010. Somatotropin in the treatment of growth hormone deficiency and Turner syndrome in pediatric patients: a review. Clin Pharmacol 2:111-122.

Rich A, Nordheim A, and Wang AH. 1984. The chemistry and biology of left-handed Z-DNA. Annu Rev Biochem 53:791-846.

Rich A, and Zhang S. 2003. Timeline: Z-DNA: the long road to biological function. Nat Rev Genet 4(7):566-572.

Richelsen B, Pedersen SB, Kristensen K, Borglum JD, Norrelund H, Christiansen JS, and Jorgensen JO. 2000. Regulation of lipoprotein lipase and hormone-sensitive lipase activity and gene expression in adipose and muscle tissue by growth hormone treatment during weight loss in obese patients. Metabolism 49(7):906-911.

Rife T, Rasoul B, Pullen N, Mitchell D, Grathwol K, and Kurth J. 2009. The effect of a promoter polymorphism on the transcription of nitric oxide synthase 1 and its relevance to Parkinson's disease. J Neurosci Res 87(10):2319-2325.

Rinn JL, and Snyder M. 2005. Sexual dimorphism in mammalian gene expression. Trends Genet 21(5):298-305.

Roberts CT, Jr., Lasky SR, Lowe WL, Jr., Seaman WT, and LeRoith D. 1987. Molecular cloning of rat insulin-like growth factor I complementary deoxyribonucleic acids:

differential messenger ribonucleic acid processing and regulation by growth hormone in extrahepatic tissues. Mol Endocrinol 1(3):243-248.

Rockman MV, and Wray GA. 2002. Abundant raw material for cis-regulatory evolution in humans. Mol Biol Evol 19(11):1991-2004.

Rogol AD. 2010. Sex steroids, growth hormone, leptin and the pubertal growth spurt. Endocr Dev 17:77-85.

Rose-John S, and Heinrich PC. 1994. Soluble receptors for cytokines and growth factors: generation and biological function. Biochem J 300 (Pt 2):281-290.

Rosenbloom AL. 2009. Mecasermin (recombinant human insulin-like growth factor I). Adv Ther 26(1):40-54.

Rosilio M, Huber-Lequesne C, Sapin H, Carel JC, Blum WF, and Cormier-Daire V. 2012. Genotypes and phenotypes of children with SHOX deficiency in France. J Clin Endocrinol Metab 97(7):E1257-1265.

Ross RJ, Esposito N, Shen XY, Von Laue S, Chew SL, Dobson PR, Postel-Vinay MC, and Finidori J. 1997. A short isoform of the human growth hormone receptor functions as a dominant negative inhibitor of the full-length receptor and generates large amounts of binding protein. Mol Endocrinol 11(3):265-273.

Ross RJ, Leung KC, Maamra M, Bennett W, Doyle N, Waters MJ, and Ho KK. 2001. Binding and functional studies with the growth hormone receptor antagonist, B2036-PEG (pegvisomant), reveal effects of pegylation and evidence that it binds to a receptor dimer. J Clin Endocrinol Metab 86(4):1716-1723.

Rothenburg S, Deigendesch N, Dittmar K, Koch-Nolte F, Haag F, Lowenhaupt K, and Rich A. 2005. A PKR-like eukaryotic initiation factor 2alpha kinase from zebrafish contains Z-DNA binding domains instead of dsRNA binding domains. Proc Natl Acad Sci U S A 102(5):1602-1607.

Rothenburg S, Koch-Nolte F, Rich A, and Haag F. 2001. A polymorphic dinucleotide repeat in the rat nucleolin gene forms Z-DNA and inhibits promoter activity. Proc Natl Acad Sci U S A 98(16):8985-8990.

Rowland JE, Lichanska AM, Kerr LM, White M, d'Aniello EM, Maher SL, Brown R, Teasdale RD, Noakes PG, and Waters MJ. 2005. In vivo analysis of growth hormone receptor signaling domains and their associated transcripts. Mol Cell Biol 25(1):66-77.

Rowlinson SW, Behncken SN, Rowland JE, Clarkson RW, Strasburger CJ, Wu Z, Baumbach W, and Waters MJ. 1998. Activation of chimeric and full-length growth hormone receptors by growth hormone receptor monoclonal antibodies. A specific conformational change may be required for full-length receptor signaling. J Biol Chem 273(9):5307-5314.

Rowlinson SW, Yoshizato H, Barclay JL, Brooks AJ, Behncken SN, Kerr LM, Millard K, Palethorpe K, Nielsen K, Clyde-Smith J et al. . 2008. An agonist-induced conformational change in the growth hormone receptor determines the choice of signalling pathway. Nature cell biology 10(6):740-747.

Rudd MF, Webb EL, Matakidou A, Sellick GS, Williams RD, Bridle H, Eisen T, and Houlston RS. 2006. Variants in the GH-IGF axis confer susceptibility to lung cancer. Genome Res 16(6):693-701.

Salmon WD, Jr., and Daughaday WH. 1957. A hormonally controlled serum factor which stimulates sulfate incorporation by cartilage in vitro. J Lab Clin Med 49(6):825-836.

Sanna S, Jackson AU, Nagaraja R, Willer CJ, Chen WM, Bonnycastle LL, Shen H, Timpson N, Lettre G, Usala G et al. . 2008. Common variants in the GDF5-UQCC region are associated with variation in human height. Nat Genet 40(2):198-203.

Sasaki A, Yasukawa H, Suzuki A, Kamizono S, Syoda T, Kinjyo I, Sasaki M, Johnston JA, and Yoshimura A. 1999. Cytokine-inducible SH2 protein-3 (CIS3/SOCS3) inhibits Janus tyrosine kinase by binding through the N-terminal kinase inhibitory region as well as SH2 domain. Genes Cells 4(6):339-351.

Sassin JF, Parker DC, Mace JW, Gotlin RW, Johnson LC, and Rossman LG. 1969. Human growth hormone release: relation to slow-wave sleep and sleep-walking cycles. Science 165(3892):513-515.

Savage MO, Attie KM, David A, Metherell LA, Clark AJ, and Camacho-Hubner C. 2006. Endocrine assessment, molecular characterization and treatment of growth hormone insensitivity disorders. Nat Clin Pract Endocrinol Metab 2(7):395-407.

Savage MO, Burren CP, and Rosenfeld RG. 2010. The continuum of growth hormone-IGF-I axis defects causing short stature: diagnostic and therapeutic challenges. Clin Endocrinol (Oxf) 72(6):721-728.

Savage MO, Hwa V, David A, Rosenfeld RG, and Metherell LA. 2011. Genetic Defects in the Growth Hormone-IGF-I Axis Causing Growth Hormone Insensitivity and Impaired Linear Growth. Front Endocrinol (Lausanne) 2:95.

Sawaya SM, Bagshaw AT, Buschiazzo E, and Gemmell NJ. 2012a. Promoter microsatellites as modulators of human gene expression. Advances in experimental medicine and biology 769:41-54.

Sawaya SM, Lennon D, Buschiazzo E, Gemmell N, and Minin VN. 2012b. Measuring microsatellite conservation in mammalian evolution with a phylogenetic birth-death model. Genome Biol Evol 4(6):636-647.

Scavo LM, Karas M, Murray M, and Leroith D. 2004. Insulin-like growth factor-I stimulates both cell growth and lipogenesis during differentiation of human mesenchymal stem cells into adipocytes. J Clin Endocrinol Metab 89(7):3543-3553.

Schally AV, Dupont A, Arimura A, Redding TW, Nishi N, Linthicum GL, and Schlesinger DH. 1976. Isolation and structure of somatostatin from porcine hypothalami. Biochemistry 15(3):509-514.

Schilbach K, and Bidlingmaier M. 2015. Growth hormone binding protein - physiological and analytical aspects. Best Pract Res Clin Endocrinol Metab 29(5):671-683.

Schilling E, and Rehli M. 2007. Global, comparative analysis of tissue-specific promoter CpG methylation. Genomics 90(3):314-323.

Schlechter NL, Russell SM, Greenberg S, Spencer EM, and Nicoll CS. 1986. A direct growth effect of growth hormone in rat hindlimb shown by arterial infusion. Am J Physiol 250(3 Pt 1):E231-235.

Schroth GP, Chou PJ, and Ho PS. 1992. Mapping Z-DNA in the human genome. Computer-aided mapping reveals a nonrandom distribution of potential Z-DNA-forming sequences in human genes. J Biol Chem 267(17):11846-11855.

Schwartz T, Behlke J, Lowenhaupt K, Heinemann U, and Rich A. 2001. Structure of the DLM-1-Z-DNA complex reveals a conserved family of Z-DNA-binding proteins. Nat Struct Biol 8(9):761-765.

Schwartzbauer G, and Menon RK. 1998. Regulation of growth hormone receptor gene expression. Mol Genet Metab 63(4):243-253.

Searle S, and Blackwell JM. 1999. Evidence for a functional repeat polymorphism in the promoter of the human NRAMP1 gene that correlates with autoimmune versus infectious disease susceptibility. J Med Genet 36(4):295-299.

Shewchuk BM, Liebhaber SA, and Cooke NE. 2002. Specification of unique Pit-1 activity in the hGH locus control region. Proc Natl Acad Sci U S A 99(18):11784-11789.

Shimajiri S, Arima N, Tanimoto A, Murata Y, Hamada T, Wang KY, and Sasaguri Y. 1999. Shortened microsatellite d(CA)21 sequence down-regulates promoter activity of matrix metalloproteinase 9 gene. FEBS Lett 455(1-2):70-74.

Shin SI, Ham S, Park J, Seo SH, Lim CH, Jeon H, Huh J, and Roh TY. 2016. Z-DNA-forming sites identified by ChIP-Seq are associated with actively transcribed regions in the human genome. DNA Res.

Silha JV, Gui Y, Modric T, Suwanichkul A, Durham SK, Powell DR, and Murphy LJ. 2001. Overexpression of the acid-labile subunit of the IGF ternary complex in transgenic mice. Endocrinology 142(10):4305-4313.

Simard M, Manthos H, Giaid A, Lefebvre Y, and Goodyer CG. 1996. Ontogeny of growth hormone receptors in human tissues: an immunohistochemical study. J Clin Endocrinol Metab 81(8):3097-3102.

Slootweg MC, Swolin D, Netelenbos JC, Isaksson OG, and Ohlsson C. 1997. Estrogen enhances growth hormone receptor expression and growth hormone action in rat osteosarcoma cells and human osteoblast-like cells. J Endocrinol 155(1):159-164.

Smuel K, Kauli R, Lilos P, and Laron Z. 2015. Growth, development, puberty and adult height before and during treatment in children with congenital isolated growth hormone deficiency. Growth Horm IGF Res 25(4):182-188.

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(4):485-489.

Steyn FJ, Huang L, Ngo ST, Leong JW, Tan HY, Xie TY, Parlow AF, Veldhuis JD, Waters MJ, and Chen C. 2011. Development of a method for the determination of pulsatile growth hormone secretion in mice. Endocrinology 152(8):3165-3171.

Steyn FJ, Tolle V, Chen C, and Epelbaum J. 2016. Neuroendocrine Regulation of Growth Hormone Secretion. Compr Physiol 6(2):687-735.

Strous GJ, van Kerkhof P, Govers R, Ciechanover A, and Schwartz AL. 1996. The ubiquitin conjugation system is required for ligand-induced endocytosis and degradation of the growth hormone receptor. EMBO J 15(15):3806-3812.

Sugathan A, and Waxman DJ. 2013. Genome-wide analysis of chromatin states reveals distinct mechanisms of sex-dependent gene regulation in male and female mouse liver. Mol Cell Biol 33(18):3594-3610.

Sun Y, Ahmed S, and Smith RG. 2003. Deletion of ghrelin impairs neither growth nor appetite. Mol Cell Biol 23(22):7973-7981.

Taka S, Gazouli M, Politis PK, Pappa KI, and Anagnou NP. 2013. Transcription factor ATF-3 regulates allele variation phenotypes of the human SLC11A1 gene. Mol Biol Rep 40(3):2263-2271.

Takahashi Y, Kipnis DM, and Daughaday WH. 1968. Growth hormone secretion during sleep. J Clin Invest 47(9):2079-2090.

Tannenbaum GS, and Martin JB. 1976. Evidence for an endogenous ultradian rhythm governing growth hormone secretion in the rat. Endocrinology 98(3):562-570.

Tartaglia M, Mehler EL, Goldberg R, Zampino G, Brunner HG, Kremer H, van der Burgt I, Crosby AH, Ion A, Jeffery S et al. . 2001. Mutations in PTPN11, encoding the protein tyrosine phosphatase SHP-2, cause Noonan syndrome. Nat Genet 29(4):465-468.

Tauber M, Ester W, Auriol F, Molinas C, Fauvel J, Caliebe J, Nugent T, Fryklund L, Ranke MB, Savage MO et al. . 2007. GH responsiveness in a large multinational cohort of SGA children with short stature (NESTEGG) is related to the exon 3 GHR polymorphism. Clin Endocrinol (Oxf) 67(3):457-461.

Thomas JT, Lin K, Nandedkar M, Camargo M, Cervenka J, and Luyten FP. 1996. A human chondrodysplasia due to a mutation in a TGF-beta superfamily member. Nat Genet 12(3):315-317.

Thomson JP, Skene PJ, Selfridge J, Clouaire T, Guy J, Webb S, Kerr AR, Deaton A, Andrews R, James KD et al. . 2010. CpG islands influence chromatin structure via the CpG-binding protein Cfp1. Nature 464(7291):1082-1086.

Toth G, Gaspari Z, and Jurka J. 2000. Microsatellites in different eukaryotic genomes: survey and analysis. Genome Res 10(7):967-981.

Troike KM, Henry BE, Jensen EA, Young JA, List EO, Kopchick JJ, and Berryman DE. 2017. Impact of Growth Hormone on Regulation of Adipose Tissue. Compr Physiol 7(3):819-840.

Tyagi M, Imam N, Verma K, and Patel AK. 2016. Chromatin remodelers: We are the drivers!! Nucleus 7(4):388-404.

Tycko B. 2010. Allele-specific DNA methylation: beyond imprinting. Hum Mol Genet 19(R2):R210-220.

Udy GB, Towers RP, Snell RG, Wilkins RJ, Park SH, Ram PA, Waxman DJ, and Davey HW. 1997. Requirement of STAT5b for sexual dimorphism of body growth rates and liver gene expression. Proc Natl Acad Sci U S A 94(14):7239-7244.

Vakili H, Jin Y, Nagy JI, and Cattini PA. 2011. Transgenic mice expressing the human growth hormone gene provide a model system to study human growth hormone synthesis and secretion in non-tumor-derived pituitary cells: differential effects of dexamethasone and thyroid hormone. Mol Cell Endocrinol 345(1-2):48-57.

Van Cauter E, Latta F, Nedeltcheva A, Spiegel K, Leproult R, Vandenbril C, Weiss R, Mockel J, Legros JJ, and Copinschi G. 2004. Reciprocal interactions between the GH axis and sleep. Growth Horm IGF Res 14 Suppl A:S10-17.

Van Dyke AL, Cote ML, Wenzlaff AS, Abrams J, Land S, Iyer P, and Schwartz AG. 2009. Chromosome 5p Region SNPs Are Associated with Risk of NSCLC among Women. J Cancer Epidemiol 2009:242151.

van Kerkhof P, Alves dos Santos CM, Sachse M, Klumperman J, Bu G, and Strous GJ. 2001. Proteasome inhibitors block a late step in lysosomal transport of selected membrane but not soluble proteins. Mol Biol Cell 12(8):2556-2566.

Veldhuis JD. 2008. Aging and hormones of the hypothalamo-pituitary axis: gonadotropic axis in men and somatotropic axes in men and women. Ageing Res Rev 7(3):189-208.

Veldhuis JD, Roemmich JN, Richmond EJ, Rogol AD, Lovejoy JC, Sheffield-Moore M, Mauras N, and Bowers CY. 2005. Endocrine control of body composition in infancy, childhood, and puberty. Endocr Rev 26(1):114-146.

Vidal OM, Merino R, Rico-Bautista E, Fernandez-Perez L, Chia DJ, Woelfle J, Ono M, Lenhard B, Norstedt G, Rotwein P et al. . 2007. In vivo transcript profiling and phylogenetic analysis identifies suppressor of cytokine signaling 2 as a direct signal

transducer and activator of transcription 5b target in liver. Mol Endocrinol 21(1):293-311.

Volpi S, Rabadan-Diehl C, Cawley N, and Aguilera G. 2002. Transcriptional regulation of the pituitary vasopressin V1b receptor involves a GAGA-binding protein. J Biol Chem 277(31):27829-27838.

Wagner K, Hemminki K, Grzybowska E, Bermejo JL, Butkiewicz D, Pamula J, Pekala W, and Forsti A. 2006. Polymorphisms in the growth hormone receptor: a case-control study in breast cancer. Int J Cancer 118(11):2903-2906.

Wan G, Mahajan A, Lidke D, and Rajput A. 2015. Bcl-2 together with PI3K p110alpha regulates cell morphology and cell migration. Cell Death Dis 6:e2006.

Wang AH, Quigley GJ, Kolpak FJ, Crawford JL, van Boom JH, van der Marel G, and Rich A. 1979. Molecular structure of a left-handed double helical DNA fragment at atomic resolution. Nature 282(5740):680-686.

Wang B, Ren J, Ooi LL, Chong SS, and Lee CG. 2005. Dinucleotide repeats negatively modulate the promoter activity of Cyr61 and is unstable in hepatocellular carcinoma patients. Oncogene 24(24):3999-4008.

Wang W, and Tao YX. 2016. Ghrelin Receptor Mutations and Human Obesity. Prog Mol Biol Transl Sci 140:131-150.

Wang YD, and Wood WI. 1995. Amino acids of the human growth hormone receptor that are required for proliferation and Jak-STAT signaling. Mol Endocrinol 9(3):303-311.

Wassenaar MJ, Dekkers OM, Pereira AM, Wit JM, Smit JW, Biermasz NR, and Romijn JA. 2009. Impact of the exon 3-deleted growth hormone (GH) receptor polymorphism on baseline height and the growth response to recombinant human GH therapy in GH-deficient (GHD) and non-GHD children with short stature: a systematic review and meta-analysis. J Clin Endocrinol Metab 94(10):3721-3730.

Waters MJ, and Brooks AJ. 2015. JAK2 activation by growth hormone and other cytokines. Biochem J 466(1):1-11.

Weedon MN, Lettre G, Freathy RM, Lindgren CM, Voight BF, Perry JR, Elliott KS, Hackett R, Guiducci C, Shields B et al. . 2007. A common variant of HMGA2 is associated with adult and childhood height in the general population. Nat Genet 39(10):1245-1250.

Wei Y, Puzhko S, Wabitsch M, and Goodyer CG. 2009a. Structure and activity of the human growth hormone receptor (hGHR) gene V2 promoter. Mol Endocrinol 23(3):360-372.

Wei Y, Puzhko S, Wabitsch M, and Goodyer CG. 2009b. Transcriptional regulation of the human growth hormone receptor (hGHR) gene V2 promoter by transcriptional activators and repressor. Mol Endocrinol 23(3):373-387.

Wei Y, Rhani Z, and Goodyer CG. 2006. Characterization of growth hormone receptor messenger ribonucleic acid variants in human adipocytes. J Clin Endocrinol Metab 91(5):1901-1908.

Weiss-Messer E, Merom O, Adi A, Karry R, Bidosee M, Ber R, Kaploun A, Stein A, and Barkey RJ. 2004. Growth hormone (GH) receptors in prostate cancer: gene expression in human tissues and cell lines and characterization, GH signaling and androgen receptor regulation in LNCaP cells. Mol Cell Endocrinol 220(1-2):109-123.

Wells RD. 2007. Non-B DNA conformations, mutagenesis and disease. Trends Biochem Sci 32(6):271-278.

Werther GA, Haynes K, Edmonson S, Oakes S, Buchanan CJ, Herington AC, and Waters MJ. 1993. Identification of growth hormone receptors on human growth plate chondrocytes. Acta Paediatr Suppl 82 Suppl 391:50-53.

Wickelgren RB, Landin KL, Ohlsson C, and Carlsson LM. 1995. Expression of exon 3-retaining and exon 3-excluding isoforms of the human growth hormone-receptor is regulated in an interindividual, rather than a tissue-specific, manner. J Clin Endocrinol Metab 80(7):2154-2157.

Wit JM, Clayton PE, Rogol AD, Savage MO, Saenger PH, and Cohen P. 2008. Idiopathic short stature: definition, epidemiology, and diagnostic evaluation. Growth Horm IGF Res 18(2):89-110.

Wit JM, Oostdijk W, Losekoot M, van Duyvenvoorde HA, Ruivenkamp CA, and Kant SG. 2016. MECHANISMS IN ENDOCRINOLOGY: Novel genetic causes of short stature. Eur J Endocrinol 174(4):R145-173.

Wit JM, Rekers-Mombarg LT, Cutler GB, Crowe B, Beck TJ, Roberts K, Gill A, Chaussain JL, Frisch H, Yturriaga R et al. . 2005. Growth hormone (GH) treatment to final height in children with idiopathic short stature: evidence for a dose effect. J Pediatr 146(1):45-53.

Wittig B, Dorbic T, and Rich A. 1989. The level of Z-DNA in metabolically active, permeabilized mammalian cell nuclei is regulated by torsional strain. J Cell Biol 108(3):755-764.

Wittig B, Dorbic T, and Rich A. 1991. Transcription is associated with Z-DNA formation in metabolically active permeabilized mammalian cell nuclei. Proc Natl Acad Sci U S A 88(6):2259-2263.

Wittig B, Wolfl S, Dorbic T, Vahrson W, and Rich A. 1992. Transcription of human c-myc in permeabilized nuclei is associated with formation of Z-DNA in three discrete regions of the gene. EMBO J 11(12):4653-4663.

Woelfle J, and Rotwein P. 2004. In vivo regulation of growth hormone-stimulated gene transcription by STAT5b. Am J Physiol Endocrinol Metab 286(3):E393-401.

Wojcik M, Krawczynska A, Antushevich H, and Herman AP. 2018. Post-Receptor Inhibitors of the GHR-JAK2-STAT Pathway in the Growth Hormone Signal Transduction. Int J Mol Sci 19(7).

Wolfl S, Wittig B, and Rich A. 1995. Identification of transcriptionally induced Z-DNA segments in the human c-myc gene. Biochim Biophys Acta 1264(3):294-302.

Wolinski K, Stangierski A, Dyrda K, Nowicka K, Pelka M, Iqbal A, Car A, Lazizi M, Bednarek N, Czarnywojtek A et al. . 2017. Risk of malignant neoplasms in acromegaly: a case-control study. J Endocrinol Invest 40(3):319-322.

Wolpert L. 2010. Arms and the man: the problem of symmetric growth. PLoS Biol 8(9).

Wood AR, Esko T, Yang J, Vedantam S, Pers TH, Gustafsson S, Chu AY, Estrada K, Luan J, Kutalik Z et al. . 2014. Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet 46(11):1173-1186.

Wu X, Liu F, Yao X, Li W, and Chen C. 2007. Growth hormone receptor expression is up-regulated during tumorigenesis of human colorectal cancer. J Surg Res 143(2):294-299.

Wyse BD, Linas SL, and Thekkumkara TJ. 2000. Functional role of a novel cis-acting element (GAGA box) in human type-1 angiotensin II receptor gene transcription. J Mol Endocrinol 25(1):97-108.

Yakar S, and Isaksson O. 2016. Regulation of skeletal growth and mineral acquisition by the GH/IGF-1 axis: Lessons from mouse models. Growth Horm IGF Res 28:26-42.

Yakar S, Liu JL, Stannard B, Butler A, Accili D, Sauer B, and LeRoith D. 1999. Normal growth and development in the absence of hepatic insulin-like growth factor I. Proc Natl Acad Sci U S A 96(13):7324-7329.

Yakar S, Werner H, and Rosen CJ. 2018. Insulin-like growth factors: actions on the skeleton. J Mol Endocrinol 61(1):T115-T137.

Yamada N, Yamaya M, Okinaga S, Nakayama K, Sekizawa K, Shibahara S, and Sasaki H. 2000. Microsatellite polymorphism in the heme oxygenase-1 gene promoter is associated with susceptibility to emphysema. Am J Hum Genet 66(1):187-195.

Yang N, Wang X, Jiang J, and Frank SJ. 2007. Role of the growth hormone (GH) receptor transmembrane domain in receptor predimerization and GH-induced activation. Mol Endocrinol 21(7):1642-1655.

Yu JH, Schwartzbauer G, Kazlman A, and Menon RK. 1999. Role of the Sp family of transcription factors in the ontogeny of growth hormone receptor gene expression. J Biol Chem 274(48):34327-34336.

Zhang JG, Farley A, Nicholson SE, Willson TA, Zugaro LM, Simpson RJ, Moritz RL, Cary D, Richardson R, Hausmann G et al. . 1999. The conserved SOCS box motif in suppressors of cytokine signaling binds to elongins B and C and may couple bound proteins to proteasomal degradation. Proc Natl Acad Sci U S A 96(5):2071-2076.

Zhou X, Benson KF, Ashar HR, and Chada K. 1995. Mutation responsible for the mouse pygmy phenotype in the developmentally regulated factor HMGI-C. Nature 376(6543):771-774.

Zhu J, He F, Hu S, and Yu J. 2008. On the nature of human housekeeping genes. Trends Genet 24(10):481-484.

Zhu T, Goh EL, Graichen R, Ling L, and Lobie PE. 2001. Signal transduction via the growth hormone receptor. Cell Signal 13(9):599-616.

Zigman JM, Nakano Y, Coppari R, Balthasar N, Marcus JN, Lee CE, Jones JE, Deysher AE, Waxman AR, White RD et al. . 2005. Mice lacking ghrelin receptors resist the development of diet-induced obesity. J Clin Invest 115(12):3564-3572.

Zizola CF, Balana ME, Sandoval M, and Calvo JC. 2002. Changes in IGF-I receptor and IGF-I mRNA during differentiation of 3T3-L1 preadipocytes. Biochimie 84(10):975-980.

Zogopoulos G, Albrecht S, Pietsch T, Alpert L, von Schweinitz D, Lefebvre Y, and Goodyer CG. 1996a. Fetal- and tumor-specific regulation of growth hormone receptor messenger RNA expression in human liver. Cancer Res 56(13):2949-2953.

Zogopoulos G, Figueiredo R, Jenab A, Ali Z, Lefebvre Y, and Goodyer CG. 1996b. Expression of exon 3-retaining and -deleted human growth hormone receptor messenger ribonucleic acid isoforms during development. J Clin Endocrinol Metab 81(2):775-782.

Zvonic S, Ptitsyn AA, Kilroy G, Wu X, Conrad SA, Scott LK, Guilak F, Pelled G, Gazit D, and Gimble JM. 2007. Circadian oscillation of gene expression in murine calvarial bone. J Bone Miner Res 22(3):357-365.