Investigating the Relationship Between *LUNATIC FRINGE* Variants and Spondylocostal Dysostosis Type-III

by

Parker Cole Wengryn

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Science

Medical Sciences - Medical Genetics

University of Alberta

Abstract

The vertebral column's bilateral symmetry plays a crucial role in respiration, ambulation, and weight bearing. Scoliosis, a lateral curvature of the spine of >10°, disrupts this anatomy and can lead to pain, poor quality of life, and in severe cases, mortality. The prevalence of scoliosis is estimated to be as high as 1-in-33, and treatment ranges from physiotherapy and structural bracing to corrective surgery. In most cases, early diagnosis and proactive care are key to limiting curve progression and minimizing treatment invasiveness. However, the underlying cause of most scoliosis cases is unknown, and screening may not occur until the patient is symptomatic. The study of monogenic causes of scoliosis can lead to the identification of new diagnostic targets within the affected cellular pathways. The NOTCH signalling pathway is critical for somite (i.e. embryonic vertebral precursor) formation and is implicated in many diseases that cause scoliosis. Spondylocostal Dysostosis Type 3 (SCD3) is caused by pathogenic variants in the NOTCH pathway gene *Lunatic Fringe* (*LFNG*). LFNG is a Golgi-resident β-1,3-N-Acetylglycosaminyltransferase that glycosylates NOTCH receptors to regulate the size, shape, and symmetry of somites and thus vertebral anatomy. This monogenic condition is characterized by abnormal anatomy of the vertebrae and ribs which leads to short stature, various extents of scoliosis, and restrictive lung disease. Very few cases have been identified, and little is known about how each causative variant affects LFNG function; however, case-to-case phenotypic variability may hint at a genotype-phenotype relationship.

In hopes of determining whether there is a relationship between SCD3 phenotype and *LFNG* variant, the first goal of this thesis was to investigate the functional consequences of two novel *LFNG* variants (c.521G>A [p.R174H]; c.766G>A [p.G256S]) presenting *in trans* in a

proband with SCD phenotype with mild scoliosis, but not clinically defined short stature nor restrictive lung disease. I assessed both variants for impaired glycosyltransferase activity, subcellular mislocalization, and aberrant protein processing. Our results indicate that the p.G256S variant is enzymatically non-functional, while the p.R174H variant is functionally less effective. Both variants were not different from wildtype (WT) in localization and processing. Our findings suggest that the hypomorphic variant (p.R174H) may have partially improved the patient's stature, as evidenced by a lower arm span-to-height ratio, increased height, and more vertebrae. However, this variant did not appear to have any effect on the severity of vertebral malformations, including scoliosis. The lack of molecular characterization studies for *LFNG* variants of other SCD3 probands hinders the generalizability of this conclusion.

To address this lack of functional testing, nine *LFNG* missense variants associated with SCD3 published before 2023, c.564C>A, c.583T>C, c.842C>A, c.467T>G, c.856C>T, c.601G>A, c.446C>T, c.521G>A, and c.766G>A, were assessed *in vitro* for subcellular localization and protein processing. Glycosyltransferase activity was quantified for the first time in the c.583T>C, c.842C>A, and c.446C>T variants. Primarily, this work indicates that all variants that prevent Golgi localization also impair protein processing, suggesting the two methods evaluate LFNG trafficking from different perspectives. It appears that the FRINGE domain is responsible for aberrant trafficking. Secondly, our data suggests that variant proximity to the catalytic residue may influence whether LFNG is improperly trafficked and/or enzymatically dysfunctional. Finally, the phenotype of the axial skeleton, but not elsewhere, may be modulated in a variant-specific fashion, supporting our previous work. More reports are

needed to continue testing this hypothesis. We anticipate our data will be used as a basis for the discussion of genotype-phenotype correlations in SCD3.

In sum, this work led to the discovery of the first two disease-implicated *LFNG* alleles with partial enzyme activity, the first *LFNG* allele to partially traffic to the Golgi and contributed to determining the effect of nine *LFNG* variants on glycosyltransferase activity, processing, and localization. With these discoveries, this work has provided support for a loci-dependant pattern of LFNG functional perturbation, a clearer perspective of LFNG trafficking experiments, and associated the larger vertebral number of an SCD3 proband with the only hypomorphic *LFNG* allele known to date. The vertebral number of SCD3 probands is likely modulated in a variant-specific fashion, and thus there does appear to be a genotype-phenotype relationship in SCD3.

Preface

Chapters 1 and 4 of this thesis are original work by Parker Wengryn. The research projects, of which this thesis is a part, received research ethics approval and consent for the imaging and radiograph presentation in Chapter 2 from the Hospital Nacional de Niños (San Jose, Costa Rica). The remaining research in Chapter 3 did not require ethical approval as per TCPS-2 (2022) Article 12.21, the University of Alberta Research Ethics Board.

Chapter 2 of this thesis has been published as: Wengryn, P., Silveira, K., Oborn, C., Soltys, C.L., Beke, A., Chacon-Fonseca, I., Damseh, N., Rodriguez, M. Q., Badilla-Porras, R., & Kannu, P. (2023). Functional Characterization of Novel *Lunatic Fringe* Variants in Spondylocostal Dysostosis Type-III with Scoliosis. *Human Mutation*, *2023*, 1–12. https://doi.org/10.1155/2023/5989733. This work was part of an international collaboration and copyrighted under, "Creative Commons 4.0©". Dr. Badilla-Porras and Dr. Rodriguez initially identified the proband and conducted the radiographic analysis at the Hospital Nacional de Niños (San Jose, Costa Rica). Dr. Chacon-Fonseca and Dr. Damseh undertook genetic investigations and gathered clinical history under the supervision of Dr. Kannu at the Hospital for Sick Children (Toronto, Canada). At the University of Alberta (Edmonton, Canada), I aided Dr. Kannu in the general design of his original idea for beginning experimental analysis. I designed the technical parameters for each experiment and collected *in silico* and *in vitro* data with contributions from Dr. Da Costa Silviera and Mrs. Soltys. I interpreted the *in vitro* data and Mr. Oborn and Mr. Beke contributed to the interpretation of *in silico* data. Mr. Beke and I conducted the statistical analysis. I drafted the original manuscript with my original scientific insight and interpretations and Dr. Kannu contributed to idea generation, editing, and refinement. The rest of the authors

contributed to the editing stage. I was responsible for revision implementation with contributions from Dr. Kannu.

Chapter 3 of this thesis has been published as: Wengryn, P., Fenrich, F., Silveria, K., Oborn, C., Mizumoto, S., Beke, A., Soltys, C.L., Yamada, S., Kannu, P. (2024). Integrative Analysis of Lunatic Fringe Variants Associated with Spondylocostal Dysostosis Type-III. *The FASEB Journal*, *38*(13), 1–16. https://doi.org/10.1096/fj.202302651RR. This work was part of an international collaboration and copyrighted under, "Creative Commons 4.0©". Dr. Kannu and I designed the project at the University of Alberta. I performed cloning and subcloning with Mrs. Fenrich and western blotted with Mrs. Fenrich and Mrs. Soltys. Dr. Silveria and I jointly undertook the immunofluorescent assays where Dr. Silveria imaged the samples and I counted the cells and processed the images. Dr. Mizumoto and Dr. Yamada collected the functional analysis results and functional analysis western blots at the University of Meijo (Kyoto, Japan) with our plasmid constructs. This was predicated on initial experimental data collected by Mrs. Fenrich and I. Mr. Beke and Mr. Oborn compiled and analyzed the *in silico* protein/genomic data. I undertook statistical analysis with Mr. Beke, Dr. Mizumoto, and Dr. Yamada. I drafted the original manuscript with my original scientific insight and interpretations and Dr. Kannu contributed significantly to idea generation, editing, and refinement. The rest of the authors contributed to the editing stage. I was responsible for revisions and corrections with contributions from Dr. Kannu.

Dedication

This document is dedicated to my parents, Bradly and Marianne Wengryn, for their endless love and limitless support of my dreams.

Acknowledgements

I would like to sincerely thank my supervisor, Dr. Peter Kannu, for taking a chance on me with what was the opportunity of a lifetime. His encouragement was genuine, and his support comprehensive, providing me with more than what was necessary to succeed. His mentorship was characterized by kindness, empathy, and mutual respect. The lessons I learned from him are ones that I cherish deeply. I hope to one day embody these as a physician, scientist, and leader.

I would like to thank all of my lab members: Dr. Karina da Costa Silveria, Dr. Asghar (Mardin) Fallah, Dr. Guoju Hong, Dr. Matthea Sanderson, Carrie-Lynn Soltys, Saima Ghafoor, Liliana Vertel, Connor Oborn, Kazette Chan, Alex Beke, Felicity Fenrich, and the many others who contributed to the positive and enriching environment that is the Kannu Laboratory. Every one of you represents the best of science; extremely clever, a lot of fun, and of course, the best kind of weird.

I would specifically like to thank two of my lab members, Carrie-Lynn Soltys and Dr. Karina da Costa Silveria. Carrie, for her consistent support and guidance: she is the cornerstone of the laboratory and her support for me professionally and personally was unwavering (as well the support for my western blots as the reigning “Western Blot Wizard”). Karina, for her willingness to sacrifice for others and her kindness while doing so: regardless of how many experiments I’d like to start at 5:00 AM, or how many “crazy ideas” I’d present to her in a 30-minute window, she was ready to get started, coffee in hand, encouraging me to continue to explore.

I am very grateful to my committee members, Dr. Fred Berry and Dr. Toshifumi Yokota, who provided expert insights and commentary throughout my graduate program. This was invaluable and provided me with many perspectives that have enhanced my assessment of problems in the laboratory and the clinic. Their genuine commitment to my professional goals was vital to the direction I took in the laboratory and my pursuit of a career in Medicine.

The support that I received from Jennifer Beattie was truly exceptional. At any time, any day, I knew that I could reach out to her for any reason, and she would ensure that my problems were solved as quickly as they began. No matter the complication, she went over and above, saving me from administrative wrath many times over. Her kindness and commitment both professionally and personally cannot be overstated.

The love and support that I received from my immediate family were critical to remaining resilient during this project. To my parents, Brad and Marianne Wengryn; my sisters, Seanna Wengryn and Amanda Boychuk; and my brother-in-law, Sean Boychuk- your encouragement and belief in me was instrumental to my success. After numerous iterations of a failed experiment, you all provided me with the foundation I needed to get back up and try again. You all provided so much more than I could ever write here, and I sincerely thank you.

Finally, I would like to thank Christopher Cooper, Brett Frederick, Tanner Kamdar, and Matthew Jackson. This exceptional group of men was there for me every step of the way, through the highest and lowest moments in my life, with big smiles and bigger hearts. The innumerable days and nights we've spent together mean more than I can express, and I hope to continue our journey together long into the future. I am blessed to call you my closest friends. Down the middle fellas, you're better than average.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1. Anatomy and Significance of the Spine and Ribs

The mammalian trunk is characterized by a repeating, regionalized set of vertebrae which anchor the ribs and form the spondylocostal axis (Moore et al., 2014). Walking, breathing, load bearing, and protection of the spinal cord are some of the many evolutionary advantages conferred by the thorax. The human vertebral column is composed of 33 vertebrae (on average), subcategorized from superior to inferior as cervical (7), thoracic (12), lumbar (5), sacral (5), and coccygeal (4) (Figure 1A). The pre-sacral vertebrae share extensive anatomical features (Figure 1B); the anterior-most portion is called the vertebral body and lies stacked between superior and inferior intervertebral cartilaginous discs. Forces are distributed by the bony vertebral bodies and absorbed by the disks, allowing for significant compressive strength. The posterior-most aspect of the vertebrae is known as the spinal arch and includes seven appendages: one posteriorly extending spinous process, two posterolaterally extending transverse processes, and four articular processes. These structures align the spine and ribs through inter-vertebral ligaments, tendons, and facet joints. Finally, within the posterior aspect of the vertebral body proximal to the spinal arch lies the vertebral foramen, a circumferential cavity which spans the vertebral column and houses the spinal cord. Extending bilaterally from the foramen are intervertebral foramina, small bore cavitations which allow for communication between the central and peripheral nervous systems. In contrast to other spinal regions, nerves which exit the thoracic foramina remain independent (i.e. do not form plexuses) and only innervate structures along their respective dermatomes. This is largely due to the anatomical structure of the ribs and intercostal regions.

Humans possess 12 bilaterally symmetrical sets of ribs which originate from thoracic vertebrae and wrap the presumptive thoracic cavity in an anteroinferior direction (Moore et al.,

2014; Figure 1C). The 7 superior-most rib sets are known as ‘true ribs’ and circumferentially connect a vertebra to the sternum through medial cartilaginous costochondral joints. The remaining ribs, also known as ‘false’ ribs, are either connected to the sternum through a superior costochondral joint (ribs 8 through 10) or extend half (rib 10) or less than a quarter (rib 11) of the thoracic circumference. Intercostal muscles lie between rib sets and house the main intercostal veins, arteries, and nerves, behind the inferior-most portion of the superior rib (the intercostal groove) along with collateral branches of all three just superior to the inferior rib. The intercostals therefore provide the space necessary for nervous and circulatory supply to the superficial aspect of the thorax.

The unique anatomical shape, flexibility, and neuromuscular functions of the spondylocostal axis contribute to a thorax which can alter its volume through expansion and contraction (Moore et al., 2014). This, along with diaphragmatic contraction and relaxation, is the basis of respiration. Therefore, structural abnormalities of the ribs and spine can restrict the respiratory capacity of the thorax. This is one of the main underlying etiologies of restrictive lung disease (RLD), an often irreversible pulmonary condition which leads to reduced quality of life and ultimately, mortality.

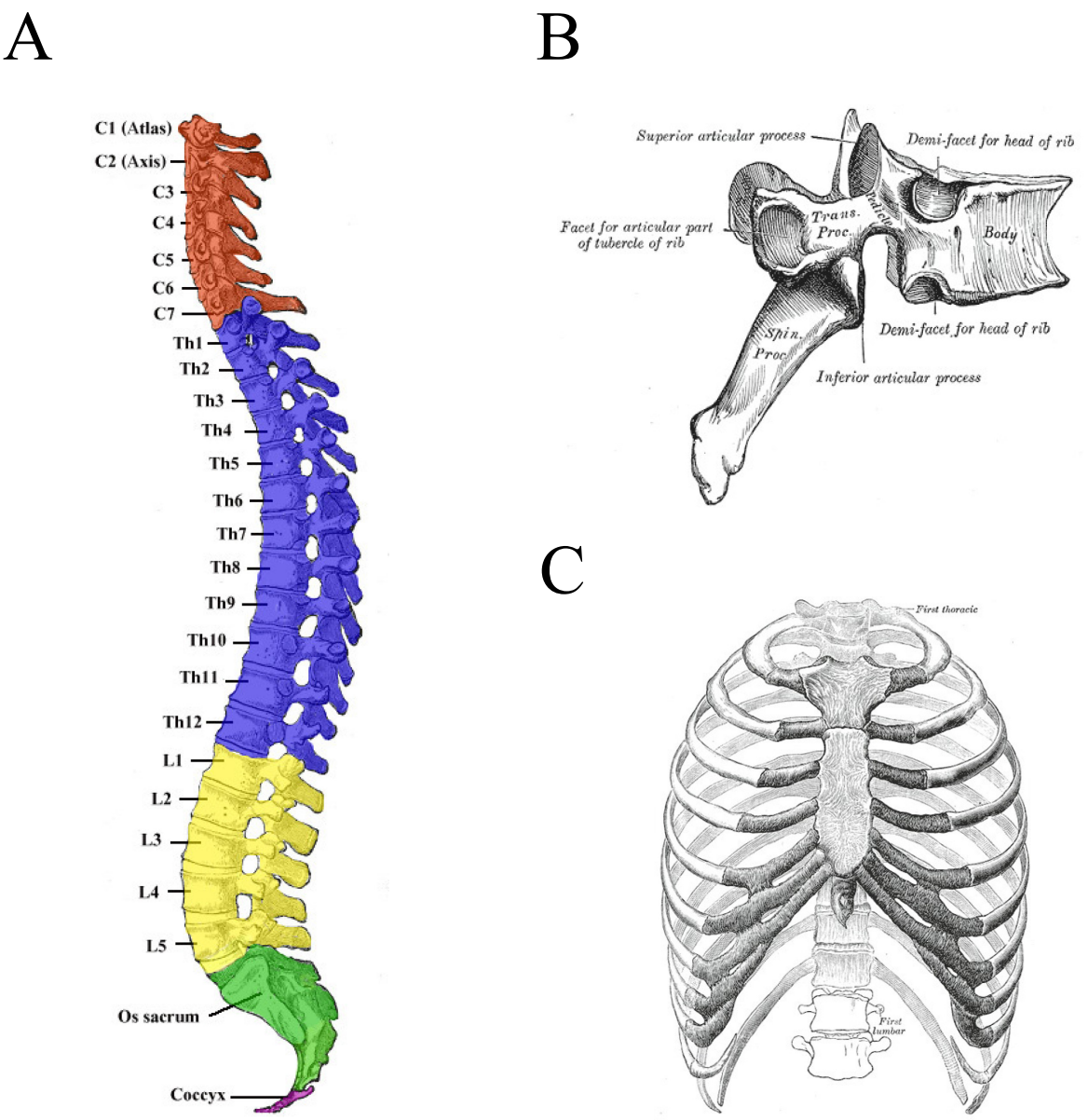


Figure 1: Anatomical Diagrams of Thoracic Structures. A) "Vertebral Column. Schematic medial view of the entire vertebral column showing its 5 regions and their curvatures. Contributed by Wikimedia Commons, Gray 111 (Public Domain)" (Waxenbaum et al., 2024). **B)** "The Thoracic Vertebrae Henry Vandyke Carter, Public Domain, via Wikimedia Commons" (Waxenbaum et al., 2024). **C)** "The Thorax; Anterior View, Henry Vandyke Carter, Public Domain, via Wikimedia Commons" (Hussain & Burns, 2024).

1.2. Scoliosis

1.2.1. Definition, Etiologies, and Incidence

The vertebral column of an anatomically normal human has 20° to 40° of cervical kyphosis (anterior curve), 40° to 60° of lumbar lordosis (posterior tilt), and <10° lateral curvature (Liebsch & Wilkie, 2018; Figure 1A). These features provide maximal flexibility, rotation, and thoracic expansion while maintaining stability and compressive strength. Deviations from these values can compromise any or all of the above factors. The pathological lateral curvature of the spine is known as scoliosis and is defined as a Cobb angle (lateral curvature) of greater than 10°. This life-altering condition can cause chronic pain, immobility, and RLD (the latter of which is a cause of mortality in rare cases) (Hawes & Weinstein, 2003). Psychologically, this cohort reports an increased prevalence of depression, negative body image, and mood disorders (Chang et al., 2016; Lin et al., 2019; Gallant et al., 2018).

The onset of scoliosis varies from infantile to geriatric and progression is case-dependent. This is largely due to the many etiologies of scoliosis, which are often (although not always) categorized as congenital, neuromuscular, degenerative, and idiopathic (Sung et al., 2021; Figure 2). Congenital scoliosis (CS) accounts for approximately 10% of diagnoses (Sebaaly et al., 2022) and is caused by segmentation defects of vertebrae (SDV's). CS can be isolated due to pathogenic variants in genes associated with vertebrae formation alone (e.g. spondylocostal dysostosis) or as the secondary result of an underlying genetic syndrome (e.g. Ehlers-Danlos Syndrome, Marfan Syndrome, VACTRL, etc.) (Giampietro, 2012; Kikanloo et al., 2019; Sebaaly et al., 2022). Here, syndrome refers to a constellation of symptoms across multiple organ systems which are often due to a genetic change (see section 1.4.5). In both cases, one or many vertebrae

form abnormally, causing the alignment of the spine to change as a result. Neuromuscular scoliosis (NS) is caused by asymmetric, degenerated, or otherwise damaged spinal muscle, leading to the lateral pull of a vertebral section (Wishart & Kivlehan, 2021). Although the incidence rate is unknown, 90% of patients with Duchenne Muscular Dystrophy (DMD), 70% of patients with Spinal Muscular Atrophy (SMA), and 20% to 60% of patients with Cerebral Palsy live with scoliosis. Degenerative scoliosis (DS) often occurs in adulthood and is due to damage to the facet joints and/or intervertebral disks (Kelly et al., 2020). Damage can occur due to trauma, chronic irritation, and/or pressure (obesity being the most common cause), and often results in exasperation of primary scoliosis. The true incidence rate is unknown due to the lack of a uniformly applied definition, but it is speculated to be the inciting factor for 30% to 60% of scoliosis cases in elderly populations (Kelly et al., 2020). Finally, idiopathic scoliosis is by far the most common etiology estimated at 80% to 90% of cases (Asher & Burton, 2006). Adolescent Idiopathic Scoliosis (AIS) is often reported to affect between 1% to 3% of children worldwide (Menger & Sin, 2023). Without grossly observable structural, muscular, or degenerative abnormalities, there has been significant investment into the study of AIS etiology surrounding potential genetic, epigenetic, endocrinological, and environmental contributions (Bagnall, 2008; Kikanloo et al., 2019). Unfortunately, many of the contributing factors remain speculatory or unknown.

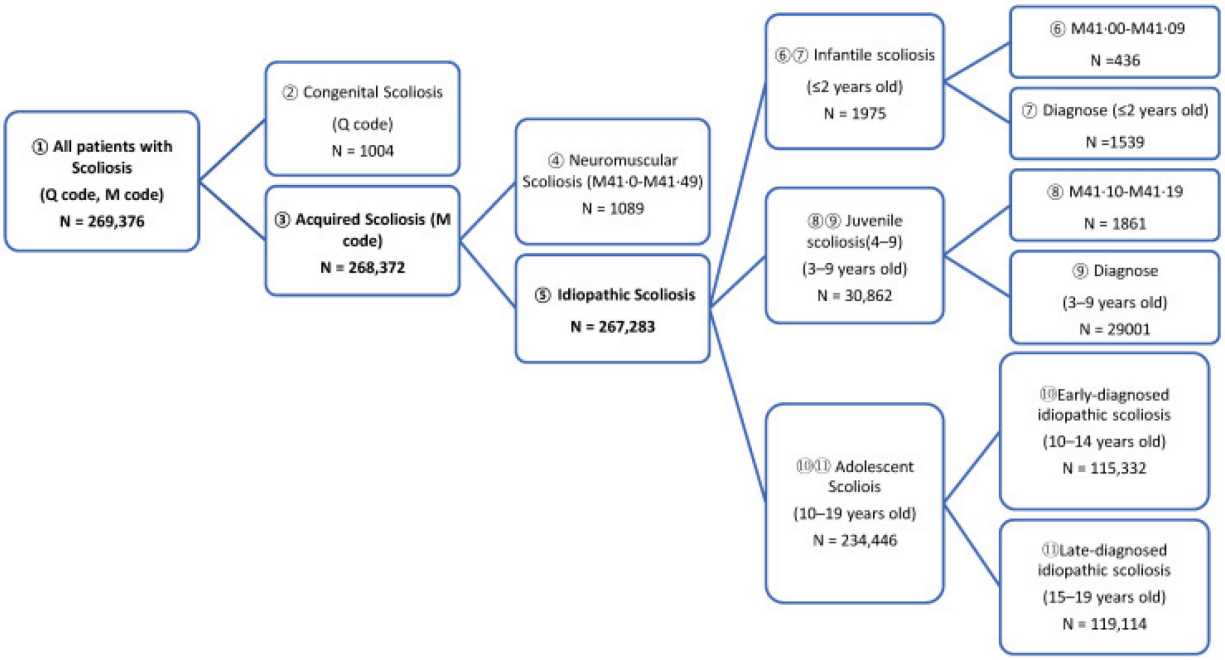


Figure 2: Categorization scheme for scoliosis. "The diagram shows the nationwide scoliosis investigation from The Health Insurance Review and Assessment Service data analysis" (Sung et al., 2021).

1.2.2. Prognosis and Management

The goal of scoliosis management is to reduce or maintain a Cobb angle of <40° at skeletal maturity (Asher & Burton, 2006; Weinstein, 1986). Lifelong care with significant effort from a multidisciplinary healthcare team is required to achieve this goal. All patients undergo longitudinal curve monitoring, batteries of genetic, physical, radiographic, and biochemical investigations, and rehabilitation/physical therapy (Asher & Burton, 2006; Kikanloo et al., 2019). This is the standard of care in 90% of AIS cases where the Cobb angle is less than 20° as the progression risk is less than 40%. However, it has been suggested that a Cobb angle of ≥25° predicts a significant risk of curve exasperation (Di Felice et al., 2018; Tan et al., 2009) and thus a more invasive treatment pathway, demonstrating the importance of early diagnosis.

Although outcomes are generally positive for less invasive strategies (Weiss et al., 2022), bracing and corrective surgery are necessary in the remaining 10% of cases, approximately 0.3% of people worldwide. Bracing is indicated for curves between 20° and 50° with a progression risk between 40% to 60% and requires the patient to wear a device which physically straightens (or prevents further progression of) the scoliotic curve (Babaee et al., 2023; Rowe et al., 1997). Optimal outcomes require months to years of wear and compliance is therefore challenging. Although efficacy data is mixed, the average reduction in Cobb angle is approximately 13° (Babaee et al., 2023; Rowe et al., 1997; Tan et al., 2009).

Surgical correction is the most invasive intervention and is indicated for patients with Cobb angles >50°. A variety of surgeries exist, with most employing metallic rods fastened to the vertebrae to straighten the curve. A recent meta-analysis reported an average curve correction of

68.4% (Babaee et al., 2023), demonstrating the efficacy of this procedure. However, intensive follow-up with significant intra- and post-operative risks renders treatment challenging.

Both interventions also have significant financial barriers, with a brace costing between $5000 (USD) and $10,000 (USD) and the average surgical cost of just over $126,000 (USD) (Bozzio et al., 2019). In 2011, this totalled $517 million (USD) in AIS surgeries in the United States alone. Although outcomes are highly variable and severity-dependent, the previous examples help to highlight that the ideal treatment for scoliosis is prevention or at least mitigation of curve exasperation. However, a prerequisite for primary intervention is a deep understanding of scoliosis etiology, and as of 2024, a vast majority of scoliosis cases remain idiopathic.

1.2.3. Current Work in AIS Etiology

Decades of research have indicated that AIS is strongly heritable, and therefore the AIS research paradigm has shifted towards evaluating both the combinatorial and isolated roles of genetic variants (Giampietro, 2012; Man et al., 2019; Terhune et al., 2022). Although genome-wide association studies have had limited successes (Man et al., 2019), the largest whole-exome sequencing study to date of 23 families suggested that AIS is largely a polygenic disease (Terhune et al., 2021). Many of the 123 genes they identified contributed to extracellular matrix formation (e.g. nine *COLLAGEN* (*COL*)-family genes) and cytoskeleton formation (e.g. *POC5 Centriolar Protein* (*POC5*) and *Kinesin Family Member 6* (*KIF6*)). However, genes involved in embryonic development of the spine (e.g. *Delta-like Ligand 3* (*DLL3*), *Ephrin A5* (*EphA5*); discussed in 1.3.3) also appeared to contribute. It was hypothesized that the accumulation of "mild" variants (i.e. those with only minor functional disturbances of the resulting protein) in

many of the gene families could be responsible for some AIS cases, although the small sample size prevented definitive, generalizable conclusions from being drawn.

The specific pathways which contribute to vertebral column anatomy have been screened in hopes of identifying contributors to AIS (De Salvatore et al., 2022; Wise, 2015). Most generally, these genes can be categorized as primary (directly form the vertebral column (see section 1.3) and secondary (influence structure once the spine is formed). An example of a primary contributor is the *T-Box Transcription Factor 6* (*TBX6*) gene which creates the boundaries of vertebral precursors, called somites, during an embryological process known as somitogenesis (see section 1.3.4) (Takeda et al., 2018). Pathogenic variants of *TBX6*, or one of five other genes involved in somitogenesis, including *Lunatic Fringe* (*LFNG*), lead to malformed vertebrae and ribs and thus congenital scoliosis in a rare disease called Spondylocostal Dysostosis (SCD) (as this is the central focus of this thesis, detailed discussion can be found in section 1.3 and beyond). An example of both a primary and secondary contributor is *Ladybird Homeobox 1* (*LBX1*), a gene which is critical to paravertebral muscle development during somitogenesis through its influence on Wingless and Int-1 (WNT) signalling (WNT discussed throughout 1.3; Kikanloo et al., 2019). Pathogenic variants not only lead to asymmetrical vertebral musculature and primary curvature, but also asymmetrical gait which can secondarily progress scoliosis after the spine has formed (Kikanloo et al., 2019). Finally, a significant secondary structural influence is the 16p11.2 chromosomal microduplication as this region contains the *SH2B Adaptor Protein 1* (*SH2B1*), a gene which decreases satiety through leptin signalling (Sadler et al., 2019). Individuals with this microduplication rarely feel satiated, leading to obesity and a 1.5-fold risk of scoliosis compared to the general population (Bochukova et al., 2010).

In sum, examining the pathways which contribute to forming or maintaining spinal structure has been effective in identifying causes of scoliosis. Furthermore, it appears likely that AIS can be of polygenic etiology due to a combination of genetic variants across common processes (Giampietro, 2012; Terhune et al., 2021). However, genes like *TBX6* and *LFNG* are directly responsible for vertebral anatomy and can lead to prominent anatomical changes such as scoliosis (Sparrow et al., 2006; Takeda et al., 2018). Investigating these genes and the pathways they contribute to are critical to understanding the extent to which genotype, vertebral anatomy, and scoliosis could be linked.

1.3. Somitogenesis Controls Axial Skeletal Development

Segmentation is the developmental process which utilizes tissue repetition to confer pattern, sequence, and symmetry to the body plan. Such arrangements directly and indirectly provide significant morphological and ambulatory diversity amongst the Annelida, Chordata, and Arthropoda phyla (Finnerty, 2003; Palmeirim et al., 2018). Illustrating the evolutionary utility of segmentation are tissues and organ systems such as the hindbrain, limbs, wings, pharynx, and vertebrae. In vertebrata, segmentation of the primordial trunk is separated into primary and secondary body formation, anatomically demarcated by the lumbosacral junction (Palmeirim et al., 2018). Although the mechanism of segmentation is controlled by the same processes, there are subtle yet significant differences in their molecular control. Primary body formation will be the focus of subsequent somitogenesis discussion due to its relevance to the formation of the thoracic cage and scoliosis.

Segmentation begins shortly after gastrulation when the embryo begins its rostrocaudal elongation (Musumeci et al., 2015). As growth proceeds, precursor structures to the axial skeleton begin to form. These structures, called somites, are spheres of mesenchymal cells wrapped in a thin layer of epithelium (Figure 3A). Somites sequentially bud in bilaterally symmetrical pairs from the paraxial mesoderm adjacent to the notochord on the dorsal aspect of the embryo. The process which governs the sequential, rhythmic, and symmetrical properties of somitic budding is known as somitogenesis.

1.3.1. The Clock and Wavefront Model: General Overview

Vertebrate somitogenesis occurs through a synchronized multiomic orchestra described by, "The Clock and Wavefront Model" (Cooker & Zeeman, 1976; Figure 3). The following is a brief overview of the Wavefront (1) and the Clock (2) based on the reviews of (Clark, 2021; Miao & Pourquié, 2024; Musumeci et al., 2015; S. Williams et al., 2019):

1) Presomitic mesodermal (PSM) cells create longitudinally opposed, pan-embryonic signalling gradient 'Wavefronts' (Figure 3B). The caudorostral gradient is characterized by Retinoic Acid (RA) signalling whereas Fibroblast Growth Factor (FGF) and WNT signalling characterize the rostrocaudal gradient. RA signalling promotes differentiation and bilateral symmetry whereas FGF and WNT help maintain multipotency. Critically, between both gradients exists an area with minimal FGF, WNT, and RA signalling and is known as the "Determination Front" (DF). The oppositional nature of these gradients contributes to the fine-tuning of DF width. Due to minimal wavefront signalling in the DF, the transcriptomic profile of this group of cells can be dictated by a molecular oscillator known as the 'Segmentation Clock'.

2) The Segmentation Clock describes a group of over 250 genes which synchronously oscillate at a species-specific tempo (e.g. humans, 5-6 hours) throughout the presomitic mesoderm (Figure 3A). Cells acted upon by the segmentation clock are directed to form somites. Expression begins at the posterior tail-bud and propagates anteriorly in a wave-like fashion until it reaches the DF. After a set clock-cycle interval, expression terminates in the tail bud, propagating the clock's conclusion throughout the PSM. The synchronicity, timing, and even propagation speed of the segmentation clock are ultimately controlled by the NOTCH pathway (see section 1.3.3).

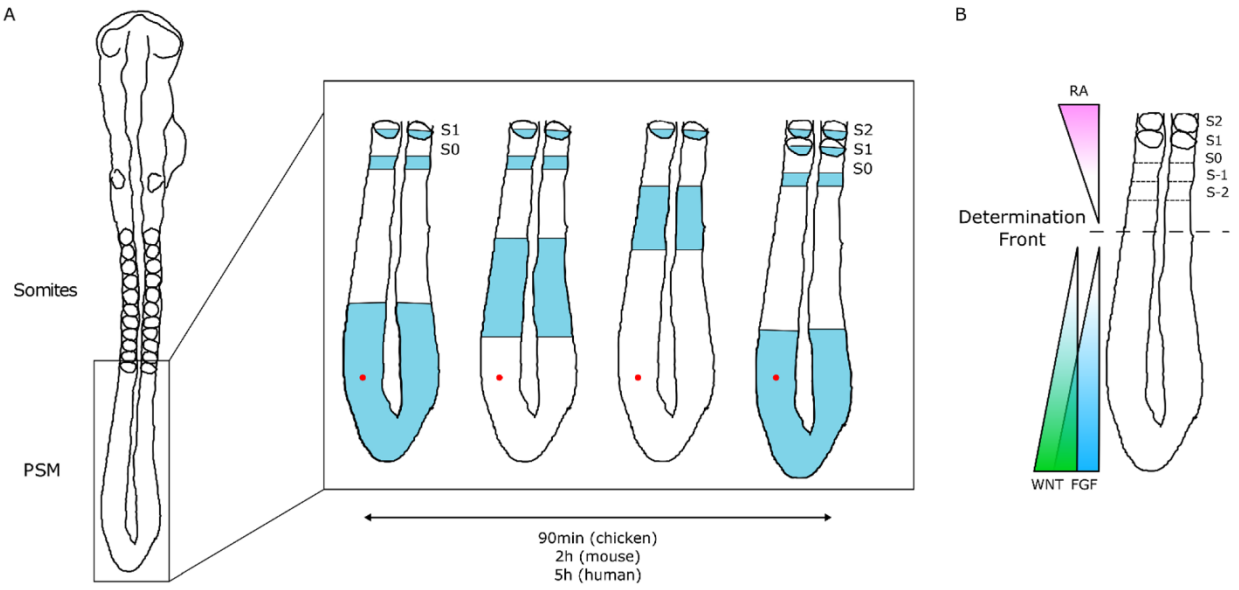


Figure 3: Diagram of the Clock and Wavefront Mechanism During Somitogenesis. "**A)** Representation of one oscillation of hairy1 gene expression (blue) in the presomitic mesoderm (PSM) of a 48h chicken embryo. A cell in the PSM (red dot) undergoes a complete cycle of gene expression activation-repression-activation as a new somite is formed in the anterior-most PSM. **B)** Opposing gradients of retinoic acid (RA) and WNT/FGF signalling converge at the determination front; rostrally, the PSM tissue is already committed to form somites." (Nóbrega et al., 2021).

Combined, these postulates suggest that as the embryo elongates rostrocaudally, the segmentation clock imposes somitic fate upon adjacent and sequential groups of PSM within a steadily advancing DF. This leads to the conclusion that the Clock and Wavefront act in tandem to set mathematical parameters which demarcate anatomical features of subsequent developmental processes. Specifically, clock interval, axial growth rate, and DF width set the size, structure, and number of somites; in turn, the structure and number of vertebrae and ribs, the stability and length of the spine, and thus the final anatomy of the thorax and its and physiological fitness (or lack thereof). Therefore, patterns in shape, size, and number of each structure can point to the stage at which development was disturbed. These clues, which often connect many levels of development, are best highlighted in historical research of animal models of somitogenesis and contemporary *in vitro* development models and will therefore be the focus of the next section.

1.3.2. Clock and Wavefront Model: A (Very) Brief History of Patterns

Historical analyses of the Clock and Wavefront Model assessed gross changes in the (supposed) somitogenic parameters of anatomical structure. Mice, chick, and zebrafish somitogenesis models were therefore commonplace, providing links between each developmental stage (Gomez & Pourquie, 2009; Pourquié, 2022). To observationally evaluate the Segmentation Clock postulate, somitic size and number were quantified in embryos of species that produce different numbers of vertebrae (zebrafish, 33; chick, 46; mouse, 50) (Gomez et al., 2008; Tam, 1981). When comparing the speed of somite creation (a proxy for clock speed) and the number that were created when correcting for DF width and elongation speed, species with faster oscillation rates often had more numerous and narrower somites, leading to more

vertebrae (Gomez et al., 2008; Gomez & Pourquie, 2009; Tam, 1981). The first experimental approach to directly support the existence of a segmentation clock utilized chick embryos that were cultured *ex vivo* and dissected along the axial plane (Palmeirim et al., 1997). One half was fixed, but both were exposed to *hairy1* (a clock-regulating gene) and then incubated for various time points. After 90 minutes, the *hairy1*-exposed half had one more somite than the control side. Further support for their suspicions were derived from the cyclic expression pattern of *hairy1* and *hairy2* mRNA as detected by a newly developed, more sensitive *in situ* hybridization technique. Although this research was based on many years of correlational evidence, this was the first time a segmentation clock was detected through experimental means.

To test the Wavefront postulate, signalling pathways associated with DF size were either amplified or diminished (Aulehla et al., 2008; Dubrulle et al., 2001; Sawada et al., 2001; Vermot et al., 2005; Wahl et al., 2007). One set of experiments increased DF width by incubating embryos (mouse and chick) with FGF-soaked beads. Indeed, as predicted by the Clock and Wavefront, somite size was reduced. The converse effect was also true as embryos with genomic deletion or chemical inhibition of FGF signalling had larger somites compared to the control. Critically, these early experiments hinted at various etiologies for axial skeleton disease and were foundational in proceeding research (Bochter et al., 2022; McIntyre et al., 2020; Nóbrega et al., 2021; Shifley & Cole, 2008; D. R. Williams et al., 2014, 2016).

Contemporary somitogenesis models utilize human and mouse stem cells of various origin to test both the Segmentation Clock and Wavefront *in vitro*. The first-generation models utilized 2D culture and visualized wave-like oscillations of a clock reporter gene called *Hairy and Enhancer of Split 7* (*HES7*) (Chu et al., 2019; Diaz-Cuadros et al., 2020; see sections 1.3.3

and 1.3.4). This approach also harnessed RNASeq, elucidating not only the developmental trajectory of cells undergoing somitogenesis, but also the hundreds of genes and their pathways responsive to (or responsible for) segmentation clock oscillations. The second-generation *in vitro* models employed 3D culture and were the first to create somite-like structures (somitoids) (Miao et al., 2023; Sanaki-Matsumiya & Ebisuya, 2022). By combining the same *HES7* fluorescent reporter system with live imaging and *in situ* hybridization, anteriorly travelling waves of clock expression were shown to demarcate the location of a future somitoid. These recent advances are the most direct supporting evidence for the Clock and Wavefront model. However, *in vitro* models have yet to create bilaterally paired somites, and instead create sprawling 'rosettes' of somites within a circular tissue or singular longitudinally oriented somites (i.e. an axioloid). The creation of third-generation models which form bilaterally paired somites is currently underway (Personal communication with Dr. Cantas Alev, April 20th, 2023).

1.3.3. Molecular Mechanisms of Somitogenesis: The Clock and NOTCH

The beginning of somitogenesis is marked by oscillation of the segmentation clock. Although its priming mechanism remains elusive, tempo, intervals, and synchronicity are modulated by NOTCH signalling. The basic pathway is as follows (Figure 4; Henrique & Schweisguth, 2019; Kopan & Ilagan, 2009): The Notch Intracellular Domain (NICD) and any of the four Notch Extracellular Domain's (NECD) are expressed and transported through the Endoplasmic Reticulum (ER) to the Golgi-apparatus of a signal receiving cell. During this time, the NECD is fused to the NICD and the NECD is glycosylated in one of many different patterns (see section 1.4). The newly formed NOTCH receptor is transported to the cell membrane where the NECD binds to a juxtaposed, signal-sending cell's Delta Like Ligand 1 (DLL1), Delta Like

Ligand 4 (DLL4), Jagged 1 (JAG1), or Jagged 3 (JAG3) ligand. Upon binding, a cleavage site within the NECD is revealed by the pulling between ligand and receptor. This site is then cleaved by A Disintegrin and Metalloprotease 17 (ADAM17). Sender cells then endocytose the ligand-NECD complex whereas receiver cells endocytose the NICD-transmembrane complex. Gamma secretase cleaves the endocytosed NICD at the transmembrane domain releasing it for nuclear translocation, allowing it to bind to the transcription factors CSL (an acronym for: CBF-1/RBPJ-κ in *Homo sapiens*/*Mus musculus* respectively, Suppressor of Hairless in *Drosophila melanogaster*, Lag-1 in *Caenorhabditis elegans*) and Mastermind-like Protein 1 (MAML). Finally, this complex transcribes a variety of target genes which will be explored in more detail next.

One would expect that regardless of which ligand binds NOTCH1 at the membrane, the same set of target genes would be transcribed as all combinations lead to nuclear NICD translocation. Counterintuitively, it is well documented that each combination potentiates different gene expression patterns within the receiving cell (Boareto et al., 2015; Kakuda et al., 2020; Nandagopal et al., 2018). For example, in mouse PSM, DLL1-NOTCH1 binding promotes a mesenchymal fate whereas DLL4-NOTCH1 binding promotes a myogenic one (Nandagopal et al., 2018). This result is not due to a qualitative difference in the NICD, but rather a quantitative difference dictated by the number and organization of ligands at the membrane. DLL1 ligands colocalize by the hundreds at the sender-cell membrane in tightly packed groups known as ‘puncta’ whereas DLL4 ligands remain interspaced. This difference means that many NECD’s must bind to the DLL1 puncta to create the tension necessary for exposure of their ADAM cleavage sites. Once the tensile force of the NECD-DLL1 cluster reaches a critical threshold,

their NICD's are simultaneously cleaved and translocated to the nucleus of the receiver cell, leading to pulsatile expression of the NOTCH pathway. In contrast, because DLL4 ligands are interspaced they lead to sustained NOTCH pathway signalling. Amongst others, pulses are generally associated with *Hairy and Enhancer of Split* (*HES*) family transcription factor expression whereas continual endocytosis is associated with *Hairy/Enhancer-of-Split Related with YRPW Motif Protein* (*HEY*) family transcription factor expression. Each of these families targets different genes which potentiate the bifurcation of cellular fates previously mentioned (M. Zhou et al., 2012). In this way, receiver-cell fate is determined by intracellular NICD concentration, a result of sender-cell ligand expression.

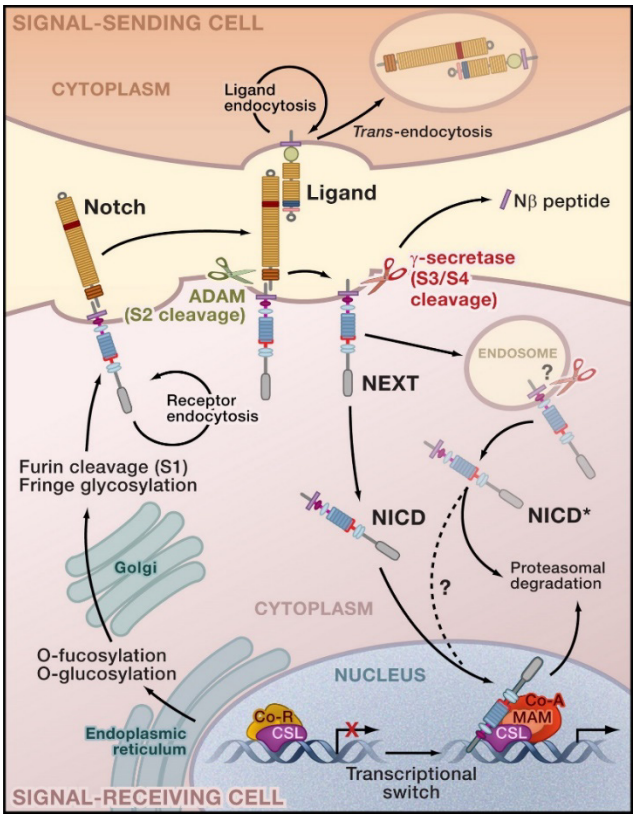


Figure 4: Diagram of the Canonical Notch Signalling Pathway. Note that LFNG is responsible for glycosylation of NOTCH receptors. LFNG is expressed in response to NOTCH signalling and localized to the Golgi (Kopan & Ilagan, 2009).

Temporary intervals of NOTCH1-DLL1 binding are responsible for maintaining segmentation clock oscillations and cell-to-cell synchrony (Boareto et al., 2015; Bochter et al., 2022; Kakuda et al., 2020; Kakuda & Haltiwanger, 2017). This interaction is promoted by glycosylation of Epidermal Growth Factor Receptor (EGFR) like repeats on NOTCH ligands and receptors; specifically, the NOTCH1-NECD, DLL1, and DLL3. Glycosylation first occurs in the ER with the addition of fucose by Protein O-Fucosyltransferase 1 (POFUT1). Homozygous variants in *POFUT1* are embryonically lethal from drosophila to mammals, demonstrating the necessity of NOTCH glycosylation for development (Irvine & Okajima, 2002; Shi & Stanley, 2002). However, the primary driver of clock oscillation is the extension of this fucose monosaccharide by a β-1,3-Nacetylglucosaminyltransferase called LFNG in the Golgi of both sender and receiver cells (Evrard et al., 1998; Rampal et al., 2005;see section 1.4.2). Although the precise interactions governing clock cycles are incompletely understood, they can be conceptualized as follows (based on the newest model produced by Bochter et al., 2022; Figure 5): A sender cell within the PSM exists in a clock 'off' state with high *DLL1* expression and low *NOTCH1* and *LFNG* expression. Here, sender cell identity is maintained by constitutive expression of DLL3, a protein which sequesters NOTCH1 away from the membrane. In response to an unknown stimulus, NOTCH1 expression rises above the threshold of DLL3 inhibition, and through canonical NOTCH signalling, results in the release of a small amount of NICD to the nucleus. One of the targets of the NICD-CSL transcription factor is *LFNG*, which acts to prevent DLL3-NOTCH1 binding *in cis* and promote NOTCH1-DLL1 binding *in trans* through DLL3/NOTCH1 glycosylation. Note that *trans* signalling is the chief function of receiver cells in the NOTCH pathway, and therefore this cell has begun its transition away from a sender state. As more glycosylated NOTCH1 is translocated to the membrane, puncti are endocytosed,

transcribing *HES7* amongst other somite-characteristic genes. However, after some time, enough HES7 is translated to perform its chief function within the clock- repression of its own transcription and the transcription of *LFNG*. As LFNG levels decrease, so too do NOTCH1-DLL1 interactions *in trans*, and DLL3 is again able to sequester NOTCH1. The cell is transitioning back to a sender state. As time goes on, DLL1 expression increases and HES7 levels fall below meaningful levels, returning the cell to the first stage of the clock cycle. There are limitations to this model, most evidently the lack of a stimulus for the initial increase in NOTCH1 expression. However, this model helps to illustrate the dynamic interplay between NOTCH pathway components and how glycosylation is key to control of the segmentation clock. Specific details of *LFNG* regulation and glycosylation are explored in section 1.4.

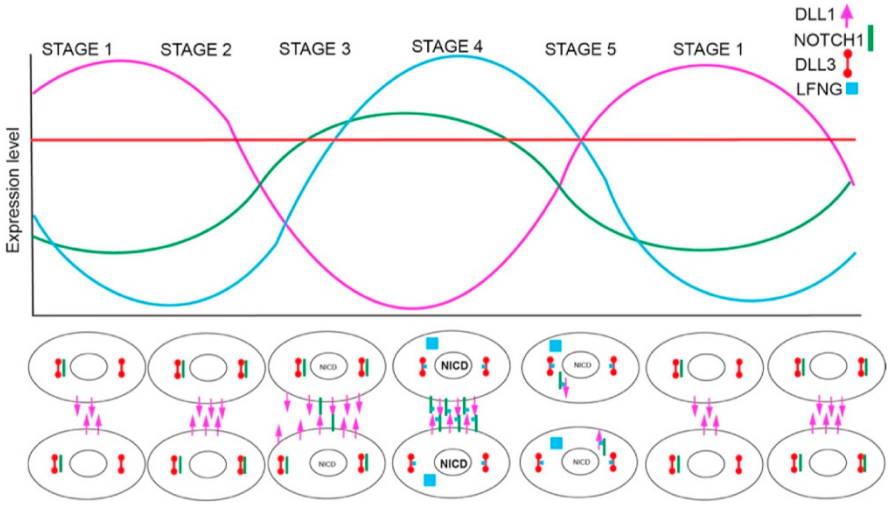


Figure 5: Schematic of Clock Coordination by LFNG and DLL3. "At the top a schematic of RNA expression levels of critical clock components. *Dll1* (pink) and *Notch1* (green) oscillate out of phase. *Lfng* (blue) transcription is triggered by NOTCH1 activation and thus *Lfng* oscillates in phase with Notch but is slightly delayed. *Dll3* expression levels (red) do not appreciably change over time. At the bottom, schematics of neighboring cells and proposed protein expression" (Bochter et al., 2022).

1.3.4. Somitogenesis: From Clock to Vertebrae

Pre-somites begin to differentiate in response to disinhibition of *Mesoderm Posterior Protein 2* (*MESP2*) activity. Although *MESP2*'s enhancer is activated throughout the PSM by *TBX6*, it is inhibited by FGF and activated by NOTCH, and thus cannot act until it reaches the DF (Oginuma et al., 2008; Yasuhiko et al., 2006). Within this region, MESP2 is expressed highly in some pre-somitic cells but minimally in others, the mechanism of which is currently unknown (Miao et al., 2023; Miao & Pourquié, 2024; Saga, 2012). $MESP2^{high}$ cells have a different transcriptomic signature from $MESP2^{low}$ cells and migrate to the anterior and posterior pre-somitic poles, respectively. This is critical for the formation of somites as the posterior-border cells express the Ephrin B2 and anterior-border cells express Ephrin A4 (Naganathan & Oates, 2020). Ephrin B2 repulses the posterior pre-somite and ephrin A4 repulses the anterior somite, both leading to the creation of an intersomitic fissure through cytosolic actin and extracellular matrix (ECM) remodelling. Once formed, anterior cells express *RIPPLY1* and *RIPPLY2*, both of which degrade TBX6, thus precisely delineating the boundary of a new somite (Oginuma et al., 2008; Zhao et al., 2015). Due to their critical role in somite formation, pathogenic variants in *MESP2*, *RIPPLY1/2*, and *TBX6* can lead to SCD3, SDV's, and scoliosis (see section 1.3.5) (Nóbrega et al., 2021).

The molecular changes described above lead to a mesenchymal-to-epithelial transition (MET), allowing somitic cells to differentiate into specialized tissues (Nóbrega et al., 2021). These tissues include the dermatome (skin), myotome (skeletal muscle), syndetome (connective tissue), and sclerotome (axial skeleton) (Musumeci et al., 2015; Nóbrega et al., 2021; Venzin & Oates, 2020). Specialization is primarily determined by proximity to other embryonic tissues.

Most relevant to vertebral development is the sclerotome which is derived from ventromedial cells most proximal to the neural tube floor plate and notochord (Fleming et al., 2015; Keynes, 2018). These cells reverse to mesenchymal tissue through an epithelial-to-mesenchymal transition (EMT) under the control of paracrine Sonic Hedgehog and Noggin signalling (Nóbrega et al., 2021). This allows the cells to migrate toward the presumptive vertebral body loci and express the sclerotome-characteristic *PAX1* and *SOX9* genes. After specialization, sclerotome from adjacent to rostral and caudal somites interlace and form a primordial vertebra (Fleming et al., 2015; Keynes, 2018). By utilizing two somites instead of one, other progenitor cells (e.g. dermatomal, myotomal, and syndetomal cells) can be utilized in the final structure. Finally, regional specification of the five spinal segments (e.g. cervical, thoracic, etc.) is determined by HOX signalling (Dubrulle et al., 2001; Hubaud & Pourquié, 2014), which will not be discussed here.

Underlying the cellular and molecular controls of somitogenesis are environmental factors such as temperature and tissue oxygenation. Temperature is an important predictive factor of somitogenesis as higher body temperature species tend to possess shorter clock intervals than those with lower temperatures (Friedmann, 1960; Schröter et al., 2008). However, this is not always the case as human and mouse embryos incubate at the same temperature but have different clock oscillation speeds (90 minutes and 5-6 hours, respectively). In attempting to explain this phenomenon, it was found that mouse PSC cells have more mitochondria per unit-volume than human PSM cells (Diaz-Cuadros et al., 2023). This correlates with a two-fold increase in [NAD+]:[NADH] ratio in mouse PSC compared to human PSC. When this ratio was decreased in human PSC somitogenesis models, the clock period increased, with the reverse

found as well. Based on these findings, it was suggested that NAD+ availability allows segmentation clock machinery to operate more efficiently, decreasing clock intervals. Therefore, the predictive value of temperature may be as a marker of redox activity (Diaz-Cuadros et al., 2023). Interestingly, ATP levels did not contribute to changing clock intervals. Although more species need to be studied to test the hypothesis more rigorously, the role that mitochondrial gene variants could play in scoliosis and other vertebral disease may be a direction for future work.

Hypoxia is a significant risk factor for the development of scoliosis and SDV's, highlighted by an altitude-dependant increase in the prevalence of congenital scoliosis (Hou et al., 2018; J. Zhou et al., 2023). Generally, it is thought that hypoxia raises the threshold of gene activity required for normal segmentation (Sparrow et al., 2012). This means that haplosufficient genes can become haploinsufficient under hypoxic conditions, promoting phenotypic penetrance of recessive alleles in heterozygous individuals. Supporting this hypothesis, the fetuses of mice with heterozygous *Mesp2*, *Hes7*, *Dll3*, *Dll1*, or *Notch1* genotypes whose mothers were exposed to hypoxia had far more mild, moderate, and severe SDV's than mothers that were exposed to normoxic conditions (Sparrow et al., 2012). Even WT mice exposed to hypoxia *in utero* presented with more SDV's than control mice. The mechanism is currently thought to stem from a hypoxic-dependent decrease in *WNT3A* expression that leads to altered wavefront signalling, secondarily inhibiting *HES7* oscillation. Although the precise mechanism is not completely understood, this could mean that heterozygosity and/or hypoxia are etiological contributors to some idiopathic scoliosis cases.

It is important to recognize that somite symmetry and structure are directly responsible for the symmetry and structure of the axial skeleton. Furthermore, due to the ubiquity of

mesodermal tissue and structural utility of the vertebral column, segmentation also indirectly contributes to organization of nervous (e.g. intercostal Plexi, dorsal/ventral/sympathetic chain ganglia, spinal cord), nephrogenic (metanephros), and pulmonary (pleuroperitoneal folds and phrenic innervation) systems, amongst others (Matsumoto et al., 2021; Nóbrega et al., 2021). Therefore, pathogenic variants in genes which control somitogenesis lead to malformations of the vertebrae and ribs and secondary pulmonary pathologies (e.g. RLD).

1.3.5. Disruption of Somitogenesis: Spondylocostal Dysostosis

Spondylocostal Dysostosis (SCD) is a monogenic condition caused by asymmetrically set somites that are incompletely segmented and limited in number (Berdon et al., 2011; Nóbrega et al., 2021; Rimon et al., 1968; Figure 6). SCD is characterized by a short vertebral column with >10 sequential SDV's along with asymmetrically set and misshapen ribs (Turnpenny et al., 2009). Scoliosis often manifests secondarily to compressive asymmetry imparted by SDV's (i.e. block, hemi, and butterfly vertebrae) whereas thoracic insufficiency presents secondary to the short and malformed thoracic cage. Both factors can lead to RLD during periods of rapid growth, the most common cause of mortality in SCD. However, with bracing, pulmonary support, patient education, and surgical intervention (if necessary), affected individuals can survive into late adulthood. The true prevalence of SCD is unknown, although a broader category of SCD-like diseases had an estimated prevalence of 1/40,000 in 1994 (Berdon et al., 2011; Martinez-Frias et al., 1994).

There are six SCD subtypes, each caused by pathogenic variants in different genes that regulate somitogenesis: SCD1- *DLL3*; SCD2- *MESP2*; SCD3- *LFNG*; SCD4- *HES7*; SCD5- *TBX6*; SCD6- *Ripply2* (Nóbrega et al., 2021). SCD1, 3, and 4 perturb the clock whereas SCD2,

5, and 6 alter post-segmentation clock somite maturation. The inheritance pattern of all SCD subtypes is autosomal recessive (except for one SCD5 family) with SCD1 being the most common at an intra-population prevalence of approximately 60% (Turnpenny et al., 2003).

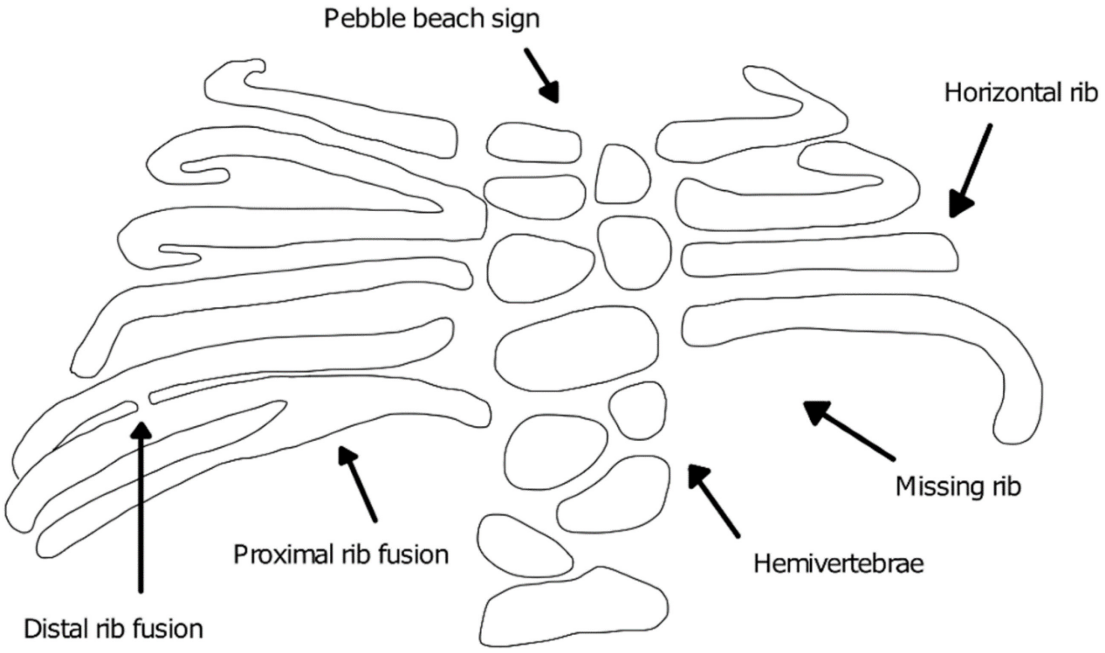


Figure 6: Diagram of Common SCD Features. Each subtype of SCD tends to present with these phenotypic manifestations, although the location and precise changes can differ (Nóbrega et al., 2021).

1.4. The Relationship Between LFNG, SCD3, and Scoliosis

The genes responsible for SCD are intrinsic to vertebral development and provide a unique window to study the etiology of scoliosis more broadly. This is because certain causes of idiopathic scoliosis may be rooted in these fundamental pathways, with small alterations in signalling leading to subtle changes in vertebral anatomy (see section 1.3.3; Giampietro, 2012). One of the subtypes- SCD3- is particularly useful in beginning to address this question. Unlike SCD's caused by genes which encode ligands (*DLL3*) or transcription factors (*MESP2*, *HES7*, *TBX6*, *Ripply2*), SCD3 is caused by variants in an enzyme-coding gene. This provides two benefits: Primarily, missense variants can dimmish enzyme activity to various degrees, and this can be measured precisely *in vitro* (Otomo et al., 2019; Sparrow et al., 2006; Takeda et al., 2018) Changes which exist on the threshold between pathogenic and benign may help to elucidate the subtle shifts in vertebral development that could cause scoliosis. Although mRNA dosage can be experimentally altered for each of the six genes, generalizability may be limited due to the lower likelihood of knock-down as opposed to missense variants in human disease (see section 1.4.5). Secondly, because both direct (e.g. enzyme activity) and downstream (e.g. signalling pathway activity) effects can be measured, a clearer link can be drawn between genetic variant, amino acid substitution, protein dysfunction, and pathway activity (see section 1.4.6). Both benefits highlight the practical utility of applying SCD3 and *LFNG* projects to the study of scoliosis etiology, granting that all SCD genes have costs and benefits.

The discovery that pathogenic *LFNG* variants were responsible for a new subset of SCD came in 2006 (Sparrow et al., 2006) in the wake of extensive research into Drosophila wing development. In 1994, a gene was discovered within the 35UZ-1 P-element which regulates the

margin between dorsal and ventral components of wings (Irvine & Wieschaus, 1994). This gene was coined *Drosophila fringe* (*fng*) and was determined to regulate the Notch signalling pathway. Three years later, homologues were discovered in mouse embryos: *Lfng*, *Manic Fringe* (*Mfng*) and *Radical Fringe* (*Rfng*) and it was determined that this group of genes encoded glycosyltransferases that modified NOTCH (Johnston et al., 1997). Interestingly, *Lfng*, but not *Mfng* nor *Rfng*, appeared to oscillate throughout mouse PSM and knocking out *Lfng* inhibited clock oscillation. Further evidence showed that the boundaries between presumptive and mature somites were no longer demarcated, demonstrating that glycosylation of NOTCH1 during somitogenesis was key to segmentation clock oscillation and somite boundary formation (Evrard et al., 1998; Johnston et al., 1997; Moloney et al., 2000; Rampal et al., 2005). Indeed, later work identified a multiomic regulatory network governed by *Lfng*, the details of which will be explored next.

1.4.1. Transcriptional Regulation of *LFNG*

Precise control of *LFNG* transcription is key to segmentation. This occurs mostly *in cis* at three major loci within *Fringe Clock Element 1* (*FCE1*), an evolutionarily conserved sequence of 110bp that lies -1.3Kb from the transcription start site (Cole et al., 2002; Irvine et al., 1991). If *FCE1* is eliminated from mouse embryos, clock oscillation is diminished. The first two loci within *FCE1* are binding sites for basic helix loop helix (bHLH) transcription factors, notably *HES7*. HES7 represses *LFNG* transcription during the clock cycle (Stauber et al., 2009). The third site promotes CSL binding, a transcription factor which is activated by NICD binding via NOTCH signalling (Henrique & Schweisguth, 2019).

The multiplicity of factors which govern *FCE1* transcription factor binding suggests that it acts not as a binary 'on/off' system, but rather as a sliding scale of *LFNG* dosage/activity (Cole et al., 2002; Oginuma et al., 2010; Shifley et al., 2008; Stauber et al., 2009; D. R. Williams et al., 2014). *LFNG* is haplosufficient in mice and humans, and the level of activity required from the 'functional' allele to maintain clock oscillation has received some investigation. Using combinations of the *Lfng*, $Lfng^{ko}$, and $Lfng^{\Delta FCE}$ alleles in transgenic mice, it was determined that LFNG affects somitogenesis in a dosage dependent fashion, reflected in the number and structure of both somites and vertebral bodies of the transgenic mice (Shifley et al., 2008). This finding was echoed soon after in a similar study, revealing a linear relationship between *LFNG* dosages and anterior skeleton (thoracic and cervical spine) length in transgenic mouse models of SCD3 (Stauber et al., 2009). It was revealed that each region of the presumptive spinal column requires a different potency of *LFNG* dosage, the highest of which are the thoracic vertebrae (Oginuma et al., 2010). This helps to explain why SDV's often noted within the thoracic trunk of SCD3 probands (Lefebvre et al., 2018; Otomo et al., 2019; Schuhmann et al., 2021; Sparrow et al., 2006; Takeda et al., 2018). Although the absolute levels of LFNG expression required are not quantifiable with this data, it suggests that the number of SCD3 vertebral abnormalities could present on a spectrum that is dependent on allelic functionality. It is interesting to consider whether isolated SDV's and/or small changes to the vertebral column that result in scoliosis are a result of partially active, yet haplosufficient genetic alleles (*LFNG* or otherwise).

Post-transcriptional regulation of *LFNG* is controlled by transcript degradation, a common mechanism for many NOTCH pathway genes during somitogenesis, including *HES7* (McIntyre et al., 2020). *LFNG*'s canonical transcript (LFNG_201, RefSeq NM_001040167.2) is

2,445bp long with a short 5' UTR, 8 protein-coding exons, and a longer 3' UTR (1087bp) (UniProtKB, 2021). The long 3'UTR hinted at a regulatory role for this region, especially in the context of its very short half-life compared to *HES7*, another critical mediator of the segmentation clock (Nitanda et al., 2014; Stauber et al., 2009; Zeiser et al., 2006). The precise mechanism by which the 3'UTR promotes earlier degradation is not well understood, but it is thought that a conserved adenine and uridine-rich (AUUUA) microRNA (miRNA) binding site may contribute (Riley et al., 2013; Wahi et al., 2017). Specifically, it was shown that mir-125a-5p is responsible for limiting *lfng* half-life during chick somitogenesis, enhancing the rate of mRNA turnover (Riley et al., 2013). However, this finding was not replicated in mouse models of somitogenesis as it was shown that mir-125a-5p knockout did not affect somitogenesis or clock timing (Wahi et al., 2017). This conflicting data raises the possibility of species-dependant post-transcriptional regulatory mechanisms.

In sum, both pre- and post-transcriptional regulation of *LFNG* play significant roles in clock interval timing, synchronicity, and somitogenesis more broadly. It is important to recognize that small changes to the many feedback loops can lead to significant changes to the formation of somites and thus vertebrae.

1.4.2. Overview of LFNG Protein Structure and Functions

Post-translational modification of LFNG is vital to clock-interval timing, and these changes are a result of specific structural elements possessed by this enzyme (Luther et al., 2009; Shifley & Cole, 2008; D. R. Williams et al., 2016). To understand these regulatory processes, it is important to first review the basic biochemical mechanisms which govern its structure and function. Note that the sequence and function of LFNG are highly conserved between species,

and thus mouse *Lfng* is often used for the analysis of SCD3-associated variants *in vitro* (Bochter et al., 2022; Luther et al., 2009; Otomo et al., 2019; Shifley & Cole, 2008; Takeda et al., 2018; D. R. Williams et al., 2016). Although not studied directly, it is believed that the difference between modelling variants with mouse or human *LFNG* is negligible, outside of small changes to the position of key residues. Therefore, for the sake of clarity, all future discussions will reference human LFNG unless otherwise stated (i.e. using Lfng to denote mouse orthologue)

LFNG is a member of GT-A fold inverting glycosyltransferase family, an evolutionarily conserved group of enzymes which utilize Rossman folding (which generally promotes sugar-binding) and divalent cations to form linkages between monosaccharides (Rini et al., 2022). The role of the β-1,3-N-Acetylglucosaminyltransferase encoded by *LFNG* is to link a Glc-NAc molecule from UDP-α-GlcNAc to an acceptor fucose, ultimately forming a GlcNAc-β-1,3-Fuc disaccharide (Luther et al., 2009; Moloney et al., 2000; Rampal et al., 2005). The full-length, 'unprocessed' protein is 356 amino acids and has two major components, an N-terminal transmembrane domain which anchors it within the Golgi lumen, and an enzymatically active C-terminal responsible for catalysis (Figure 7). However, the N-terminal can be cleaved, and processed LFNG is 266 amino acids in length. Both unprocessed and processed LFNG are enzymatically active (Rampal et al., 2005; Sparrow et al., 2006).

The catalytic component of LFNG is characterized by the FRINGE domain (amino acids 108-358) which hosts an active site with three key motifs: a DxD divalent cation (Mn^{2+}) binding site (p.D200-p.D201-p.D202), the catalytic residue (p.D290), and a Mn^{2+} chelating residue (p.H314) (UniProtKB, 2021). It is thought that these motifs work in tandem, with p.D290 sequestering a proton from a fucose molecule previously localized to the binding pocket by Mn^{2+}

(Luther et al., 2009; Sinnott, 1990; Ünligil & Rini, 2000). Proton sequestration allows the newly formed nucleophile (fucose) to attack the anomeric carbon of an adjacent Glc-NAc, inverting its orientation and creating the β-glycosidic linkage. The Mn^{2+} chelating residue (p.H314) plays a minor role in stabilizing the structure throughout catalysis. Of note, UDP is released as a byproduct of this reaction at a 1:1 ratio. Contemporary *in vitro* quantifications of LFNG glycosyltransferase activity utilize this relationship to determine the effect of a variant on enzyme function (Otomo et al., 2019; Takeda et al., 2018). By incubating post-reaction UDP-laden (or UDP-absent, in the case of null variants) supernatant with UDP-dependant luciferase, luminescence can then be quantified and compared to null and WT controls to determine relative activity.

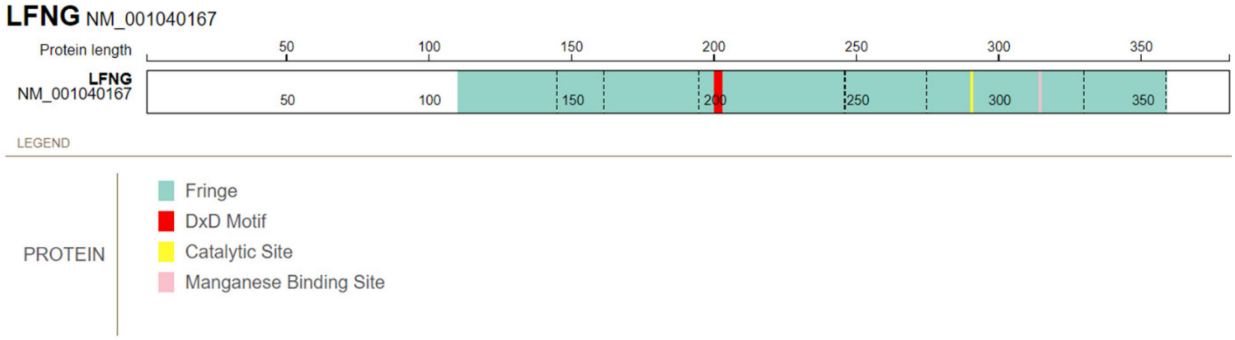


Figure 7: Model of LFNG Primary Structure. This model is based on LFNG's canonical transcript NM_001040167 (hg19). Numbers indicate the amino acid position in LFNG's primary structure. The N-terminal anchor is depicted in white to the left of the enzymatically active Fringe domain (turquoise). The DxD motif (red) is comprised of the p.D200-D201-D202 residues, the catalytic site (yellow) is residue p.D290, and the manganese binding site (pink) is residue p.H314.

This figure was adapted from the freely available demonstration program, "ProteinPaint: Web application for visualizing genomic data (SJ-15-0021©)" created by St. Jude Research Hospital (https://proteinpaint.stjude.org/)

It should be noted that these assays provide a raw value of total reaction activity over a set period and therefore are useful in assessing whether a variant completely inhibits, or fails to inhibit enzyme function. However, kinetic analysis provides a much more nuanced perspective of *LFNG* variants, particularly those which retain partial enzymatic activity. Such variants have been identified through targeted mutagenesis of the 'short loop' (encoding p.C168, p.A176, p.S178, p.C179, p.D289, p.D290, and p.C291) and the 'long loop' (encoding p.T253, p.S228, and p.P251) of LFNG. Sparing the intricate details, both loops are on opposite sides of the active site and change in conformation during catalysis. It was shown that variants within the short loop were more likely to retain Km but diminish Vmax whereas variants in the long loop tended to elevate Km beyond physiological limits. In other words, this work suggests that short-loop variants are likely to reduce the total number of reactions whereas long-loop variants tend to prevent them altogether. Therefore, screening for the functional activity of any short-loop variants identified in SCD3 probands may be of value in determining whether they can modulate SCD3 presentation in the same way gene dosage does. Although this model has yet to be tested by other groups, it does provide a broad tool to contextualize the enzymatic effects of *LFNG* variants in different regions.

As opposed to the enzymatically active C terminal, the first 86 amino acids of LFNG encode an α-helix that acts as a single-pass, type-II transmembrane domain (UniProtKB, 2021; Figure 7). By anchoring LFNG to the Golgi membrane, it orients the FRINGE domain towards the Golgi lumen, the major site of NOTCH protein glycosylation (Shifley & Cole, 2008). There has yet to be an SCD3 case study that identifies a normally localized LFNG variant (Otomo et al., 2019; Sparrow et al., 2006; Takeda et al., 2018) (note: these variants often inhibit enzyme

activity as well). Although it is not known where the variants mislocalize to, the ER appears to be a likely candidate as it is well known that protein misfolding due to perturbing amino acid substitution prevents normal trafficking (Hegde & Zavodszky, 2019). Indeed, previous evidence from disease-causing variants in Glycosyltransferase 8 Domain–Containing Protein 1 (GLT8D1), another GT-A fold family protein, have demonstrated ER-retention as a result of misfolding (Tsai et al., 2021). However, another work has suggested that mislocalization may be a result of a yet-undiscovered domain within the short-loop of LFNG that promotes translocation to the Golgi (Sparrow et al., 2006). This suggestion was based on the discovery of such a region within the Drosophila protein O-fucosyltransferase protein (Okajima et al., 2005). As very few variants have been studied, there is not yet enough data to support either of the hypotheses.

1.4.3. Post-Translational Regulation of LFNG

The presence of a conserved, dibasic cleavage site (p.R86 – p.A87 – p.R88 – p.R89) between the transmembrane and FRINGE domains hinted at the possibility of a protease-dependant regulatory mechanism. These sites are targeted by the Subtilisin-like proprotein convertase (SPC) family enzymes, and thus western blot analysis was conducted from the media of *Lfng*-transfected cells with various amounts of the SPC inhibitor α1-PDX (Shifley & Cole, 2008). As the concentration increased, the intensity of the 45kDa band (uncleaved/unprocessed/pre-pro-LFNG) increased whereas the 35kDa band decreased. Subsequently, co-transfection of *Lfng* with members of the SPC family indicated that SPC6 cleaved the dibasic site most efficiently and that *SPC6* is co-expressed with *Mesp2* at the end of the clock cycle in pre-somites of mouse embryos. These results suggested that cleavage of the Golgi-anchor leads to LFNG exocytosis, finely tuning the amount of time LFNG can potentiate

clock oscillation. Indeed, later work which fused the dibasic-cleavage-site free transmembrane domain of *RFNG* to the N-terminal of *LFNG* supported this model (D. R. Williams et al., 2016). These *R/LFNG* heterozygous mice had disorganized somites with large anterior compartments (D. R. Williams et al., 2016), likely due to excess NOTCH signalling desynchronizing the clock and creating a disproportionate number of $MESP2^{high}$ cells, respectively (Miao et al., 2023). The result of this was a pronounced SCD-like phenotype, demonstrating the key role of post-translational regulation of LFNG. This data suggests that variants which inhibit processing could cause SCD phenotype and predicts that excess cleavage may also be pathogenic. Therefore, the investigation of protein processing is critical in the assessment of *LFNG* variants associated with SCD3 phenotype. The precise mechanism by which LFNG is exocytosed remains unknown.

The relationship between LFNG localization and protein processing has not yet been directly discussed. However, previous data has shown that a variant which caused mislocalization was not secreted (c.564C>A [p.F188L]) (Sparrow et al., 2006). Although only one variant has been evaluated, if more variants are found to behave in this way, it would suggest that variants which cause mislocalization necessarily prevent processing. This is because if LFNG is not in the Golgi lumen, it cannot be cleaved by SPC6 (Shifley & Cole, 2008). Therefore, immunofluorescence localization assays and western blots may simply detect LFNG trafficking in different ways. More work is needed to test this hypothesis as the cause of mislocalization and misprocessing are unknown and only one variant has been evaluated with both methods (Sparrow et al., 2006).

1.4.4. Overview of Variant Nomenclature

The previous section highlights that all aspects of *LFNG* structure, function, and regulation contribute to somitogenesis, and therefore each level requires consideration in the context of variant interpretation. Not only does this create a clearer understanding of the proband's disease, allowing for more accurate management, but it also informs more broadly about normal and abnormal development. Before exploring the effect of variants on *LFNG* and their implication in SCD3, a review of common knowledge variant nomenclature in medical genetics will be provided.

The largest genetic changes (from 1000bp to entire chromosomes), sometimes referred to as regional or structural variants, are deletions and duplications (Jackson et al., 2018). Deletions remove a large portion, if not the entirety of an allele whereas duplications lead to the multiplication of an allele. Deletions are very often pathogenic as the allele, and thus its product has been completely removed. Duplications can be pathogenic in some contexts, the details of which will not be discussed here. Insertion-Deletions (Indels) are similar in that DNA is added or removed from a locus, but these changes are less than 1000bp, and often one or two nucleotides. Changes which add or subtract one or two shift the gene's reading frame (i.e. frameshifts) often leading to early stop codons (a.k.a. nonsense variants), truncating the protein and causing disease. Point variants are changes to one nucleotide and can occur within coding or non-coding regions of a gene in a disease context. They are classified as silent, missense, and nonsense. Silent variants change a nucleotide but not the resulting amino acid, missense variants change the resulting amino acid, and nonsense variants encode a premature stop codon at that site. Therefore, disease can arise from silent variants through splice site abrogation, codon bias, or

transcription start site changes (Liu et al., 2021) whereas nonsense variants lead to a truncated, non-functional protein (Jackson et al., 2018).

Missense variants are unique in that the effect they have depends on which amino acid is encoded in place of the original (Jackson et al., 2018). Generally, 'conservative' substitutions within the same amino acid class (e.g. non-polar to non-polar) are likely to be less damaging whereas 'non-conservative' changes of one amino acid class to another (e.g. acidic to polar) tend to be more damaging (Z. Zhang et al., 2012). However, this general rule is not always the case. For example, variants in the *Collagen Type II Alpha 1 Chain* (*COL2A1)* gene that cause glycine to serine substitutions are highly damaging as serine's large size destabilizes collagens quaternary triple-helix leading to osteogenesis imperfecta (Marini et al., 2017).

The effect of a genetic variant on overall gene function is typically categorized with "Mullers Morphs" (Muller, 1932). Amorphs (a.k.a. Nullomorphs, null variants) are alleles with complete loss of gene function due to dysfunctional protein product (a.k.a. loss of function) or loss of expression (a.k.a. knock-out), and therefore often cause disease. Hypomorphs are alleles with partial activity due to tapered expression (a.k.a. knock-down) or lowered functional capacity (a.k.a. partially active), and their role in disease is gene and context-dependent. Hypermorphs are alleles with enhanced activity due to increased expression or increased catalytic activity, and neomorphs are alleles with a new function due to a change in protein structure. Both are interchangeably referred to as "gain-of-function" variants and therefore it is important to consider context before assuming functional result. Therefore, the contribution of these alleles to disease is also context-dependent. Finally, antimorphs (a.k.a. dominant-negative variants) are typically found in autosomal dominant disease as these variants are epistatic to their

partner allele(s). One commonly used generic example is an antimorphic allele of a homodimerized transcription factor that prevents quaternary structure formation, preventing the WT gene product from functioning.

Although deletions and frameshifts are very likely to cause disease, the unique effect of each amino acid encoded by a missense variant requires significant experimental evidence to determine the functional effect (Richards et al., 2015). This process can extend a patient's 'diagnostic odyssey' by years or more. To address this limitation, *in silico* algorithms (e.g. CADD, PolyPhen-2, REVEL, etc.) have been created to predict the likelihood that a variant (missense or otherwise) is pathogenic or benign (Adzhubei et al., 2013; Ioannidis et al., 2016; Rentzsch et al., 2019). This is typically calculated with a combination of factors such as population-based variant incidence, evolutionary conservation, and structural modelling. At the poles of this spectrum, *in silico* analysis is very useful and can quickly determine the functional effect of a variant. However, missense variants occasionally fall somewhere in the middle, with their effect on protein structure, and thus disease, uncertain.

1.4.5. Variant Classification

To address the difficulty in determining the effect of a genetic variant on human health, The American College of Medical Genetics and Genomics (ACMG) created the "Standards and Guidelines for the Interpretation of Sequence Variants" (Richards et al., 2015). Sparing the details, *in silico* results, allelic frequency, homo/heterozygosity, type of variant (i.e. frameshift, silent, etc.), variant loci, *in vitro*/*in vivo* results, proband phenotype, and parent phenotype, amongst others, are utilized to develop a likelihood of pathogenicity. Based on these factors, a variant can be classified into one of the following five categories: Pathogenic, Likely Pathogenic,

Variant of Uncertain Significance (VUS), Likely Benign, and Benign. A variant which is classified as pathogenic or likely pathogenic is treated as disease-causing, and therefore clinical intervention and prenatal screening are recommended. Variants classified as benign and likely benign are treated as non-contributory to disease, and other etiologies are considered. However, VUS pose a particular challenge as their contribution to disease is unknown. These are typically missense variants due to the large variety of changes that they can impart (see section 1.4.4). Although it is recommended to monitor a proband with VUS closely, clinical action and prenatal screening are not recommended. This prevents the utilization of genetic information for reproductive risk counselling and can prevent patients from receiving curative treatment options. At this time, the ACMG guideline recommends attempting to reclassify the VUS by gathering more information that was not previously available. If all clinical or otherwise quickly attainable data has been gathered, *in vitro* functional testing is the next step, hence the 'diagnostic odyssey' for probands with VUS.

1.4.6. Considerations for LFNG Variant Interpretation

The investigation of *LFNG* variants typically requires *in vitro* functional analysis due to the high prevalence of VUS missense variants in SCD3 probands. To this end, previous work has studied the effect of variants on glycosyltransferase activity, protein localization, protein processing, and NOTCH signalling activation (Otomo et al., 2019; Sparrow et al., 2006; Takeda et al., 2018) as each of these four experiments can justify ACMG PS3 criteria (functional evidence via well-established functional assay) (Richards et al., 2015). The glycosyltransferase assay determines whether LFNG is catalytically active after isolation and *in vitro* reaction, the localization assay utilizes immunofluorescence microscopy to determine whether LFNG is

compartmentalized to the Golgi, and the protein processing assay determines whether LFNG is being cleaved by densitometric analysis of western blot images (Otomo et al., 2019; Sparrow et al., 2006; Takeda et al., 2018). The absence of any of these processes can indicate that LFNG is unable to glycosylate NOTCH receptors, preventing segmentation clock oscillation and thus explaining the SCD3 phenotype. On the contrary, the NOTCH signalling assay uses a luciferase reporter gene under the control of an NICD-responsive promoter to determine the effect of an LFNG variant on NOTCH pathway activity (Sparrow et al., 2006).

In deciding which experiments to use, it is important to consider whether the data provides evidence of pathway activity or protein function. For example, the NOTCH signalling assay cannot determine why an *LFNG* variant leads to dysfunction, just that it prevents NOTCH signalling. However, the other three experimental readouts are the consequences of a variant on the intrinsic properties of LFNG function. Therefore, these experiments indirectly provide evidence of pathway activity, and thus pathogenicity, through direct functional evidence. For this reason, previous research tends to interrogate functional parameters over NOTCH pathway activity (Otomo et al., 2019; Sparrow et al., 2006; Takeda et al., 2018).

It is important to recognize that the interpretation of functional experiments may require nuance in cases of hypomorphic alleles. Although all SCD3 variants published until 2021 completely perturb enzymatic function (see section 1.4.7; Table 3), others which have not yet been associated with SCD3 are hypomorphic (Luther et al., 2009). This is problematic as hypomorphic alleles may not fulfill PS3 criteria as the threshold of LFNG function required for somitogenesis is unknown. In these cases, a nuanced discussion of this outcome would be

necessary, contextualizing the functional results with the phenotype (Brnich et al., 2019; Kanavy et al., 2019).

Outside of *in vitro* experimentation, rudimentary *in silico* analyses are often also conducted to elucidate the fundamental interactions governing protein dysfunction (Schuhmann et al., 2021; Sparrow et al., 2006). These studies often incorporate protein modelling and conservation analysis (amongst others) to explain why certain variants perturb protein function in a specific fashion. It has been suggested that amino acid substitutions proximal to the active site (tertiary structure, not primary) perturb function independent of structure, whereas those farther perturb function as a result of structure (Luther et al., 2009). If supported, this suggests that structural perturbation is responsible for mislocalization (see section 1.4.3), and that mislocalized variants should lack enzymatic activity. Conversely, some variants which lack enzymatic activity could normally localize. This is important because alleles with hypomorphic function could arise from the short loop, localizing normally with partial enzymatic activity. It will be important to study more *LFNG* variants, especially those with atypical phenotypes, to further test the hypotheses discussed here.

As alluded to previously, a significant limitation in the study of *LFNG* variants is the lack of published data. As of 2021, only six cases were published in the scientific literature and this work often did not provide complete clinical and experimental reports (Lefebvre et al., 2018; Otomo et al., 2019; Schuhmann et al., 2021; Sparrow et al., 2006; Takeda et al., 2018). For example, phenotypic data such as: age, sex, consanguinity, height (absolute and relative), arm-span-to-height ratio, mid-parental height, vertebral body number, rib number, scoliosis presence/cobb angle, and descriptions of both axial and non-axial characteristics, were

incompletely collected (Table 2, Table 3, and Table S2). Furthermore, localization data, protein processing, and glycosyltransferase activity were only reported for one, four, and four of the variants, respectively (see Table 2). This data is vital for understanding the SCD3 genotype-phenotype relationship and will need to be collected in the future. Each of these reports, including those published between the initiation of this work and 2024 (not including our work in 2023) will be briefly summarized in the next section.

1.4.7. Review of SCD-III Publications to Date

The first SCD-III patient was described in 2006 (Sparrow et al., 2006). The proband presented with multiple SDV's, an arm span 20% longer than height, short trunk, non-progressive scoliosis, and camptodactyly. After a negative screen for variants in *DLL3* and *MESP2*, they identified homozygous c.564C>A (p.F188L) *LFNG* missense variants. The variant caused the loss of phenylalanine, a residue which is conserved throughout all Fringe proteins from mammals to drosophila. *In vitro,* there was a complete loss of enzymatic function, and the protein was improperly localized and not secreted into media. Steric hindrance and/or aromatic cluster disruption of the active site were proposed as mechanisms of protein dysfunction.

In 2018, two more SCD-III patients were described (Lefebvre et al., 2018). Sanger sequencing indicated compound heterozygous variants (c.583T>C [p.W195R], c.842C>A [p.T281K]) in the first patient and homozygous frameshifts (c.44dupG [p.A16Rfs*135]) in the second. Similarities between these and the previous case are severe shortening of the spine and asymmetrical ribs, as well as multiple segmentation defects of the vertebrae (SDV's) anterior of the lumbosacral junction. A somewhat pronounced difference was the presence of angular vertebral bodies that appeared as a 'pebble beach' in the neonatal patient (see Figure 6). This

phenotype appears to be characteristic of SCD-III although it is unknown why. The team did not report any *in vitro*, *in silico*, or biochemical analyses for any of the variants.

The fourth proband, an adolescent male, was found to have compound heterozygous *LFNG* mutations (c.467T>G [p.L156R], c.856C>T [p.R286W]) (Takeda et al., 2018). *In vitro* enzymatic assay indicated a complete loss of glycosyltransferase activity. This patient presented with a more severe scoliosis and milder SDV's as compared to previous cases. The authors believe that the severe scoliosis is due to the larger spacing between SDV's, but they do not explore this hypothesis any further. Prior to this study, the patient was diagnosed with CS and not SCD, demonstrating that misdiagnosis is not uncommon in this population.

The fifth proband was also compound heterozygous for *LFNG* variants (c.372delG [p.K124Nfs*21], c.601G>A [p.D201N]) (Otomo et al., 2019). Using the same *in vitro* enzymatic assay, their data indicated a loss of function due to the c.601G>A allele which encodes a highly conserved aspartic acid residue involved in substrate binding (p.D200; see section 1.4.2). They speculated that the c.372delG causes nonsense-mediated mRNA decay but did not explore this hypothesis experimentally. Phenotypically, the patient presented with SDV's and proximal rib fusions, but no scoliosis. This may be due to the proband's age during the investigation (nine months).

The final SCD-III case published before September 2021 was a 17-year-old male with homozygous *LFNG* missense variants (c.446C>T, [p.T149I]) (Schuhmann et al., 2021). *In silico*, it is predicted that the variant causes a loss of function due to steric hindrance of the active site. However, the hypothesis has yet to be tested *in vitro*. The proband has multiple thoracic, proximal rib fusions on the left side, as well as hemi, block, and pebble beach vertebrae spanning

the vertebral column. There is also a prominent left scoliosis. Interestingly, this proband presented with absence epilepsy, solitary pelvic kidney, uterine dysgenesis, Mayer-Rokitansky-Küster-Hauser syndrome, and inner ear deafness. This is the first time that *LFNG* variants have been associated with a significant number of extra-vertebral features, although consanguinity may have led to the accumulation of other contributory recessive alleles.

After the initiation of this work, and excluding results reported hereafter, a final group of three probands were described (Lecca et al., 2023). Like all previous cases, the probands presented with severely shortened vertebral columns, multiple SDV's, and rib malformations. The proposita passed at one month due to RLD. She possessed novel compound heterozygous *LFNG* variant's, c.863dup (p.D289*) and c.1063G>A (p.D355N), which were considered damaging and VUS, respectively (note discussion in 1.4.5 regarding the typical classifications of nonsense versus missense variants). The second proband is a three-year-old female who presented with classic SCD3 phenotype along with joint laxity, bradycardia, and unilateral hearing loss, amongst other uncommon features. Sequencing revealed a compound heterozygous genotype (c.521G>T [p.R174L]; c.890T>G [p.V297G]), both of which were predicted to be damaging but classified as VUS. The final proband is a five-year-old female with SCD3 phenotype and extra-vertebra manifestations like camptodactyly, diaphragmatic hernia, and RLD. Homozygous intronic *LFNG* variants were identified five nucleotides upstream of the splice acceptor site of exon 6 (c.822-5C>T). *In vitro* analysis indicated a significant decrease in mRNA expression, likely due to the creation of CUGBP Elav-Like Family Member 1 (CELF1) binding sites. This protein is known to regulate splicing and expression of mRNA, potentially

suggesting a role for this protein in the proband's disease. Note that it is unlikely that CELF1 regulates WT *LFNG* transcripts *in vivo* as they do not contain this binding region.

1.5. Research Goal, Aims, and Hypotheses

Scoliosis causes significant hardship ranging from pain and diminished quality of life to high-risk surgeries and even mortality (Hawes & Weinstein, 2003; Kebaish et al., 2011; Kwon et al., 2021; Sung et al., 2021). The logistical and economic challenges of management further burden the up to 1-in-33 who suffer from this progressive disease. Key to limiting these burdens and limiting Cobb angle progression is early intervention and thus early recognition (Turnpenny et al., 2009; White et al., 2020). However, most scoliosis cases present with an idiopathic etiology, stifling the clinician's diagnostic and prognostic capacity, often leading to the replacement of preventative management strategies with therapeutic ones. Identification of the underlying cause can allow for increased screening measures for at-risk populations and a more clearly defined prognosis, allowing the patient and healthcare team to limit curve progression and enhance healthcare outcomes.

Studying rare diseases associated with vertebral abnormalities offers a unique window into the developmental pathways which govern truncal anatomy. By determining one gene that causes scoliosis, other members of the same pathway can then be investigated for their role in spinal development. Furthermore, by investigating the monogenic causes of scoliosis, a more refined, nuanced understanding of its regulatory role may be elucidated. *LFNG* is a great candidate due to its significant role in trunk formation and its association with SCD3. Therefore, in the hopes of more deeply understanding how changes in vertebral column development can lead to scoliosis,

the overarching goal of this work is to explore the relationship between genotype and phenotype in SCD3 probands. This will be accomplished with *in vitro* and *in silico* investigations of *LFNG* variants in probands with SCD3 phenotypes or diagnoses. The following sequence of questions will serve to implicitly direct the body and will be addressed explicitly in the conclusion: 1) Do the variants studied here satisfy PS3 criteria? How does this affect ACMG classification? 2) How does each pathogenic variant alter LFNG function? 3) Is there a link between variant location and type of functional alteration? 4) Is there a link between *LFNG* variant and SCD3 presentation?

Chapter 2 intends to determine whether two, novel *LFNG* VUS (c.766G>A [p.G256S]; c.521G>A [p.R174H]) presenting *in trans* (i.e. compound heterozygous) are disease-causing in a proband with SCD3 phenotype. To test whether the variants satisfy PS3 criteria, glycosyltransferase activity, protein processing, and subcellular localization will be tested *in vitro*. It is predicted that the two variants are pathogenic as both variants affect highly conserved nucleotides and the probands phenotype is mostly consistent with SCD. The aims and hypotheses are as follows:

1) Quantitatively determine the effect of each variant on glycosyltransferase activity. This will be achieved *in vitro* by purifying LFNG enzyme, incubating it with a glycosyltransferase reaction mix, and measuring the luminescent output of a UDP-powered luciferase assay. It is hypothesized that glycosyltransferase activity will not be significantly different from an untransfected control as the substitutions are within the short loop (p.R174H) and long loop (p.G256S) of LFNG, both of which are critical for enzyme function (Luther et al., 2009).

2) Semi-quantitatively determine the effect of each variant on protein processing. This will be achieved with densitometric analysis of western blots from cell lysate of NIH-3T3 cells transiently transfected with *LFNG-HA* plasmids. It is hypothesized that the unprocessed (43 kDa) and processed (34 kDa) bands of both variants will not statistically vary from WT. This is because they lie close to the active site and may not perturb LFNG function as a result of structural changes (Luther et al., 2009).
3) Qualitatively investigate the effect of each variant on LFNG subcellular localization. This will be achieved through immunofluorescent imaging of anti-HA and anti-GM130 antibody signals in NIH-3T3 cells transiently transfected with *LFNG-HA*. It is hypothesized that the anti-HA and anti-GM130 signals will colocalize if protein processing is normal, but not if protein processing is diminished. This is because if LFNG is processed appropriately, then it must be cleaved by the Golgi-resident SPC6 (Shifley & Cole, 2008). If it is not being processed, then it is not being cleaved by SPC6, and thus should not be within the Golgi lumen.

In hopes of creating a more complete discussion surrounding the genotype-phenotype relationship, this chapter will also provide a clinical description of the proband, genetic and proteomic analysis of each variant, and *in silico* protein modelling to assess potential structural changes. Furthermore, this chapter aims to produce a preliminary discussion on the relationship between protein processing and localization.

The goal of Chapter 3 is to investigate the genotype-phenotype relationship more deeply by determining the effect of all seven *LFNG* variants associated with SCD3 that were published with radiographic findings before December 2021, including two identified in Chapter 2 (Tables

2 and 3). All nine variants will be tested *in vitro* for their effect on protein processing and subcellular localization, and three previously untested variants will be assessed for glycosyltransferase activity. *In silico* structural modelling and 3D distance measuring will also be undertaken for all variants. Broadly, it is predicted that all variants associated with arm-span-to-height ratios of >1.1, less than 20 vertebral bodies, and clinically defined short stature (<2 SD), will completely lack LFNG function, fully inhibit glycosyltransferase activity and/or fail to localize to the Golgi. This prediction is based on the results of Chapter 2 which propose that the SCD3-probands longer trunk was due to a partially active, hypomorphic *LFNG* allele (c.521G>A [p.R174H]) (Wengryn et al., 2023). Finally, in line with previous research (Luther et al., 2009; Wengryn et al., 2023), it is predicted that variants closest to the active site will be less likely to cause mislocalization than those which are farther away. The aims and hypotheses are listed below and partially recreate the wording of Chapter 2's aims:

1) Quantitatively determine the distance between each variant and the DxD (p.D201) motif, as well as each variant and the catalytic residue (p.D290). This will be achieved by modelling LFNG *in silico* and measuring α-carbon to α-carbon distances between each of the points and the variants. It is hypothesized that the farthest variants will be mislocalized and misprocessed whereas the closest will not. This is because variants far from the active site may perturb LFNG as a result of structural changes, and misfolding could cause ER retention (Luther et al., 2009; Tsai et al., 2021). Therefore, it is further hypothesized that glycosyltransferase activity will be inhibited regardless of proximity, as both gross structural changes and active site inhibition could lead to loss of enzyme activity.

2) Quantitatively determine the effect of the c.583T>C [p.W195R], c.842C>A [p.T281K], and c.446C>T [p.T149I] variants on glycosyltransferase activity. This will be achieved *in vitro* by purifying LFNG enzyme, incubating with a glycosyltransferase reaction mix, and measuring the luminescent output of a UDP-powered luciferase assay. It is hypothesized that these three variants will not be statistically different from null due to their proximity to the short (p.T149, p.W195R) and long (p.T281K) loops of LFNG (Luther et al., 2009; Wengryn et al., 2023), and clinical associations with probands who have fewer vertebrae (<20) and short stature.
3) Semi-quantitatively determine the effect of each variant on protein processing. This will be achieved by densitometric analysis of western blots from cell lysate of NIH-3T3 cells transiently transfected with LFNG-HA plasmids. It is hypothesized that the unprocessed (43 kDa) and processed (34 kDa) bands of variants farthest from the active site will not statistically vary from WT as they may not have gross structural alterations (Luther et al., 2009).
4) Qualitatively investigate the effect of all variants on LFNG subcellular localization. This will be achieved through immunofluorescent imaging of anti-HA and anti-GM130 antibody signals in NIH-3T3 cells transiently transfected with LFNG-HA. It is hypothesized that the anti-HA and anti-GM130 signals will colocalize if protein processing is normal, but not if protein processing is diminished. This is because if LFNG is processed appropriately, then it must be cleaved by the Golgi-resident SPC6, in line with previous work and that which is reported in Chapter 2 (Shifley & Cole, 2008; Wengryn et al., 2023). If it is not being processed, then it is not being cleaved by SPC6, and thus should not be in the Golgi. Finally, borrowing from Aims 1 and 3,

it is hypothesized that variants closest to the active site will be processed normally whereas those farther away will be misprocessed.

Like the previous chapter, this work also includes an investigation of evolutionary conservation*, in silico* pathogenicity predictions, and protein modelling to create a more comprehensive approach to assess the genotype-phenotype relationship.

Chapter 2: Functional Characterization of Novel *Lunatic Fringe* Variants in Spondylocostal Dysostosis Type-III with Scoliosis

[1]Parker Wengryn, [1]Karina da Costa Silveira, [1]Connor Oborn, [1]Carrie-Lynn Soltys, [3]Alex Beke, [2,5]Inara Chacon-Fonseca, [2]Nadirah Damseh, [4]Marco Quesada Rodriguez, [4]Ramses Badilla-Porras, [1]Peter Kannu

[1]University of Alberta, Department of Medical Genetics, Edmonton, Canada

[2]University of Toronto, Department of Medical Genetics, Toronto, Canada

[3]University of Alberta, Department of Medicine, Edmonton, Canada

[4]Hospital Nacional de Niños, Medical Genetics and Metabolics, San José, Costa Rica

[5]Lakeridge Health Oshawa, Oshawa, Canada

Peter Kannu is the corresponding author.

2.1. Introduction

Scoliosis, the pathological lateral curvature of the spine, can be a life-threatening condition that significantly impairs quality of life (Hawes & Weinstein, 2003). Complications include chronic pain, restrictive lung disease, and pulmonary hypertension. This cohort also experiences an increased prevalence of depression (Chang et al., 2016; Lin et al., 2019), negative body image (Gallant et al., 2018), and mood disorders. Although incidence varies by demographic, 8% of Americans 40 or older presented with scoliosis as of 2011 (Kebaish et al., 2011). Despite significant research investment, the etiology of scoliosis is largely unknown and termed idiopathic. Studying single-gene disorders associated with scoliosis can provide insights into genetic pathways which contribute to normal spinal development and thus are useful in identifying unexpected yet relevant factors which cause scoliosis.

Central to scoliosis pathobiology are somites. Somites are transient, embryonic spheres that rostrocaudally bud from the paraxial mesoderm in lateral pairs from days 20-35 during human development (Nóbrega et al., 2021; S. Williams et al., 2019). These multipotent mesenchymal progenitor cells differentiate into the sclerotome, which gives rise to the vertebrae and ribs. Somitogenesis is characterized by the symmetrical, rostrocaudal elongation of the embryo during which ‘the Clock and Wavefront’ mechanism acts (Chu et al., 2019; Diaz-Cuadros et al., 2020; Hubaud & Pourquié, 2014; Sanaki-Matsumiya et al., 2022). The ‘Wavefront’ is an area of minimal Fibroblast Growth Factor, Wingless-and-Int-1, and Retinoic Acid signaling that permits the six-hour ‘clock’ gene expression oscillations to differentiate precursor cells into the anterior and posterior somite (Evrard et al., 1998; Johnston et al., 1997; Matsuda et al., 2020; Sanaki-Matsumiya et al., 2022; Yoshioka-Kobayashi et al., 2020; N. Zhang

et al., 2002). As the embryo elongates posteriorly, it forces the midline posteriorly as well, allowing new progenitor cells to be acted upon. Somite symmetry is necessary for vertebral symmetry.

Clock oscillation is chiefly regulated by differential ligand binding to NOTCH1 during canonical NOTCH signaling (Boareto et al., 2015; Kakuda et al., 2020; Nandagopal et al., 2018). LFNG, a Golgi-resident β-1,3-N-acetylglucosaminyltransfease, elongates fucose residues on Epidermal-Growth-Factor (EGF) like repeats of the NOTCH1 Extracellular Domain (NECD). LFNG enhances NOTCH1 heterodimerization with Delta-Like Ligand 1 (DLL1) (Kakuda & Haltiwanger, 2017; Nandagopal et al., 2018) causing pulses of NOTCH-Intracellular Domain (NICD) endocytosis. The pulsatile nature of this process is critical for the expression of genes which encode the anterior somite (Evrard et al., 1998; Johnston et al., 1997; Sanaki-Matsumiya et al., 2022; Yoshioka-Kobayashi et al., 2020; N. Zhang et al., 2002). In contrast, the genes encoding the posterior somite depend on sustained NICD endocytosis, which is mediated through the repression of *LFNG* by HES7 and increased heterodimerization of NOTCH1 with Delta-Like Ligand 4 (DLL4) (Bessho et al., 2001, 2003). Thus, LFNG plays a crucial role in controlling somite polarization and organization.

Spondylocostal Dysostosis Type-III (SCD3) is an autosomal recessive condition characterized by disorganized somites that result in short stature, missing, fused or hemivertebrae, missing and fused ribs, scoliosis, and respiratory failure caused by adult-onset restrictive lung disease (Lefebvre et al., 2018; Otomo et al., 2019; Schuhmann et al., 2021; Sparrow et al., 2006; Takeda et al., 2018) Biallelic pathogenic variants in *LFNG* have been reported in six probands in different reports (Lefebvre et al., 2018; Otomo et al., 2019;

Schuhmann et al., 2021; Sparrow et al., 2006; Takeda et al., 2018) (RefSeq NM_001040167.2), three of which whose variants have been characterized *in vitro*. Of the six described probands, three were identified with homozygous variants and the rest were compound heterozygotes. The c.564C>A (p.F188L) pathogenic *LFNG* variant was present in the homozygous state and led to protein mislocalization and impaired enzymatic function (Sparrow et al., 2006). Perturbance of both function and localization decreased NOTCH activation and is suggested to have impaired oscillatory signaling during somitogenesis. Takeda et al. (2018) described compound heterozygous *LFNG* variants (c.467T>G [p.L156R], c.856T>G [p.R286W]) which caused diminished glycosyltransferase activity. This lack of activity was also suggested to have caused aberrant somitogenesis. The final SCD3 case characterized at a molecular level described compound heterozygous *LFNG* variants (c.372delG [p.K124Nfs*21], c.601G>A [p.D201N]) (Otomo et al., 2019). The frameshift variant was assumed to cause nonsense-mediated mRNA decay whereas the missense variant was found to diminish enzymatic function. Both variants would have prevented *LFNG* from regulating NOTCH signalling during somitogenesis. Both missense and frameshift *LFNG* variants caused SCD3 by inhibiting N-acetylglucosamine transfer to fucose on the NECD of NOTCH1.

Here, we describe the phenotype and characterization of two novel *LFNG* variants (c.766G>A [p.G256S]; c.521G>A [p.R174H]) from a proband with an SCD3 phenotype. The variants are located *in trans* and have been classified as variants of uncertain significance (VUS). To support their role in causing the SCD3 phenotype, we present *in vitro* data assessing glycosyltransferase activity, protein localization, and pre-pro-protein processing. Our analysis led to the identification of the first hypomorphic *LFNG* variant c.521G>A (p.R174H) associated

with disease, and characterization of the first compound heterozygote with both loss of function c.766G>A (p.G256S) and hypomorphic c.521G>A (p.R174H) variants. These findings provide new insights into the pathogenesis of SCD3 and highlight the importance of *LFNG* in somite development.

2.2. Materials and Methods

2.2.1. Exome sequencing

Research-based exome sequencing was conducted with blood samples collected from the proband and parents. Sequencing was undertaken at the SickKids Hospital Center for Applied Genomics on the Illumina platform using the Hiseq4000 sequencer (100 bp PE, 6G/sample) and SureSelect V6 library prep kit (Agilent, USA) with 100 mean coverage. The reads were mapped to hg38 and genotypes were called using GATK 4. *Vcf* files were analyzed using the Franklin website (https://franklin.genoox.com/clinical-db/home). Filtering and prioritization of variants (indels, nonsense, missense, and splice variants) were conducted considering variants with frequencies lower than 1% in gnomAD (Genome Aggregation Database — http://gnomad.broadinstitute.org/) and EVS (Exome Variant Server— http://evs.gs.washington.edu/EVS/).

2.2.2. *In Silico* Variant Analysis

Variant pathogenicity was assessed under the ACMG classification system (Richards et al., 2015) with the web tools SIFT (Sorting Intolerant From Tolerant, (https://sift.bii.a-star.edu.sg/), Polyphen2 (http://genetics.bwh.harvard.edu/pph2/), Align-GVGD (http://agvgd.hci.utah.edu/agvgd_input.php), CADD (Combined Annotation Dependent Depletion, https://cadd.gs.washington.edu/info), and REVEL (Rare Exome Variant Ensemble Learner, https://sites.google.com/site/revelgenomics/).

2.2.3. *In Silico* Protein Modelling

The freely available online workflow ColabFold (Mirdita et al., 2022) was used to generate AlphaFold2 (Jumper et al., 2021) derived structural models of LFNG wildtype, LFNG p.G256S, and LFNG p.R174H. One FASTA sequence for each variant was folded *ab initio* without any templates or relaxation steps. MMseqs2 was used for multiple sequence alignment. Five versions of each structure were folded and ranked by overall pLDDT scores, then predicted finalized structures were visualized in the PyMOL Molecular Graphics System, Version 2.0 (Schrödinger, LLC).

2.2.4. Cloning and Subcloning

Full-length *hLFNG* was amplified from *hLunatic Fringe VersaClone cDNA Shuttle Vector* (R&D, RDC1570; RefSeq NM_001040167.2 [clone NP_001035257]) through PCR (F: 5'-ATGCTCAAGCGTTGTGGACGAC-3; R: 5-GAAGATGGCAGTGCGGGGAC-3') and identity was confirmed with Sanger Sequencing (University of Alberta Molecular Biology Core). The amplicon was gel extracted (Qiaex II Gel Extraction Kit, Qiagen, Cat: 20021) following manufacturers protocol and subcloned into *pCR2.1-TOPO* (Thermo Fisher, K450002). Mach1 T1 Phage-Resistant Chemically Competent E. *coli* (Fisher Scientific, C86203) were used for transformation following the manufacturer's protocol without 5-bromo-4-chloro-3-indolyl-beta-D-galacto-pyranoside (XGAL) due to LFNG toxicity. Plasmid DNA was extracted through miniprep (Wizard Plus SV Miniprep DNA Purification System, Promega, A1460) and then digested with XhoI (Anza, Invitrogen) to allow for screening of insert directionality.

pCR2.1-TOPO-hLFNG and *p3XFLAG-myc-CMV24* (Sigma) were digested with HindIII (Anza, Invitrogen) and XbaI (Anza, Invitrogen), then gel purified. The backbone of *p3XFLAG-myc-CMV24* and *hLFNG* (with short 5' and 3' *pCR2.1-TOPO* sequences) were ligated with T4 DNA ligase (Hereafter referred to as, *LFNG-myc*) (Rapid DNA Ligation Kit, K1422) and then used to transform Mach1 E. *coli*. Plasmid DNA was extracted through miniprep and confirmed with Sanger Sequencing.

LFNG-myc and *3XFLAG-LFNG* (a kind gift from Dr. Shuji Mizumoto) (Takeda et al., 2018) underwent site-directed mutagenesis to obtain the c.564C>A (F: 5'-CGTGGAGTATGACCGCTTAATCGAGTCCGGCA-3'; R: 5'-TGCCGGACTCGATTAAGCG GTCATACTCCACG-3'), c.766G>A (F: 5'-GTTTGCCACGGGCAGCGCTGGCTTCTG-3'; R: CAGAAGCCAGCGCTGCCCGTGGCAAAC-3'), and c.521G>A (F: 5'-CGCCCACAGCCACCAGGCGCTGT-3'; R: 5'-ACAGCGCCTGGT GGCTGTGGGCG-3') (QuikChange XL Site-Directed Mutagenesis Kit, Agilent, Cat: 200516-5) variants. We then transformed Mach1 E. *coli* and plasmid DNA was extracted through miniprep and confirmed with Sanger Sequencing.

Wildtype and variant *LFNG-myc* constructs were in-fusion cloned (In-Fusion Snap Assembly Master Mix with Competent Cells, Takara, 638952; a gift from Dr. Serhiy Havralov and Dr. Ordan Lehmann) into the *PCDNA3.1* backbone of *TRPML1-HA* (Venkatachalam et al., 2006) (Gifted by Craig Montell, Addgene plasmid #18825; n2t.net/addgene:18825; RRID: Addgene 18825) as per manufacturers instructions with Mach1 E. *coli* transformation. Plasmid DNA was extracted through miniprep and midiprep (ZymoPURE II Plasmid Midiprep Kit, Zymo, Cat: D4201) following manufacturers' protocols and confirmed with Sanger Sequencing.

2.2.5. Cell Culture

Human embryonic kidney 293T (HEK293T) cells (ATCC, CRL-11268) were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Gibco, 11995-065) with 10% Fetal Bovine Serum (Sigma, F1051) and 1% Pen/Strep + Glutamine (Gibco, 10378-016). NIH-3T3 cells (ATCC, CRL-1658) were cultured in DMEM with 10% Calf Serum (ATCC, 302030) and 1% Penicillin-Streptomycin-Glutamine (Gibco, 10378-016). Cells were regularly tested for mycoplasma contamination via PCR.

2.2.6. Western Blotting

$2.5x10^5$ NIH-3T3 cells were plated on 6-well plates for 24 hours then transiently transfected with each of the *LFNG-HA* constructs using Lipofectamine LTX (Thermo Fisher, 15338100), following the manufacturer's instructions. After 48 hours, cells were scraped with 1% NP40 lysis buffer (20mM Tris pH 7.4, 5 mM EDTA, 10 mM $Na_4P_2O_7$, 100mM NaF), phosphatase inhibitor (Millipore, 524628), protease inhibitor (Sigma, P8340), and sodium orthovanadate. The cells were then sheared with a 26½ gauge needle and cleared by centrifugation at 1200 x G for 30 minutes. Each sample was quantified via bicinchoninic acid assay (BCA) and unadjusted protein samples were flash-frozen with liquid nitrogen or dry ice and placed at -80°C for long-term storage. Concentration-adjusted samples were run on an SDS gel, transferred to a nitrocellulose membrane, and blocked in a 5% Skim-Milk solution. Blots rotated overnight in primary antibody at 4C (Mouse anti-FLAG-M2 [Sigma F1804], 1:1000; Mouse anti-β-Actin [ABCAM AB6276, AC-15], 1:1000; Rabbit anti-HA [Cell Signalling C29F4], 1:1000.) The next morning, blots were incubated in either HRP-linked Goat anti-Mouse

IgG (Cell Signalling 7076S, 1:1000) or HRP-linked Goat anti-Rabbit IgG (Cell Signalling, 7074S) secondary antibody for one hour. Chemiluminescence was achieved through the Western Lighting ECL Plus kit (Perkin Elmer) on a BioRad Chemidoc. Densitometry was conducted in Image Lab 6.1 (BioRad) and exported to Microsoft Excel for aggregation.

2.2.7. Functional Assay

Functional analysis was conducted as previously reported with *3XFLAG-hLFNG* (Takeda et al., 2018). The protocol and plasmid were both generous gifts from Dr. Shuji Mizumoto and the methods description partly reproduces the wording. 2.0 x 10^6 HEK293T cells were plated in 10cm dishes coated with poly-l-lysine (Sigma, P4707) for 22 hours, then transiently transfected with each of the *3XFLAG-LFNG* constructs using Lipofectamine 3000 (Thermo Fisher, L3000001). After 68 hours, the media was incubated with 90µL of anti-FLAG M2 Affinity Gel (Sigma, A2220) for two hours at 4°C. 10µL of FLAG-bound resin was then incubated with 40µL of reaction mixture (50 mM 2-(N-morpholino) ethanesulfonic acid-NaOH (pH 6.5)), 10 mM $MnCl_2$, 0.1 mM UDP-GlcNAc (Promega, V7071), and 1 mM p-nitrophenyl-α-L-fucose (Sigma, N3628)) for 2 hours at 37°C. After returning to room temperature, 25µL of supernatant was mixed with 25µL of UDP-Detection Mix as per manufacturer's instructions for one hour at room temperature. Luminescence was quantified with a TD-20/20 Luminometer (Turner Systems) in 1.5mL Eppendorf tubes (Fisher Scientific) and compared to a standard curve of UDP-concentration/luminescent intensity.

2.2.8. Immunofluorescence Microscopy

1×10^5 NIH-3T3 cells were seeded on glass coverslips coated with poly-l-lysine in 6-well plates for 24 hours. The next day, media was replaced, and cells were transfected with each of the *LFNG-HA* plasmid constructs using Lipofectamine LTX Reagent following the manufacturer's instructions. After 24 hours, media was removed, and cells were fixed with 4% paraformaldehyde at room temperature for 15 minutes, then permeabilized with 0.1% Triton X-100 (Thermo Fisher, A16046.0F) for five minutes at room temperature. Samples were then blocked with 3% bovine serum albumin (BSA; Sigma, A9418-5G) for one hour, washed with phosphate buffered saline (PBS) 3X, and incubated with Rabbit anti-HA (Cell Signalling C29F4, 1:800) and Mouse anti-GM130 (BD 610623, 1:800) primary antibodies for two hours. Cells were washed with PBS 3X and incubated with secondary antibodies for one hour (Goat anti-mouse Alexa Fluor 594 [Thermo Fisher A11005, 1:500]; goat anti-rabbit Alexa Fluor 488 [Thermo Fisher A11034, 1:500]). Cells were washed again with PBS 3X, stained with Hoechst 33342 at 1µg/ml (Molecular Probes, #H-3570) for 15 minutes, washed with PBS 3X, and mounted on slides with ProLong Gold Antifade Mountant (Invitrogen, P36930). Fluorescence microscopy was conducted at 60X in the Spinning Disk Confocal Microscope (Quorum Technologies) at the Cell Imaging Core at the University of Alberta. Images were analyzed and processed with FIJI software.

2.2.9. Statistical Analysis

To compare three or more groups at once, a One-Way ANOVA was performed. Test assumptions were verified before test performance to maintain the internal validity of the results. Upon receiving a significant output from the One-Way ANOVA ($p < 0.05$), Bonferroni post-hoc

statistical analyses were conducted. Statistical tests were performed in Microsoft Excel and figures were created in GraphPad Prism version 8.0.0 for Windows.

2.3. Results

2.3.1. Clinical History

The proband was born by caesarian section after an uncomplicated pregnancy at 41 weeks of gestation. He did not require resuscitation and his APGAR scores were 9^{1} and 9^{5}. His head circumference and weight were at or above the 50^{th} percentile and his length fell between the 10^{th} to 25^{th} percentile. He had a wide anterior and posterior fontanelle with a separated sagittal suture, abnormal spinal curvature with short neck, congenital torticollis, and a wide and short thorax. A thoracic spina bifida was present. Cardiovascular and respiratory examinations were normal. Investigations, including head and abdominal ultrasound, were also normal. Spinal computed tomography scans revealed large vertebral segmentation anomalies with rib malalignment and fusions (Figure 8A, 8B). At one year and eight months, spinal radiographs showed vertebral segmental defects and rib fusions (9 right ribs and 10 left) (Figure 8C, 8D). He had a left 38-degree scoliosis at L1/S1, 44-degree lordosis at L1/S1, and 32-degree kyphosis at T2/T12. On his last evaluation at two years and four months, his height was between the 10^{th} to 25^{th} percentile, and the arm span-to-height ratio was 1.075. He had an asymmetric thorax with pectus carinatum but no major respiratory issues. There was no limitation of neck and limb movement, and he was developmentally appropriate.

His parents were a healthy (mother aged 35 years and father aged 39 years) and non-consanguineous couple of Costa Rican background. His father had two healthy children with a

prev[illegible]. There was n[illegible]ry of sig[illegible]e. Parental hei[illegible]nd

150 [illegible]r and [illegible]ely.

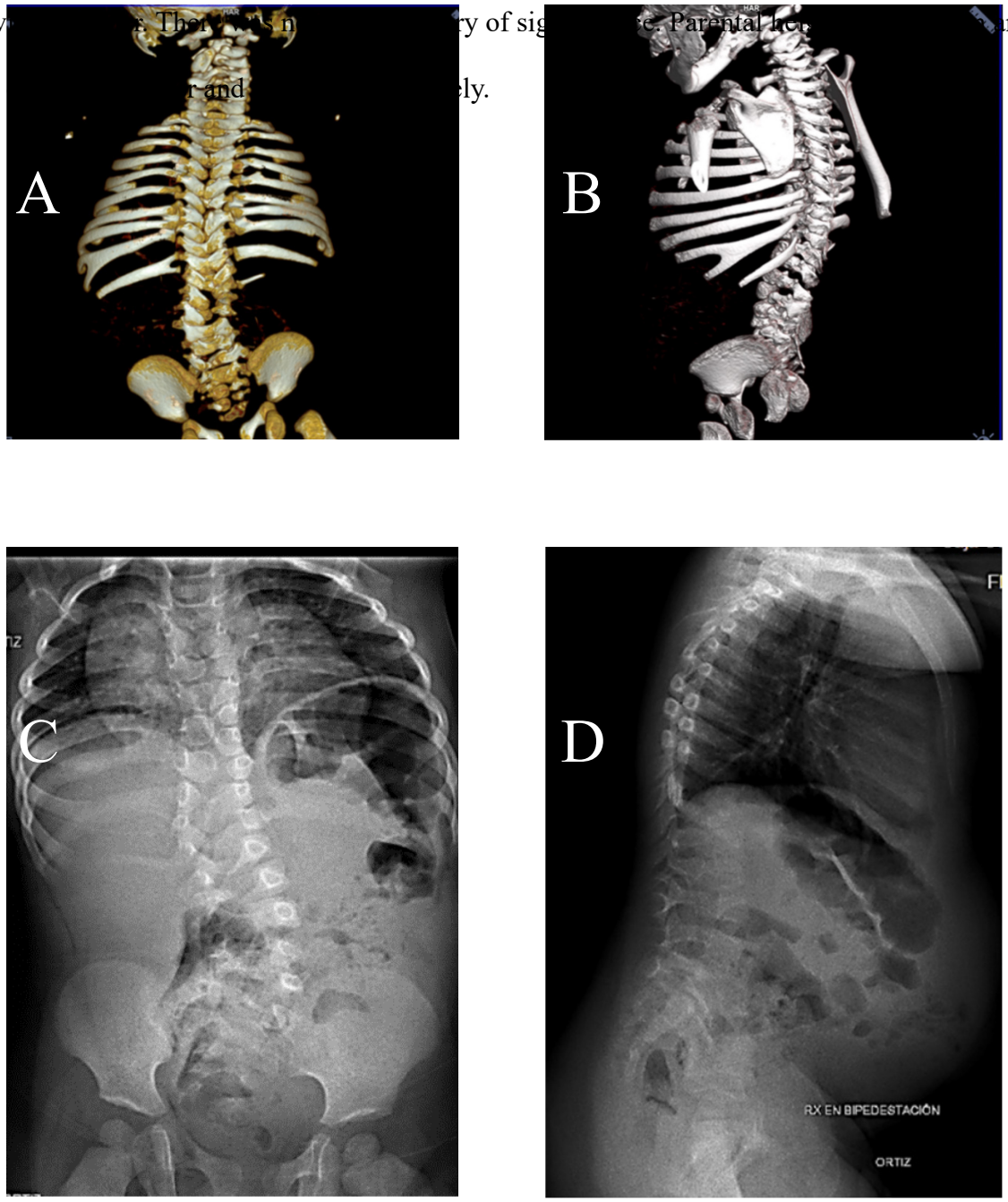


Figure 8: Computed Tomography and X-Ray Imaging of the SCD3 Proband. A) Posterior CT image at three months. Note the disorganized vertebrae in the cervical, lumbar, and sacral spine and the right-thoracic-hemivertebrae. There are 10 left and 9 right asymmetrically set ribs

with posterior proximal fusion and posterior distal fission. **B)** 45-degree clockwise posterior CT image at three months. Note posterior rib fusion and fission as well as asymmetrical vertebral shape. **C)** Posterior X-Ray at one year eight months. The posterior spine is characterized by the "Pebble Beach Sign", angulated vertebrae, and scoliosis. **D)** Lateral X-Ray at one year and eight months. Note the disorganized vertebral bodies in the thoracic and lumbar spine.

2.3.2. Genetic and Proteomic Analysis of Novel *LFNG* Variants

Whole exome sequencing revealed two *LFNG* variants c.521G>A (p.R174H) and c.766G>A (p.G256S) (RefSeq NM_001040167.2). These two variants were novel and *in trans* (see Table 1). The first missense change, located in exon 3 of *LFNG*, was paternally inherited. This change replaced arginine with histidine at codon 174 of the LFNG protein (p.R174H) and was rare in population databases (7-2525258-G-A, gnomAD v3.1.2, 6.57×10^{-6}). The second missense variant was maternally inherited and in exon 5. The glycine at codon 256 was replaced with serine (p.G256S). This variant was also rare in population databases (7-2525715-G-A, gnomAD v3.1.2, 6.57×10^{-6}).

Both variants were analyzed in SIFT, PolyPhen-2, Align-GVGD, CADD, and REVEL to generate *in silico* estimations of possible protein disruption (Table 1). SIFT determines outcomes of substitutions by sequence homology and physicochemical properties where arbitrary outcome scores of <0.05 are predicted to be deleterious (Ng & Henikoff, 2001). The p.R174H substitution was predicted to be tolerated with a score of 0.22, whereas the p.G256S substitution was predicted to be deleterious with a score of 0.00. Using PolyPhen-2, we assessed the effect of amino acid substitutions by analyzing sequence and structural homology as well as existing Pfam annotations. Scores are given from 0.0 - 1.0 as increasing probabilities that a variant is deleterious (Adzhubei et al., 2013). The p.R174H and p.G256S substitutions were given scores ranging from 0.746-0.970 and 1.00-1.00, respectively. Align-GVGD was then employed to

assess the effect of each substitution in the context of proteomic sequence alignment and biophysical amino acid properties (Tavtigian et al., 2006). Scores are classed into one of seven categories from Class C0 (Benign) to Class C65 (Pathogenic). The p.R174H substitution was categorized as Class C25 (GV 0, GD 28.2), whereas the p.G256S substitution was categorized as Class C55 (GV 0, GD 55.3). CADD was utilized to determine the relative genome-wide deleteriousness of each variant (Rentzsch et al., 2019). Both p.R174H and p.G256S scored 25.6, indicating they were within the top 0.4% of deleterious human genome variants. Finally, we used REVEL, a program which aggregates 13 variant-prediction algorithms into one score which is scaled from 0.0 (benign) to 1.0 (pathogenic) (Ioannidis et al., 2016). p.R174H and p.G256S received scores of 0.515 and 0.889, respectively. Both c.521G>A (p.R174H) (PM2, PP2, PP4) and c.766G>A (p.G256S) (PM2, PP2, PP3, PP4) fell outside of ACMG variant classification guidelines and thus were classified as VUS (Richards et al., 2015).

To qualitatively examine whether these variants cause structural perturbation *in silico*, we generated AlphaFold2-based structural models of LFNG WT, p.R174H, and p.G256S and aligned them in PyMOL2 (Figure 9). No structural changes were identified as both p.R174H and p.G256S LFNG substitutions maximally overlapped with WT LFNG.

Variant	c.766G>A	c.521G>A
DNA change (GRCh38)	NC_000007.14:g.2525715G>A	NC_000007.14:g.2525258G>A
ClinVar ID	1003507	1062453
SNPdb ID	rs1437427476	N/A
LOVD DB ID	LFNG_000035	LFNG_000034
gnomAD v3.1.2	6.57×10^{-6}	6.57×10^{-6}
Protein	p.G256S	p.R174H
SIFT	Likely Deleterious	Tolerated
PolyPhen 2	Damaging	Possibly Damaging
AlignGVGD	Class-C55	Class-C25
CADD	25.6 (3.67*)	25.6 (3.66*)
REVEL	0.889	0.515

Table 1: Novel *LFNG* Variants Associated with the SCD3 Proband. *LFNG* transcript- RefSeq NM_001040167.2 (MANE select) (Clone NP_001035257). * CADD raw score.

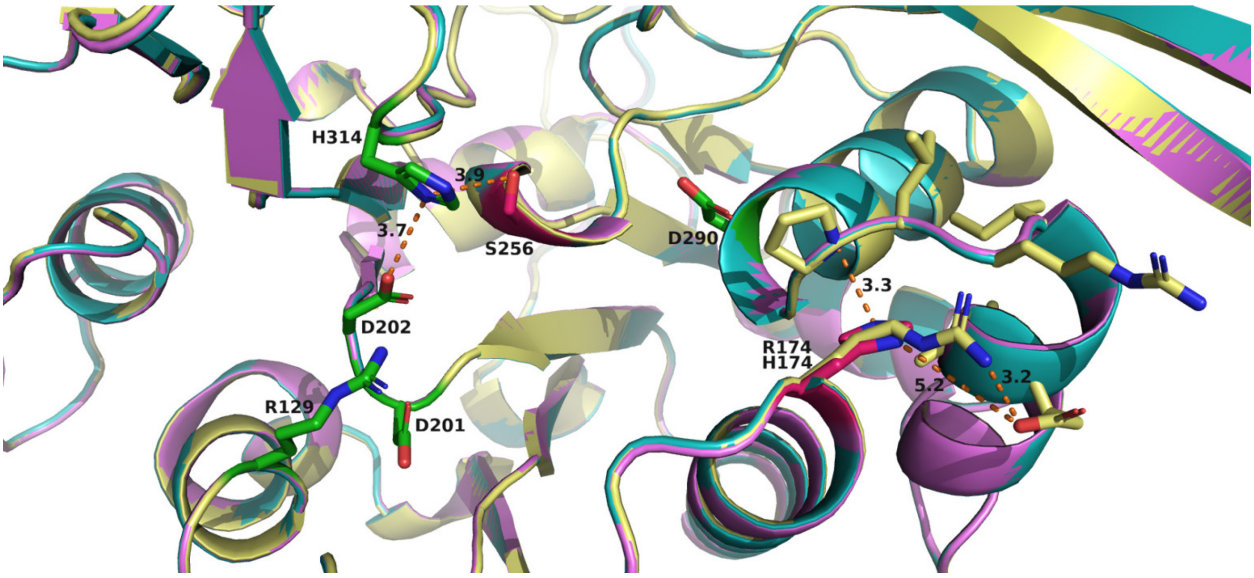


Figure 9: AlphaFold2 models of wildtype, p.R174, p.G256S LFNG enzymes with highlighted residues. Wildtype residues are labeled yellow, p.R174 is labeled purple, and p.G256S is labeled turquoise. Important active site residues have been identified with a green color, p.H314 and p.D202 chelate Mn^{2+}, p.R129 and p.D201 are responsible for substrate binding, and finally, D290 is the catalytic residue. Both substituted residues are highlighted with a bright pink color at 174 and 256. Position 174 demonstrates canonical arginine positioning overlaid with the histidine substitution. Measurements are taken in angstroms to show relative positioning in a 3D space and important nearby residues.

2.3.3. The Novel *LFNG* Alleles are Functionally Hypomorphic and Null

To investigate the effect of each *LFNG* variant *in vitro*, glycosyltransferase activity, protein processing, and intracellular localization were assessed. Previous work suggests that these factors contribute to SCD-III manifestation (Otomo et al., 2019; Shifley & Cole, 2008; Sparrow et al., 2006; Takeda et al., 2018; D. R. Williams et al., 2016). Glycosyltransferase activity was assessed using previously reported methods (Takeda et al., 2018). Briefly, HEK293T cells were transiently transfected with a shortened version of *LFNG* cDNA fused to a 3' pre-pro-trypsin leader sequence (*3XFLAG-hLFNG*). Western blotting ensured similar secretion between conditions (Figure 10A). The WT, p.G256S, and p.R174H conditions appeared similar in band intensity whereas the functionally inactive variant p.F188L (Sparrow et al., 2006) was not secreted into media. We used the empty vector (EV) backbone as the negative control in line with the work of Takeda et al., (2018).

The glycosyltransferase reaction was undertaken by incubating 3XFLAG-LFNG bound anti-FLAG agarose resin with a reaction mixture containing UDP-GlcNAc (donor) and p-nitrophenyl-α-L-fucose (acceptor). Here, one UDP molecule was released for each GlcNAc transfer, and this UDP was quantified via luminescence post-reaction. Quantitative analyses indicated that WT luminescence was significantly more intense than p.R174H ($p = 6.5 \times 10^{-3}$), p.G256S ($p = 2.5 \times 10^{-3}$), and E.V. ($p = 2.5 \times 10^{-3}$) (Figure 9B). The data also indicated that p.R174H luminescence was significantly more intense than both p.G256S ($p = 7.5 \times 10^{-6}$) and EV ($p = 7.5 \times 10^{-6}$). Ultimately, the data indicated that p.R174H Glc-NAc transferase activity lied between WT and inactive levels, and it was the first observation of a patient-derived LFNG substitution causing hypomorphic function.

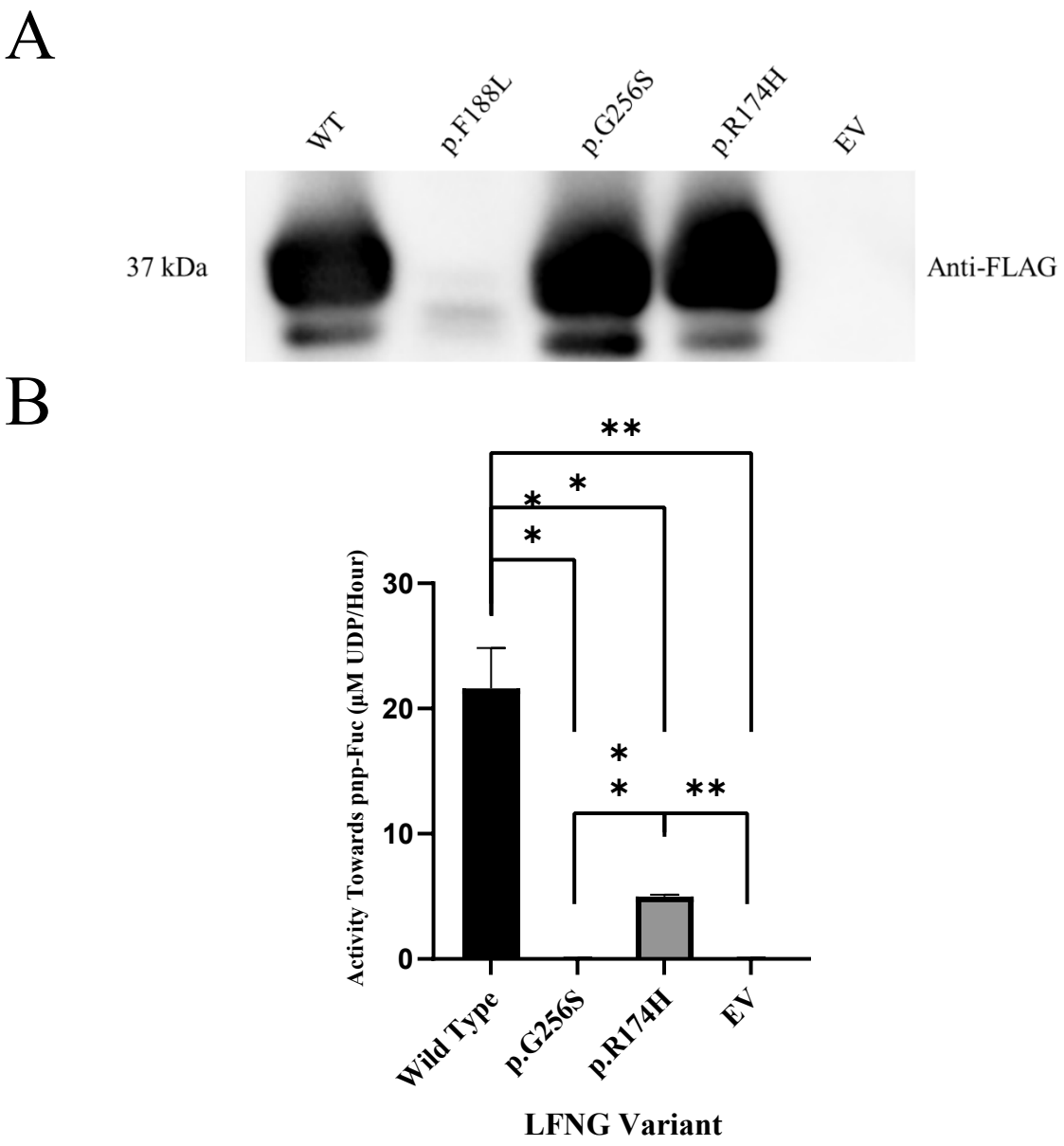


Figure 10: The p.G256S and p.R174H LFNG Substitutions are Null and Hypomorphic, Respectively. A) Western blot of anti-FLAG immunoprecipitated 3XFLAG-LFNG protein. 3XFLAG-LFNG was purified with anti-FLAG agarose resin, disassociated with SDS-buffer, and 20μL of each sample was run on a 10% SDS gel. Protein was detected with anti-FLAG primary antibody and HRP-linked anti-Mouse IgG secondary. **B)** Glc-NAc transferase activity of LFNG variant protein isolated from media. One-way ANOVA indicated a significant difference between groups ($F(3, 8) = 40.6$, $p = 3.5 \times 10^{-5}$) and Bonferroni-adjusted ($\alpha = 8.3 \times 10^{-3}$) post-hoc analyses indicated that WT was significantly different from p.R174H, p.G256S, and E.V. p.R174H was significantly more intense than p.G256S and EV. There was no statistically significant difference between p.G256S and EV ($p = 6.8 \times 10^{-1}$). Values are means +- S.E of three independent experiments (N = 3) plated in triplicate. *$p < 8.3 \times 10^{-3}$ **$p < 2.5 \times 10^{-3}$.

2.3.4. Protein Processing Appears Normal with p.G256S and p.R174H LFNG

With the identification of a partially active *LFNG* variant, it was necessary to rule-out that aberrant pre-pro-processing or protein mislocalization could cause the phenotype. LFNG contains an N-terminal Type-II transmembrane domain that is cleaved at (K/R)XX(K/R) sites by Subtilisin-like Proprotein Convertase family (SPC) proteins, specifically SPC6 (Shifley & Cole, 2008; D. R. Williams et al., 2016). LFNG possesses two cleavable sites within its 86-residue transmembrane domain, RGRR (37 to 40) and RARR (83 to 86). Pre-pro-processing leads to LFNG exocytosis and allows for fine-tuning of clock cycle oscillations (Shifley & Cole, 2008; D. R. Williams et al., 2016). Concordantly, these works suggest that aberrant cleavage impairs somitogenesis and causes SCD3 phenotype.

To qualitatively and semi-quantitatively assess each variant's effect on protein processing, intracellular protein lysate from NIH-3T3 cells transiently transfected with 3' *HA*-fused *LFNG* expression plasmids (*LFNG-HA*) were western blotted. The predicted molecular weights of LFNG-HA were 42.8 kDa (pre-pro-LFNG), 38.4 kDa (pro-LFNG), and 34.1 kDa (processed LFNG). The previously characterized p.F188L LFNG substitution (Sparrow et al., 2006) was also assessed to test whether protein processing and localization were linked. Qualitatively, both p.F188L and p.G256S exhibited increased pre-pro-protein band intensity compared to WT (Figure 11A). However, densitometric analyses indicated that only p.F188L was significantly different from WT ($p = 9.0 \times 10^{-3}$) (Figure 11B). There were no bands in the pre-processed 38.4 kDa range for any of the conditions (Figure 11A). Finally, WT, p.G256S, and p.R174H had faint bands at the processed-protein weight of 34.1 kDa (Figure 11A). Qualitatively, p.F188L lacked this signal whereas p.G256S was more intense (Figure 11A).

Statistical analysis indicated a significant difference in band intensity between WT and p.F188L ($p = 9.0 \times 10^{-9}$) (Figure 11C).

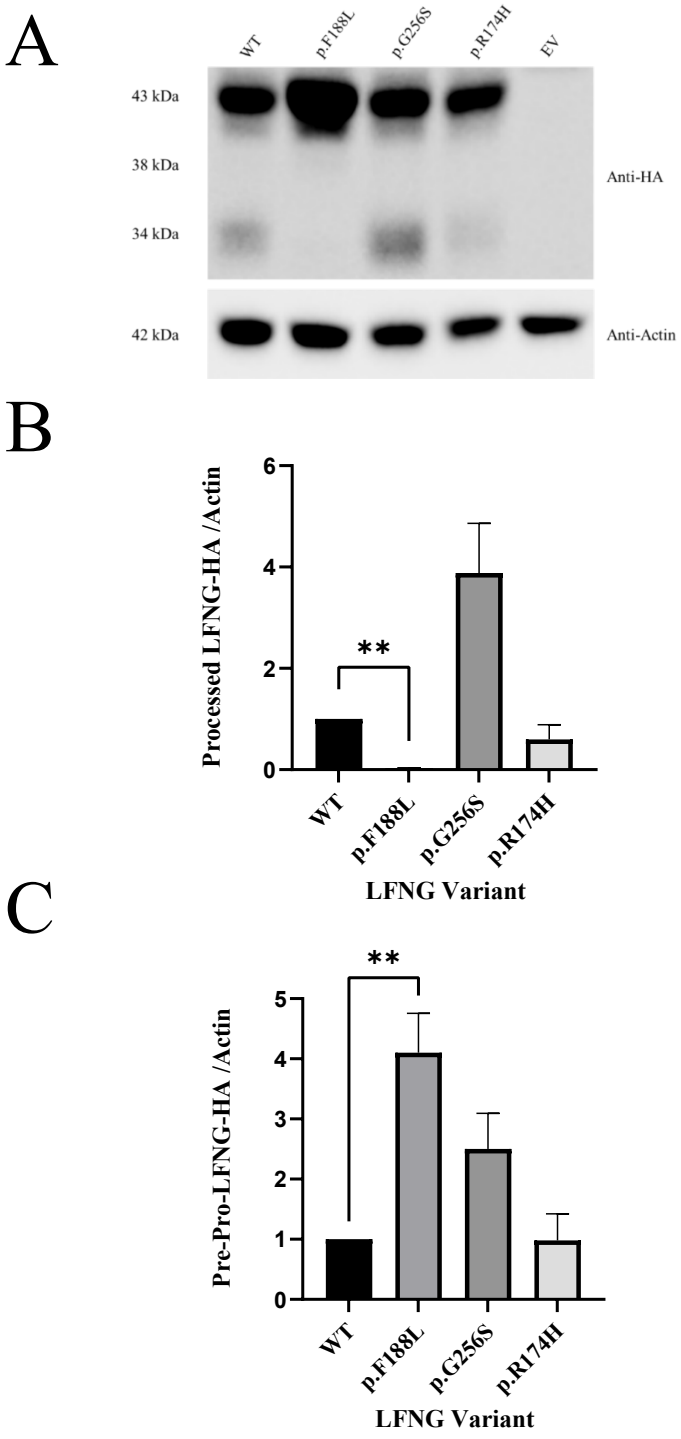


Figure 11: The p.G256S and p.R174H LFNG Substitutions Do Not Lead to Aberrant Protein Processing. A) Western blot of cell-lysate from LFNG-HA transiently transfected NIH-3T3 cells. Cell lysates were run on a 12% SDS gel and membranes were incubated with anti-HA primary then anti-Rabbit IgG secondary antibodies, or anti-Actin primary then anti-Mouse IgG secondary antibodies. **B, C)** Densitometry analysis of pre-pro-LFNG-HA (B) or processed (C) LFNG-HA bands. Band intensities were normalized to actin. One-Way ANOVA indicated a statistically significant difference between all conditions in both pre-pro-LFNG-HA ($F(3, 8) = 9.1$, $p = 5.9 \times 10^{-3}$) and processed LFNG-HA ($F(3, 8) = 21.2$, $p = 4.4 \times 10^{-5}$). Bonferroni-adjusted ($\alpha = 1.3 \times 10^{-2}$) post-hoc analyses indicated the pre-pro-LFNG-HA p.F188L bands were significantly more intense than WT, and the processed LFNG-HA p.F188L bands were significantly less intense than WT. No other conditions were significantly different from WT. Values are means +- S.E of three independent experiments (N = 3) plated in triplicate. $^{**}p < 1.3 \times 10^{-2}$.

2.3.5. The p.G256S and p.R147H LFNG Substitutions Do Not Affect Subcellular Localization

Finally, we asked whether the variants could cause mislocalization of LFNG (Moloney et al., 2000; Sparrow et al., 2006). NIH-3T3 cells were transiently transfected with the *LFNG-HA* constructs. p.F188L was employed as a negative control as it is the only substitution previously shown to cause LFNG mislocalization (Sparrow et al., 2006). Qualitative analysis indicated anti-HA and anti-GM130 signal colocalization in the WT, p.G256S, and p.R174H conditions (Figure 12A, 12C, 12D). The anti-HA and anti-GM130 signals did not colocalize in the p.F188L condition, supporting previous results (Figure 12B) (Sparrow et al., 2006).

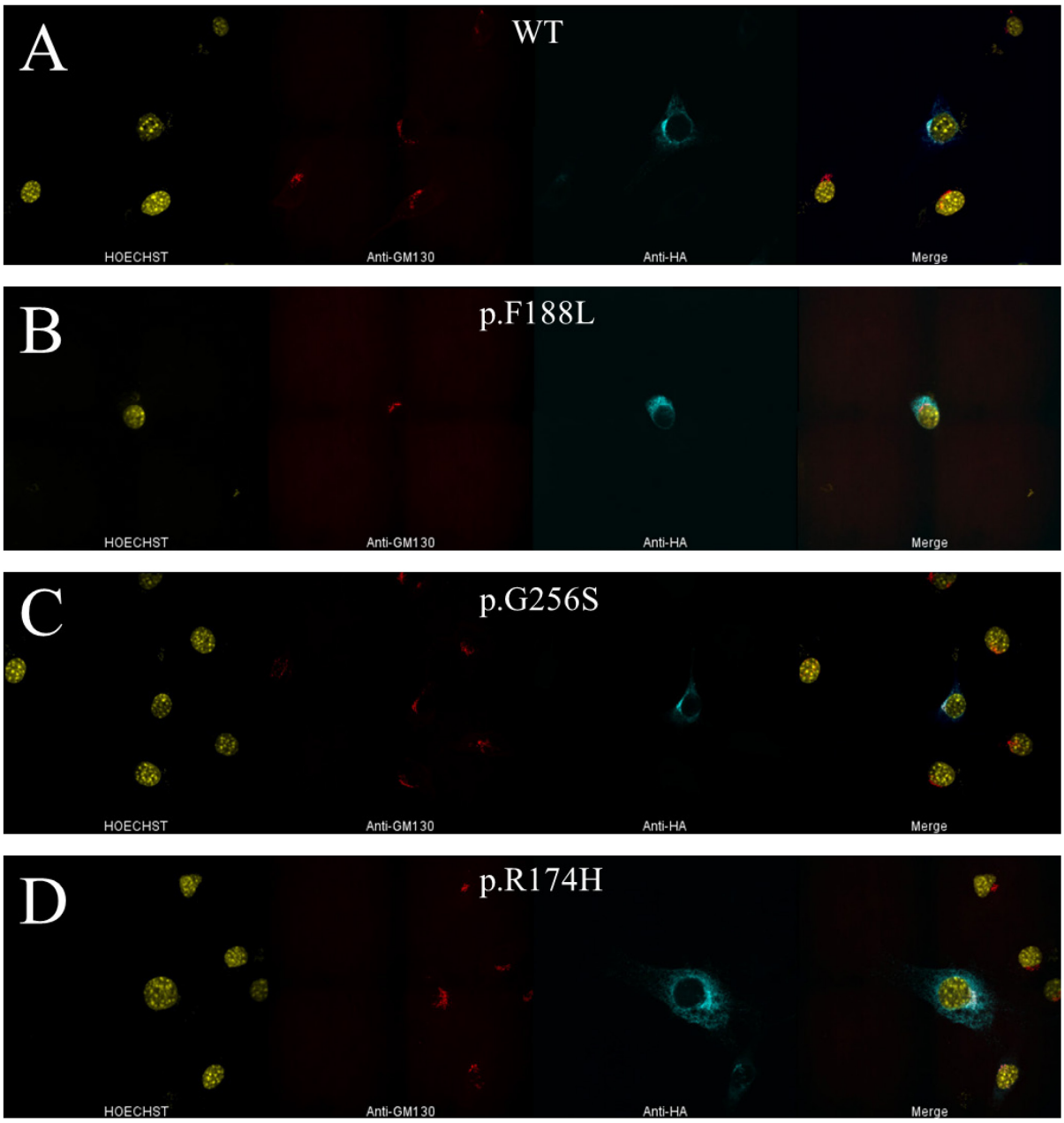


Figure 12: The p.G256S and p.R174H LFNG Substitutions Localize to the Golgi. WT **(A),** p.F188L **(B)**, p.G256S **(C)**, and p.R174H **(D)** LFNG-HA variants transiently transfected NIH-3T3 cells. Cells were incubated with HOECHST, rabbit anti-HA, and mouse anti-GM130 primary antibodies then again with fluorescent secondary antibodies. The fluorescent signal was artificially altered to be colour-blind friendly. Note the qualitative absence of signal overlap in B). Images are representative samples of three independent experiments (N = 3). Scale bar, 30μm.

2.4. Discussion

LFNG is of interest due to its critical role in somitogenesis and the pronounced, yet varied presentation caused by pathogenic variants. Here, we identified a proband with SCD3 phenotype and two *LFNG* VUS according to the ACMG classification system. Since SCD3 is an autosomal recessive disorder, the recurrence risk is 25%. However, VUS cannot be used for clinical decision-making to inform reproductive risk management. Our work therefore aimed to determine whether the p.G256S and p.R174H LFNG substitutions were pathogenic. Our functional data indicate that the p.G256S and p.R174H substitutions are null and hypomorphic, respectively. This data acts as PS3 (functional) ACMG evidence for each variant, elevating their status to likely pathogenic (Richards et al., 2015). Therefore, our data support the hypothesis that these *LFNG* variants cause the SCD3 phenotype enabling the parents to access prenatal genetic testing or preimplantation genetic diagnosis in a subsequent pregnancy.

2.4.1. *In silico* Mechanisms of Functional Perturbance

The biochemical mechanism of p.R174H hypomorphism appears to be reminiscent of a Manic Fringe (MFNG)-like substitution. Specifically, the PolyPhen-2 multiple sequence alignment of Fringe enzymes showed that p.R174H substitutes a Lunatic-specific conserved residue (arginine) for a Manic-specific conserved residue (histidine) (Supplementary Figure 1). MFNG is significantly less active than LFNG (Moloney et al., 2000), and the functional data support that p.R174H follows this trend (Figure 10B). *In vitro*, it would be interesting to test if: 1) histidine interacts with similar residues in both cases, 2) if there are similar structural features of both enzymes, and 3) if these interactions are the cause of lowered activity. These tests could

aid in identifying amino acid changes which potentiate the increased activity of LFNG relative to MFNG and RFNG. From an evolutionary perspective, this may begin to elucidate the varied developmental roles of Fringe family proteins across species.

Structurally, the core of each AlphaFold2 model has maximal overlap with WT (Figure 9). However, these models only account for visual structural changes, not functional ones. For example, although the p.R174H substitution is conservative, it may cause changes in hydrogen bonding due to the Golgi's slightly acidic pH (6.0 to 6.8) (Wu et al., 2000). This is because histidine is deprotonated within this pH range whereas arginine is protonated. Such an effect may account for both maximal structural overlap and varied results across *in silico* functional analyses (Table 1). Contrastingly, although p.G256S is a non-conservative substitution, its location may allow for a WT-like structure (Figure 9) since residue 256 exists in a small bend to the opening of LFNG's active site. When glycine is substituted with serine, the new hydroxyl group may interact with the adjacent p.H314 which facilitates Mn2+ cofactor coordination (Figure 9). The distance between p.H314 and its native binding partner p.D202 is similar to the distance between p.H314 and p.S256 (3.7Å and 3.9Å respectively). Therefore, serine's highly polar hydroxyl group could destabilize p.H314 Mn^{2+} coordination or even sequester it on the 'wrong' side of this residue. This hypothesis would account for structural similarities (Figure 9) and damaging effects of the substitution (Table 1; Figure 10B).

2.4.2. Mislocalization Prevents Protein Processing

We were unable to attribute mislocalization or aberrant pre-pro-processing to disease in this study. Both p.G256S and p.R174H were normally pre-pro-processed and present in the Golgi (Figures 11 and 12). Interestingly, however, the negative control p.F188L was aberrantly

processed and mislocalized (Figures 10 and 11). Previous work only indicated that this variant mislocalizes (Sparrow et al., 2006). Here, we propose that the p.F188L variant is not pre-pro-processed nor secreted because it is mislocalized. Primarily, there is an intracellular accumulation of pre-pro-LFNG-HA (Figures 11A, 11B) and an absence of processed LFNG-HA (Figure 11A, 11C). Furthermore, the p.F188L protein did not colocalize with the Golgi (Figure 12B). This suggests that the variant protein was not translocated to the Golgi, and therefore was not cleaved by the Golgi-resident SPC6. Crucially, it is unlikely that processed LFNG p.F188L is absent due to secretion because 3XFLAG-hLFNG p.F188L was absent from media in our work (Figure 10A) and elsewhere (Sparrow et al., 2006; Takeda et al., 2018). The combination of evidence suggests that p.F188L inhibits the ability of LFNG to translocate to the Golgi thus preventing its processing and secretion. It was previously hypothesized that this residue may play a role in ER export and Golgi transport (Sparrow et al., 2006), and current data supports this hypothesis. Future work should investigate molecular mechanisms behind the unknown role of this region. Pre-pro-processing is a key modifier of clock timing during somitogenesis, and it has been strongly suggested that the perturbance could lead to malformed somites (Shifley & Cole, 2008; D. R. Williams et al., 2016). Therefore, it will also be important to test whether the mislocalization of otherwise functional LFNG causes SCD3 phenotypes in humans.

2.4.3. Partial Rescue of the Probands Phenotype

Our evidence suggests that impaired glycosyltransferase activity alone is responsible for the proband's phenotype. Importantly, the proband in this study has two *LFNG* variants *in trans*; one of these variants is hypomorphic (p.R174H) and the other is null (p.G256S). It is interesting to consider whether partial activity from the hypomorphic variant may have modestly rescued

the proband's phenotype. Such an effect has been documented in murine models of SCD3 which found that increasing *LFNG* dosage from knock-down to WT increases the rate of somitogenesis and the number of vertebrae (D. R. Williams et al., 2014).

In our proband, we counted over 20 vertebral bodies; this is more than what has been previously reported in others with loss of function *LFNG* variants. In these cases, the vertebral number appears to be less than 17 (Otomo et al., 2019; Takeda et al., 2018). In addition to a greater number of vertebral bodies, our proband is also taller (between 10^{th} and 25^{th} percentile) than other reported cases ($<5^{th}$ percentile) (Sparrow et al., 2006; Takeda et al., 2018; Otomo et al., 2019). Parental heights for many of the cases are not reported. Finally, the arm span-to-height ratio in our proband was indicative of a more preserved trunk length (1.075) compared to previous reports (1.203) (Sparrow et al., 2006). In the context of SCD3, arm span-to-height ratios closer to one indicate a lengthened trunk and thus a less pronounced phenotype. Together, these clinical features may highlight a partial rescue of trunk length, potentially due to the p.R174H substitution. However, the very small sample size and lack of certain molecular/phenotypic data prevent any definitive conclusions from being drawn. More evidence will need to be gathered to test whether vertebrae number and trunk length can be recused by hypomorphic alleles in humans.

We did not observe a difference in vertebral body morphology compared to other reported cases with loss of function *LFNG* variants. Primarily, the "Pebble Beach Sign", a manifestation of rounded, offset vertebrae, was present (Figure 8C) (compare to Sparrow et al., 2006; Lefebvre et al., 2018; Otomo et al., 2019; Schuhmann et al., 2021) (Turnpenny et al., 2003). A variety of segmentation defects, hemivertebrae, and rib abnormalities were also

observed (Figure 8). Although we noted milder scoliosis in our proband, we cannot predict if this deformity will remain stable or progress with time (compare with Sparrow et al., 2006; Takeda et al., 2018; Schuhmann et al., 2021) (Kebaish et al., 2011; Kwon et al., 2021). Therefore, with current evidence, it is unreasonable to suggest that segmentation defect severity is affected by the hypomorphic allele. Many more cases will need to be identified and fully molecularly characterized to test if segmentation defects, including scoliosis, are modulated by partially active LFNG enzyme.

2.4.4. Future Considerations in the Context of Variant Characterization

Our work highlights the importance of functional analyses for VUS alongside the difficulty in translating ambiguous functional results to the clinic. Here, we demonstrated that c.766G>A (p.G256S) causes a complete loss of LFNG function whereas c.521G>A (p.R174H) does not (Figure 10). The former substitution could reasonably be associated with pathogenicity when homozygous based on current evidence and previous reports (Otomo et al., 2019; Sparrow et al., 2006; Takeda et al., 2018); however, the latter is more difficult to predict. Our interpretation of the functional analysis, and that of many other genes, is limited by the current understanding of pathomechanisms (Kanavy et al., 2019). Since the enzymatic activity threshold at which LFNG causes SCD3 is not known, the PS3 criteria are not entirely informative for c.521G>A. Therefore, we cannot discount the possibility of a milder SCD3 phenotype associated with this allele in a homozygous state. *LFNG* is not highly constrained (Z = 0.62; gnomAD 2.1.1) and thus missense variants that lead to partial activity are worth further investigation. It would be scientifically and clinically informative to demarcate the threshold of LFNG activity required for proper somitogenesis.

Finally, although PS3 data can be required to inform probands and families of their reproductive risk status if they harbour VUS, these experiments are often costly, time-consuming, and technically difficult. In this study, this was largely due to creating, optimizing, and performing experiments with multiple epitope tags to ensure reliable and viable data. Therefore, we suggest that future research aims to identify high-throughput, reliable functional assays to limit lengthy and expensive variant characterization projects. One example of such an assay is Saturation Prime Editing; a novel, high-throughput assay which can characterize hundreds to thousands of variants simultaneously (Erwood, Bily, et al., 2022). Applying these novel techniques to genes that cause rare diseases could aid in limiting the cost and time commitment of variant characterization projects. Further, such projects would aid in providing informed consent during genetic counseling and reproductive risk assessment.

2.5. Conclusions

In this work, we were able to characterize two novel *LFNG* variants, c.521G>A (p.R174H) and c.766G>A (p.G256S). p.G256S is functionally null whereas p.R174H is functionally hypomorphic. To our knowledge, this is the first hypomorphic *LFNG* allele associated with human SCD3. Our evidence suggests that this hypomorphic variant may have partially rescued the proband vertebral number and trunk length. We conclude that heightened LFNG activity may play a role in human SCD3 phenotypic variance. In the future, we hope to molecularly characterize all *LFNG* variants in the literature simultaneously to study the relationship between gene, protein, and pathological variance.

2.6. Acknowledgements

The authors would first like to thank the family of our patient and SickKids Hospital Centre for Applied Genomics. We would like to thank Dr. Shuji Mizumoto for his invaluable guidance, access to protocols, and material availability which allowed us to gather our functional data. We would also like to thank the Department of Medical Genetics, Department of Biochemistry stores, and Molecular Biology Core at the University of Alberta for their technical and practical support. P.W. would like to thank Dr. Fred Berry and Dr. Toshifumi Yokota for their scientific guidance during the course of the project. Finally, P.W. would like to thank Patrick Wachowich for inspiring him to pursue science in the hope of creating a better world for us all.

2.7. Funding Statement

This work was supported by grants from the Shriners Chair in Scoliosis, the Women and Children's Health Research Institute at the University of Alberta, and the Rare Disease Foundation. This work was also supported by the Alberta Innovates Graduate Student Scholarship, the Alberta Graduate Excellence Scholarship, and the Graduate Entrance Scholarship.

2.8. Ethics Statement

Ethical approval for the publication and radiographs were obtained through the Hospital Nacional de Niños, San José, Costa Rica.

2.9. Conflicts of Interest

R.B.P. declares a conflict of interest as a consultant for Sanofi Pharmaceuticals. None of the other authors declare a conflict of interest.

2.10. Data Availability

The multiple sequence alignment data used to support the findings of this study are included within the supplementary information file. The genetic and phenotypic data used to support the findings of this study have been deposited in the LOVD repository (Individual #00431334). The data were also submitted in ClinVar (Accession numbers VCV001062453.5 and VCV001003507.5) by Invitae.

2.11. Bibliography

Adzhubei, I., Jordan, D. M., & Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Current Protocols in Human Genetics, SUPPL.76.* https://doi.org/10.1002/0471142905.hg0720s76

Bessho, Y., Hirata, H., Masamizu, Y., & Kageyama, R. (2003). Periodic repression by the bHLH factor Hes7 is an essential mechanism for the somite segmentation clock. *Genes and Development, 17*(12), 1451–1456. https://doi.org/10.1101/gad.1092303

Bessho, Y., Miyoshi, G., Sakata, R., & Kageyama, R. (2001). Hes7: a bHLH-type repressor gene regulated by Notch and expressed in the presomitic mesoderm. *Genes to Cells, 6*(2), 175–185.

Boareto, M., Jolly, M. K., Lu, M., Onuchic, J. N., Clementi, C., & Ben-Jacob, E. (2015). Jagged-delta asymmetry in Notch signaling can give rise to a sender/receiver hybrid phenotype. *Proceedings of the National Academy of Sciences of the United States of America, 112*(5), E402–E409. https://doi.org/10.1073/pnas.1416287112

Chang, W. P., Lin, Y., Huang, H. L., Lu, H. F., Wang, S. T., Chi, Y. C., Hung, K. S., & Chen, H. Y. (2016). Scoliosis and the subsequent risk of depression. *Spine, 41*(3), 253–258. https://doi.org/10.1097/BRS.0000000000001187

Chu, L. F., Mamott, D., Ni, Z., Bacher, R., Liu, C., Swanson, S., Kendziorski, C., Stewart, R., & Thomson, J. A. (2019). An In Vitro Human Segmentation Clock Model Derived from Embryonic Stem Cells. *Cell Reports, 28*(9), 2247-2255.e5. https://doi.org/10.1016/j.celrep.2019.07.090

Diaz-Cuadros, M., Wagner, D. E., Budjan, C., Hubaud, A., Tarazona, O. A., Donelly, S., Michaut, A., Al Tanoury, Z., Yoshioka-Kobayashi, K., Niino, Y., Kageyama, R., Miyawaki, A., Touboul, J., & Pourquié, O. (2020). In vitro characterization of the human segmentation clock. *Nature, 580*(7801), 113–118. https://doi.org/10.1038/s41586-019-1885-9

Erwood, S., Bily, T. M. I., Lequyer, J., Yan, J., Gulati, N., Brewer, R. A., Zhou, L., Pelletier, L., Ivakine, E. A., & Cohn, R. D. (2022). Saturation variant interpretation using CRISPR prime editing. *Nature Biotechnology 2022*, 1–11. https://doi.org/10.1038/s41587-021-01201-1

Evrard, Y. A., Lun, Y., Aulehla, A., Gan, L., & Johnson, R. L. (1998). Lunatic fringe is an essential mediator of somite segmentation and patterning. *Nature, 394*(6691), 377–381. https://doi.org/10.1038/28632

Gallant, J. N., Morgan, C. D., Stoklosa, J. B., Gannon, S. R., Shannon, C. N., & Bonfield, C. M. (2018). Psychosocial Difficulties in Adolescent Idiopathic Scoliosis: Body Image, Eating Behaviors, and Mood Disorders. *World Neurosurgery, 116*, 421-432.e1. https://doi.org/10.1016/j.wneu.2018.05.104

Hawes, M. C., & Weinstein, S. L. (2003). Health and Function of Patients with Untreated Idiopathic Scoliosis [2] (multiple letters). *Journal of the American Medical Association, 289*(20), 2644–2645. https://doi.org/10.1001/jama.289.20.2644-a

Hubaud, A., & Pourquié, O. (2014). Signalling dynamics in vertebrate segmentation. *Nature Reviews Molecular Cell Biology, 15*(11), 709–721. https://doi.org/10.1038/nrm3891

Ioannidis, N. M., Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., Musolf, A., Li, Q., Holzinger, E., Karyadi, D., Cannon-Albright, L. A., Teerlink, C. C., Stanford, J. L., Isaacs, W. B., Xu, J., Cooney, K. A., Lange, E. M., Schleutker, J., Carpten, J. D., … Sieh, W. (2016). REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *American Journal of Human Genetics, 99*(4), 877–885. https://doi.org/10.1016/j.ajhg.2016.08.016

Johnston, S. H., Rauskolb, C., Wilson, R., Prabhakaran, B., Irvine, K. D., & Vogt, T. F. (1997). A family of mammalian Fringe genes implicated in boundary determination and the Notch pathway. *Development, 124*(11), 2245–2254.

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kakuda, S., & Haltiwanger, R. S. (2017). Deciphering the Fringe-mediated Notch Code: Identification of activating and inhibiting sites allowing discrimination between ligands. *Developmental Cell, 40*(2), 193–201. https://doi.org/https://doi.org/10.1016/j.devcel.2016.12.013

Kakuda, S., LoPilato, R. K., Ito, A., & Haltiwanger, R. S. (2020). Canonical notch ligands and fringes have distinct effects on NOTCH1 and NOTCH2. *Journal of Biological Chemistry, 295*(43), 14710–14722. https://doi.org/10.1074/jbc.RA120.014407

Kanavy, D. M., McNulty, S. M., Jairath, M. K., Brnich, S. E., Bizon, C., Powell, B. C., & Berg, J. S. (2019). Comparative analysis of functional assay evidence use by ClinGen Variant Curation Expert Panels. *Genome Medicine, 11*(1). https://doi.org/10.1186/s13073-019-0683-1

Kebaish, K. M., Neubauer, P. R., Voros, G. D., Khoshnevisan, M. A., & Skolasky, R. L. (2011). Scoliosis in adults aged forty years and older: Prevalence and relationship to age, race, and gender. *Spine, 36*(9), 731–736. https://doi.org/10.1097/BRS.0b013e3181e9f120

Kwon, J. W., Chae, H. W., Lee, H. S., Kim, S., Sung, S., Lee, S. bin, Moon, S. H., Lee, H. M., & Lee, B. H. (2021). Incidence rate of congenital scoliosis estimated from a nationwide health insurance database. *Scientific Reports, 11*(1), 1–11. https://doi.org/10.1038/s41598-021-85088-7

Lefebvre, M., Dieux-Coeslier, A., Baujat, G., Schaefer, E., Judith, S. O., Bazin, A., Pinson, L., Attie-Bitach, T., Baumann, C., Fradin, M., Pierquin, G., Julia, S., Quélin, C., Doray, B., Berg, S., Vincent-Delorme, C., Lambert, L., Bachmann, N., Lacombe, D., … Thevenon, J. (2018). Diagnostic strategy in segmentation defect of the vertebrae: a retrospective study of 73 patients. *Journal of Medical Genetics, 55*(6), 422–429. https://doi.org/10.1136/jmedgenet-2017-104939

Lin, T., Meng, Y., Ji, Z., Jiang, H., Shao, W., Gao, R., & Zhou, X. (2019). Extent of Depression in Juvenile and Adolescent Patients with Idiopathic Scoliosis During Treatment with Braces. *World Neurosurgery, 126*, e27–e32. https://doi.org/10.1016/j.wneu.2019.01.095

Matsuda, M., Yamanaka, Y., Uemura, M., Osawa, M., Saito, M. K., Nagahashi, A., Nishio, M., Guo, L., Ikegawa, S., Sakurai, S., Kihara, S., Maurissen, T. L., Nakamura, M., Matsumoto, T., Yoshitomi, H., Ikeya, M., Kawakami, N., Yamamoto, T., Woltjen, K., … Alev, C. (2020). Recapitulating the human segmentation clock with pluripotent stem cells. *Nature, 580*(7801), 124–129. https://doi.org/10.1038/s41586-020-2144-9

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nature Methods, 19*(6), 679–682. https://doi.org/10.1038/s41592-022-01488-1

Moloney, D. J., Panin, V. M., Johnston, S. H., Chen, J., Shao, L., Wilson, R., Wang#, Y., Stanley, P., Irvine, K. D., Haltiwanger, R. S., & Vogt, T. F. (2000). Fringe is a glycosyltransferase that modifies Notch. *Nature, 406*, 369–375. https://doi.org/https://doi.org/10.1038/35019000

Nandagopal, N., Santat, L. A., LeBon, L., Sprinzak, D., Bronner, M. E., & Elowitz, M. B. (2018). Dynamic Ligand Discrimination in the Notch Signaling Pathway. *Cell, 172*(4), 869-880.e19. https://doi.org/10.1016/j.cell.2018.01.002

Ng, P. C., & Henikoff, S. (2001). Predicting deleterious amino acid substitutions. *Genome Research, 11*(5), 863–874. https://doi.org/10.1101/gr.176601

Nóbrega, A., Maia-Fernandes, A. C., & Andrade, R. P. (2021). Altered Cogs of the Clock: Insights into the Embryonic Etiology of Spondylocostal Dysostosis. *Journal of Developmental Biology, 9*(5), 1–14. https://doi.org/10.3390/jdb9010005

Otomo, N., Mizumoto, S., Lu, H. F., Takeda, K., Campos-Xavier, B., Mittaz-Crettol, L., Guo, L., Takikawa, K., Nakamura, M., Yamada, S., Matsumoto, M., Watanabe, K., & Ikegawa, S. (2019). Identification of novel LFNG mutations in spondylocostal dysostosis. *Journal of Human Genetics, 64*(3), 261–264. https://doi.org/10.1038/s10038-018-0548-2

Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., & Kircher, M. (2019). CADD: Predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Research, 47*(D1), D886–D894. https://doi.org/10.1093/nar/gky1016

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., & Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine, 17*(5), 405–424. https://doi.org/10.1038/gim.2015.30

Sanaki-Matsumiya, M., Matsuda, M., Gritti, N., Nakaki, F., Sharpe, J., Trivedi, V., & Ebisuya, M. (2022). Periodic formation of epithelial somites from human pluripotent stem cells. *Nature Communications, 13*(2325), 1–14. https://doi.org/10.1038/s41467-022-29967-1

Schuhmann, S., Koller, H., Sticht, H., Kraus, C., Krumbiegel, M., Uebe, S., Ekici, A. B., Reis, A., & Thiel, C. T. (2021). Clinical and molecular delineation of spondylocostal dysostosis type 3. *Clinical Genetics, 99(*6), 851–852. https://doi.org/10.1111/cge.13952

Shifley, E. T., & Cole, S. E. (2008). Lunatic fringe protein processing by proprotein convertases may contribute to the short protein half-life in the segmentation clock. *Biochimica et*

Biophysica Acta - Molecular Cell Research, 1783(12), 2384–2390. https://doi.org/10.1016/j.bbamcr.2008.07.009

Sparrow, D. B., Chapman, G., Wouters, M. A., Whittock, N. V., Ellard, S., Fatkin, D., Turnpenny, P. D., Kusumi, K., Sillence, D., & Dunwoodie, S. L. (2006). Mutation of the Lunatic Fringe gene in humans causes spondylocostal dysostosis with a severe vertebral phenotype. *American Journal of Human Genetics, 78*(1), 28–37. https://doi.org/10.1086/498879

Takeda, K., Kou, I., Mizumoto, S., Yamada, S., Kawakami, N., Nakajima, M., Otomo, N., Ogura, Y., Miyake, N., Matsumoto, N., Kotani, T., Sudo, H., Yonezawa, I., Uno, K., Taneichi, H., Watanabe, K., Shigematsu, H., Sugawara, R., Taniguchi, Y., … Ikegawa, S. (2018). Screening of known disease genes in congenital scoliosis. *Molecular Genetics and Genomic Medicine, 6*(6), 966–974. https://doi.org/10.1002/mgg3.466

Tavtigian, S. v., Deffenbaugh, A. M., Yin, L., Judkins, T., Scholl, T., Samollow, P. B., de Silva, D., Zharkikh, A., & Thomas, A. (2006). Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *Journal of Medical Genetics, 43*(4), 295–305. https://doi.org/10.1136/jmg.2005.033878

Turnpenny, P. D., Whittock, N., Duncan, J., Dunwoodie, S., Kusumi, K., & Ellard, S. (2003). Novel mutations in DLL3, a somitogenesis gene encoding a ligand for the Notch signalling pathway, cause a consistent pattern of abnormal vertebral segmentation in spondylocostal dysostosis. *Journal of Medical Genetics, 40*(5), 333–339. https://doi.org/10.1136/jmg.40.5.333

Venkatachalam, K., Hofmann, T., & Montell, C. (2006). Lysosomal localization of TRPML3 depends on TRPML2 and the mucolipidosis-associated protein TRPML1. *Journal of Biological Chemistry, 281*(25), 17517–17527. https://doi.org/10.1074/jbc.M600807200

Williams, D. R., Shifley, E. T., Braunreiter, K. M., & Cole, S. E. (2016). Disruption of somitogenesis by a novel dominant allele of Lfng suggests important roles for protein processing and secretion. *Development, 143*, 822–830. https://doi.org/10.1242/dev.128538

Williams, D. R., Shifley, E. T., Lather, J. D., & Cole, S. E. (2014). Posterior skeletal development and the segmentation clock period are sensitive to Lfng dosage during somitogenesis. *Developmental Biology, 388*(2), 159–169. https://doi.org/10.1016/j.ydbio.2014.02.006

Williams, S., Alkhatib, B., & Serra, R. (2019). Development of the axial skeleton and intervertebral disc. In *Current topics in developmental biology* (pp. 50–82). Elsevier. https://doi.org/10.1016/bs.ctdb.2018.11.018

Wu, M. M., Llopis, J., Adams, S., Mccaffery, J. M., Kulomaa, M. S., Machen, T. E., Moore, H.-P. H., & Tsien, R. Y. (2000). Organelle pH studies using targeted avidin and fluorescein-biotin. *Chemistry & Biology, 7*(3), 7(3), 197–209.

Yoshioka-Kobayashi, K., Matsumiya, M., Niino, Y., Isomura, A., Kori, H., Miyawaki, A., & Kageyama, R. (2020). Coupling delay controls synchronized oscillation in the segmentation clock. *Nature, 580*(7801), 119–123. https://doi.org/10.1038/s41586-019-1882-z

Zhang, N., Norton, C. R., & Gridley, T. (2002). Segmentation defects of Notch pathway mutants and absence of a synergistic phenotype in lunatic fringe/radical fringe double mutant mice. *Genesis, 33*(1), 21–28. https://doi.org/10.1002/gene.10081

Chapter 3: Integrative Analysis of Lunatic Fringe Variants Associated with Spondylocostal Dysostosis Type-III

[1,2]Parker Wengryn, [1,3]Felicity Fenrich, [1]Karina da Costa Silveira, [1]Connor Oborn, [4]Shuji Mizumoto, [1,2]Alexander Beke, [1]Carrie-Lynn Soltys, [4]Shuhei Yamada, [1]Peter Kannu.

[1]University of Alberta, Department of Medical Genetics, Edmonton, Alberta, Canada.

[2]University of Alberta, Department of Medicine, Edmonton, Alberta, Canada

[3]University of Guelf, Department of Molecular and Cellular Biology, Guelf, Ontario, Canada

[4]Meijo University, Department of Pathobiochemistry, Nagoya, Aichi, Japan

Peter Kannu is the corresponding author.

816 Medical Sciences Building, University of Alberta

Edmonton, Alberta, Canada, T6G 2H7

Phone: (780) 492-9044

Email: kannu@ualberta.ca

3.1. Introduction

The structure of vertebrae and ribs directly results from the symmetrical formation of their precursor structures, somites, in a process called somitogenesis. Somitogenesis occurs early in human embryogenesis and is tightly regulated by the segmentation clock (Chu et al., 2019; Dequéant & Pourquié, 2008; Evrard et al., 1998; Matsuda et al., 2020; Miao et al., 2023; Sanaki-Matsumiya et al., 2022; S. Williams et al., 2019). A clock cycle lasts six hours and results in the formation of a new pair of somites which bud bilaterally from the paraxial mesoderm (Matsuda et al., 2020; Sanaki-Matsumiya et al., 2022; Yoshioka-Kobayashi et al., 2020). Somitogenesis progresses in a rostrocaudal fashion until all somitic pairs are formed.

The NOTCH signaling pathway is integral to the control of the clock mechanism and thus regulates somitogenesis(Kakuda et al., 2020; Nóbrega et al., 2021). *LUNATIC FRINGE* (*LFNG*), a Golgi-resident β-1,3-*N*-acetylglucosaminyltransferase, modifies the NOTCH receptor extracellular domain by transferring *N*-acetylglucosamine (GlcNAc) to specific fucose residues (Evrard et al., 1998; Johnston et al., 1997; Kakuda & Haltiwanger, 2017; Moloney et al., 2000; Venzin & Oates, 2020). The glycosylation process modifies how NOTCH receptors interact with NOTCH ligands, essentially recalibrating the signaling mechanism. Resetting the clock mechanism is pivotal for establishing both the anterior and posterior intra-somitic halves, which are crucial for the development of a symmetrical spondylocostal axis (Musumeci et al., 2015; Shifley et al., 2008; Shifley & Cole, 2008).

Finely tuned regulation of *LFNG* expression and LFNG processing is required to ensure equal somitic partitioning (Shifley & Cole, 2008; D. R. Williams et al., 2016). LFNG is initially synthesized as a 43 kDa preproprotein with two key regions: an N-terminal, type-II

transmembrane domain (amino acids 1-86), and an enzymatically active FRINGE domain (amino acids 108-358) (Johnston et al., 1997; Moloney et al., 2000; Rampal et al., 2005). Following translation, LFNG is relocated to the Golgi apparatus, where its transmembrane region serves as an anchor for its FRINGE domain, which faces the Golgi lumen. It is at this location that LFNG exerts its influence on NOTCH receptors. Subtilisin-like proteinase 6 (SPC6) cleaves the hydrophobic tether at amino acid 86, liberating the processed LFNG (34 kDa) into the extracellular environment within a brief period (Shifley & Cole, 2008; D. R. Williams et al., 2016). This exocytosis process finely tunes segmentation clock timing, facilitating the symmetrical formation of somites.

Biallelic pathogenic variants in *LFNG* cause Spondylocostal Dysostosis Type-III (SCD3) (OMIM: 609813) (Lefebvre et al., 2018; Otomo et al., 2019; Schuhmann et al., 2021; Sparrow et al., 2006; Takeda et al., 2018; Wengryn et al., 2023). A subtype of the larger SCD disease family (Matsumoto et al., 2021; Nóbrega et al., 2021), SCD3 is characterized by short stature and segmentation defects of the vertebrae (SDV) including hemivertebrae, block vertebrae, missing vertebrae, scoliosis, and rib fusion abnormalities (Dunwoodie, 2009). Abnormal vertebrae formation results in scoliosis and affects the development of the thorax, compromising lung development and potentially leading to restrictive lung disease. Genotyping is often required to distinguish between SCD subtypes due to their phenotypic similarities. Within SCD3, varied presentation includes the prevalence and severity of scoliosis, vertebral number, and the location of the spinal segmentation defects (Otomo et al., 2019; Takeda et al., 2018; Wengryn et al., 2023). Less frequent physical findings include arachnodactyly, camptodactyly, and auditory impairment (Lecca et al., 2023; Schuhmann et al., 2021; Sparrow et al., 2006). Perinatal

mortality is uncommon with most affected individuals surviving to adulthood. Previous work has suggested that different *LFNG* variants may modulate phenotypic presentation (Wengryn et al., 2023). There, it was hypothesized that a hypomorphic *LFNG* genotype resulted in relatively increased activity that would have damaged vertebral column formation less than a null genotype. However, with only 13 cases published in the scientific literature (Lecca et al., 2023; Lefebvre et al., 2018; Otomo et al., 2019; Schuhmann et al., 2021; Sparrow et al., 2006; Takeda et al., 2018; Wengryn et al., 2023), it is difficult to draw firm conclusions from the limited data.

The genotypic and functional data of nine *LFNG* missense variants associated with SCD3 published prior to 2023 are listed in Tables 2 and 3. Analysis of potential genotype-phenotype correlations centers on four aspects of the LFNG enzyme: 1) glycosyltransferase activity, 2) protein localization, 3) protein processing, 4) NOTCH signaling activation (Otomo et al., 2019; Sparrow et al., 2006; Takeda et al., 2018; Wengryn et al., 2023). Previous research has primarily utilized the first three approaches to investigate *LFNG* variant pathogenicity as these can provide a specific mechanism for the fourth. However, not all approaches have been used consistently across the published *LFNG* variants. In this study, we aim to functionally characterize nine *LFNG* missense variants associated with SCD3 published with radiographic findings (Tables 2, 3) in terms of glycosyltransferase activity, subcellular localization, and protein processing. The goal is to provide *in vitro* evidence that gauges the relative function of each *LFNG* variant, filling experimental gaps from previously published literature to provide general correlations with available SCD3 phenotypic data. We hope that this will form the basis of future work aimed at correlating updated, thorough, and uniform clinical data with the molecular findings collected here, initiating discussion surrounding the genotype-phenotype relationship.

Publication	Variant	DNA change (GRCh38)	ClinVar ID	SNPdb ID	LOVD DB ID	gnomAD 3.1.2	Splice AI
Sparrow et al. 2006	c.564C>A	NC_000007.14: g.2525301C>A	6999	rs104894024	LFNG_000030	0.00	0.01
Lefebvre et al. 2018	c.583T>C	NC_000007.14: g.2525415T>C	N/A	N/A	N/A	0.00	0.00
Lefebvre et al. 2018	c.842C>A	NC_000007.14: g.2526264C>A	N/A	N/A	N/A	0.00	0.01
Takeda et al. 2018	c.467T>G	NC_000007.14: g.2524729T>G	N/A	N/A	N/A	0.00	0.00
Takeda et al. 2018	c.856C>T	NC_000007.14: g.2526278C>T	N/A	rs752671299	N/A	0.00	0.00
Otomo et al. 2019	c.601G>A	NC_000007.14: 2525432:G>A	619139	rs1211456697	LFNG_000032	0.00	0.00
Schuhmann et al. 2021	c.446C>T	NC_000007.14: g.2524708C>T	N/A	rs867769213	N/A	0.00	0.00
Wengryn et al. 2023	c.521G>A	NC_000007.14: g.2525258G>A	1062453	N/A	LFNG_000034	6.57×10^{-6}	0.01
Wengryn et al. 2023	c.766G>A	NC_000007.14: g.2525715G>A	1003507	rs1437427476	LFNG_000035	6.57×10^{-6}	0.07

Table 2: *LFNG* Variants of Interest and Genetic Data. GnomAD 3.1.2: Genome Aggregation Database (version 3.1.2). LOVD ID: Leiden Open Variation Database Identifier. SNPdb-ID: Single Nucleotide Polymorphism Database Identifier.

Publication	Protein	REVEL	GERP	RMSD	Predicted Structural Change?	Predicted Distance (Å) to p.D201	Predicted Distance (Å) to p.290D	NOTCH Signalling Assay Activity	Enzyme Activity	Subcellular Localization	PPT-PC Secretion	N-terminal Cleaved
Sparrow et al. 2006	p.F188L	0.51	4.93	N/A	No‡	16.1	17.8	Null	Null‡	Mislocalized†	No‡	No
Lefebvre et al. 2018	p.W195R	0.90	5.32	0.392	Yes	16.8	19.0	N/A	Null	Mislocalized	No	No
Lefebvre et al. 2018	p.T281K	0.62	4.70	0.476	Yes	11.8	24.5	N/A	Null	Mislocalized	No	No
Takeda et al. 2018	p.L156R	0.74	4.83	0.521	Yes	26.1	16.3	N/A	Null‡	Mislocalized	No‡	No
Takeda et al. 2018	p.R286W	0.73	4.70	N/A	No	11.7	24.2	N/A	Null‡	Mislocalized + Golgi	Yes‡	Yes
Otomo et al. 2019	p.D201N	0.84	5.32	0.318	No	15.3	0.0	N/A	Null‡	Mislocalized	No‡	No
Schuhmann et al. 2021	p.T149I	0.68	3.95	0.542	Yes	17.3	10	N/A	Partial	Mislocalized	No	No
Wengryn et al. 2023	p.R174H	0.52‡	5.18	N/A	No‡	10.1	17.4	N/A	Partial‡	Golgi†	Yes‡	Yes†
Wengryn et al. 2023	p.G256S	0.89‡	5.31	0.285	No	8.7	8.6	N/A	Null‡	Golgi†	Yes‡	Yes†

Table 3: *LFNG* Variants of Interest with Structural and Functional Consequences *In Silico* and *In Vitro.* †Supporting Previous Evidence. ‡Documented elsewhere. All structural predications were generated with Alpha-Fold. Residue distance predictions were measured from α-carbon to α-carbon with the WT model. GERP: Genomic Evolutionary Rate Profiling. PPT-PC: Pre-Pro-Trypsin Leader Sequence in Plasmid Construct. REVEL: Rare Exome Variant Ensemble Learner. RMSD: Root-Mean-Square of Atomic Deviation.

3.2. Materials and Methods

3.2.1. Cloning and subcloning

The creation of the *3XFLAG-LFNG* plasmid construct is described elsewhere (Takeda et al., 2018; Wengryn et al., 2023). The creation of the *LFNG-HA* plasmid construct is also described elsewhere (Wengryn et al., 2023). Site-directed mutagenesis (QuikChange XL Site-Directed Mutagenesis Kit, Agilent, Santa Clara, CA, USA, Cat: 200516-5) was undertaken with both constructs to obtain the desired point variants (see Table S1) following the manufacturers protocol. We then transformed Mach1 T1 phage-resistant chemically competent *E. coli* (Thermo Fisher Scientific, Whitby, ON, CAN, Cat: C86203) and extracted clonal plasmid DNA by miniprep (Wizard Plus SV Miniprep DNA Purification System, Promega, Madison, WI, USA, Cat: A1460). The plasmids were screened for point variant induction and random mutagenesis with Sanger sequencing (University of Alberta Molecular Biology Core).

3.2.2. Cell culture

The cell culture protocols for NIH-3T3 (ATCC, Manasses, VI, USA, Cat: CRL-1658, RRID: CVCL_0594) are described elsewhere (Wengryn et al., 2023). COS-7 cells (JCRB Cell Bank, Osaka, JPN, Cat: JCRB9127, RRID: CVCL_0224) were cultured in high glucose DMEM (FujiFilm Waco, Osaka, JPN, Cat: 044-29765) and 10% FBS (Hyclone, Logan, UT, USA, Cat: SH30396.03). All cell lines were tested for mycoplasma contamination via PCR.

3.2.3. *In silico* genetic analysis

REVEL, GERP, and SpliceAI scores were obtained for each of the nine genetic variants through the web tool Franklin (Genoox, Palo Alto, CA, USA, https://franklin.genoox.com/clinical-db/home, RRID: SCR_024670).

3.2.4. *In silico* protein analysis

The protocol for proteomic analysis is described in our prior publication (Wengryn et al., 2023). Briefly, wildtype and variant FASTA files were fed into ColabFold v1.5.2 without any relaxation steps and the top-ranked structure for each variant was chosen for visual analysis based on pLDDT score. Visual analysis was performed in the PyMOL Molecular Graphics System, Version 2.5 (Schrödinger, New York, NY, USA, RRID: SCR_000305).

3.2.5. Ortholog/Paralog Comparisons

Accession codes for LFNG, MFNG, and RFNG were retrieved from UniprotKB (http://www.uniprot.org/help/uniprotkb, RRID: SCR_004426) and Uniparc (http://www.uniprot.org/uniparc, RRID: SCR_005818). Chosen sequences were aligned with MAFFT (http://mafft.cbrc.jp/alignment/server, RRID: SCR_011811) and then visualized with the MVIEW tool [https://desmid.github.io/mview, RRID: SCR_024129]). Initially, 156 orthologous sequences that shared 50% sequence conservation with Q8NES3 (LFNG, Homo sapiens) were compared (Uniref50 [http://www.uniprot.org/help/uniref, RRID: SCR_010646]; Uniparc) from which six LFNG sequences from human, cow, mouse, rat, African clawed frog, xenopus, and zebrafish were selected to consolidate orthologous comparison (Uniref50; UniprotKB/Swiss-Prot [https://www.expasy.org/resources/uniprotkb-swiss-prot, RRID: SCR_021164]). Furthermore,

four isoform sequences of human LFNG, two isoform sequences of human MFNG, and the sequence for human RFNG were used in paralogous analysis (Uniprotkb/swiss-prot). The aforementioned sequences were labeled with their corresponding accession codes accordingly (Figure S3A).

3.2.6. LFNG-HA Western Blotting

The protocol for western blotting is reported elsewhere (Wengryn et al., 2023). Briefly, 15 µL of 1µg/µL protein lysate (15µg total protein per lane) was run on a 12% SDS-polyacrylamide gel and blotted on nitrocellulose membrane. Membranes were blocked in 5% skim milk solution for one hour and incubated overnight with mouse anti-β-actin (Santa Cruz, Freemont, CA, USA, Cat: 47778, RRID: AB_2714189; 1:1000) or rabbit anti-HA (Cell Signaling Technology, Danvers, MA, USA, Cat: 3724, RRID: AB_1549585, 1:1000). The following morning, blots were incubated for one hour in HRP-linked goat anti-mouse IgG (Cell Signaling Technology, Cat: 7076, RRID: AB_330924, 1:1000) or HRP-linked goat anti-rabbit IgG (Cell Signaling Technology, Cat: 7074, RRID: AB_2099233, 1:1000) and imaged on a ChemiDoc MP Imaging System (BioRad, Mississauga, ON, CAN, Cat: 12003154). Densitometry was performed in BioRad Image Lab 6.1 Software (http://www.bio-rad.com/en-us/sku/1709690-image-lab-software, RRID: SCR_014210).

3.2.7. LFNG-HA Immunofluorescence Microscopy

The protocol for immunofluorescence microscopy is reported elsewhere (Wengryn et al., 2023). Briefly, NIH-3T3 cells were transiently transfected with each of the *LFNG-HA* plasmids and incubated for 20 hours. The following day, the cells were fixed, permeabilized, blocked, and

incubated with rabbit anti-HA (Cell Signaling Technology, Cat: 3724, RRID: AB_1549585, 1:800) and mouse anti-GM130 (BD Biosciences, Mississauga, ON, CAN, Cat: 610822, RRID: AB_398141, 1:800) primary antibodies. Cells were then treated with goat anti-mouse Alexa Fluor 594 (Thermo Fisher Scientific, Cat: A-11005, RRID: AB_2534073, 1:500), and goat anti-rabbit Alexa Fluor 488 (Thermo Fisher Scientific Cat: A-11034, RRID: AB_2576217, 1:500), followed by HOECHST 3342 (Thermo Fisher Scientific, Cat: H-3570, 1 μg/mL). Slides were mounted before imaging at 60X on the WaveFX Spinning Disk Confocal Microscope (Quorum Technologies, Cambridge, ON, CAN) at the University of Alberta Cell Imaging Core.

3.2.8. Functional Analysis and 3XFLAG-LFNG Western Blotting

The protocol for functional analysis is reported elsewhere (Otomo et al., 2019; Takeda et al., 2018; Wengryn et al., 2023) and the description of the method partly replicates the wording. 5 µg of *3XFLAG-LFNG* plasmid was transfected into CV-1 in Origin with SV40 genes (COS-7) cells (~80% confluency) on 100-mm culture dish using Lipofectamine 3000 (Thermo Fisher Scientific, Cat: L3000001) according to the manufacturer's instructions. Three days after transfection, protein was isolated, incubated with 10 µl of anti-DYKDDDDK agarose resin (FujiFilm Wako, Cat: 012-22781), eluted at 95°C with SDS buffer, then western blotted. The membrane was then incubated overnight at 4˚C with anti-DYKDDDDK antibody (FujiFilm Wako, Cat: 018-22381, RRID: AB_10659453). The bound antibody was detected the next morning with IRDye 680RD donkey anti-mouse IgG (Li-Cor, Lincoln, NA, USA, Cat: 926-68072, RRID: AB_10953628) on an Amersham ImageQuant Fluor 800 (Cytiva, Tokyo, JPN, Cat: 29399484). The amount of recombinant 3XFLAG-LFNG protein was estimated with a

standard curve of 3XFLAG-tagged Bacterial Alkaline Phosphatase with Image Quant (GE Life Sciences, Chicago, IL, RRID: SCR_014246).

The Glc-NAc-transferase assay mixture contained 10 µl of enzyme-bound anti-FLAG affinity resins (100~200 ng protein of the recombinant LFNG), 50 mM 2-(N-morpholino)-ethanesulfonic acid-NaOH (pH 6.5) (Nacalai tesque, Kyoto, JPN, Cat: 21623-26), 10 mM $MnCl^2$ (Nacalai tesque, Cat: 21211-45), 0.1 mM UDP-GlcNAc (Promega, Cat: V7071) as the sugar donor substrate, and 1 mM *p*-nitrophenyl-α-L-fucose (pNP-Fuc; Sigma, Cat: N3628) as the sugar acceptor in a total volume of 50 µL. The reaction mixture was incubated at 37 ˚C for 4 hours. The UDP reaction product released from UDP-GlcNAc was mixed with UDP detection reagent from the UDP-Glo Glycosyltransferase Assay (Promega, Cat: V6961). The newly synthesized ATP was measured using a luciferase/luciferin reaction, and the luminescent signal was detected using an EnSpire multiwell plate reader (PerkinElmer, Waltham, MA, USA, Cat: 6057420).

3.2.9. Statistical Analysis

One-way ANOVA was performed to determine whether a statistically significant difference between 3 or more groups exists ($p = 0.05$). Upon validation, Bonferroni-adjusted post-hoc analysis was employed to quantify differences between specific groups. Data analyses were performed in Excel (Microsoft, Redmond, WA, USA, https://www.microsoft.com/en-gb, RRID: SCR_016137) and GraphPad Prism version 10.0.0 for Windows (GraphPad Software, Boston, MA, USA, https://www.graphpad.com, RRID: SCR_002798).

3.3. Results

3.3.1. *LFNG* Variants Lead to Inhibition of Protein Processing

Protein processing is a critical aspect of LFNG post-translational regulation during somitogenesis. Improper processing is associated with SCD3-like phenotype in murine models (D. R. Williams et al., 2016), but only one human variant (c.564C>A) [p.F188L]) was previously shown to inhibit this process (Sparrow et al., 2006; Wengryn et al., 2023). To test protein processing for each of the nine variants, western blot was undertaken with cell lysate obtained from NIH-3T3 cells transiently transfected with variant-induced HA-tagged *LFNG* plasmids (Figure 13). Two of the nine variants were previously shown to process normally (p.R174H, p.G256S) and were used as positive controls (Wengryn et al., 2023). One of the nine variants (p.F188L) was shown to prevent processing and was used as a negative control. In the six experimental conditions, there were variable yet statistically insignificant differences between the amount of unprocessed variant LFNG and WT (MW = 43 kDa) (Figure 13A, B). The three control conditions (p.R174H, p.F188L, and p.G256S) also did not yield significant differences in unprocessed protein. Interestingly, five variants lead to decreased cleavage of LFNG (MW = 34 kDa) compared to WT: p.T149I ($p = 5.0 \times 10^{-8}$), p.L156R ($p = 5.8 \times 10^{-5}$), p.W195R ($p = 9.2 \times 10^{-10}$), p.D201N ($p = 1.7 \times 10^{-5}$), and p.T281K ($p = 6.3 \times 10^{-5}$) (Figure 13C, D). No significant difference was observed between p.R286W and WT ($p = 1.9 \times 10^{-2}$) as the Bonferroni-adjusted $\alpha = 5.56 \times 10^{-3}$ (Figure 13D). The amount of the p.F188L control was significantly decreased, as previously reported ($p = 2.4 \times 10^{-8}$), whereas the differences between the p.R174H and p.G256S controls were not statistically significant (Figure 13C, D). Therefore, this data indicated that only three (p.R174H, p.G256S, and p.R286W) of the nine LFNG variants were normally processed.

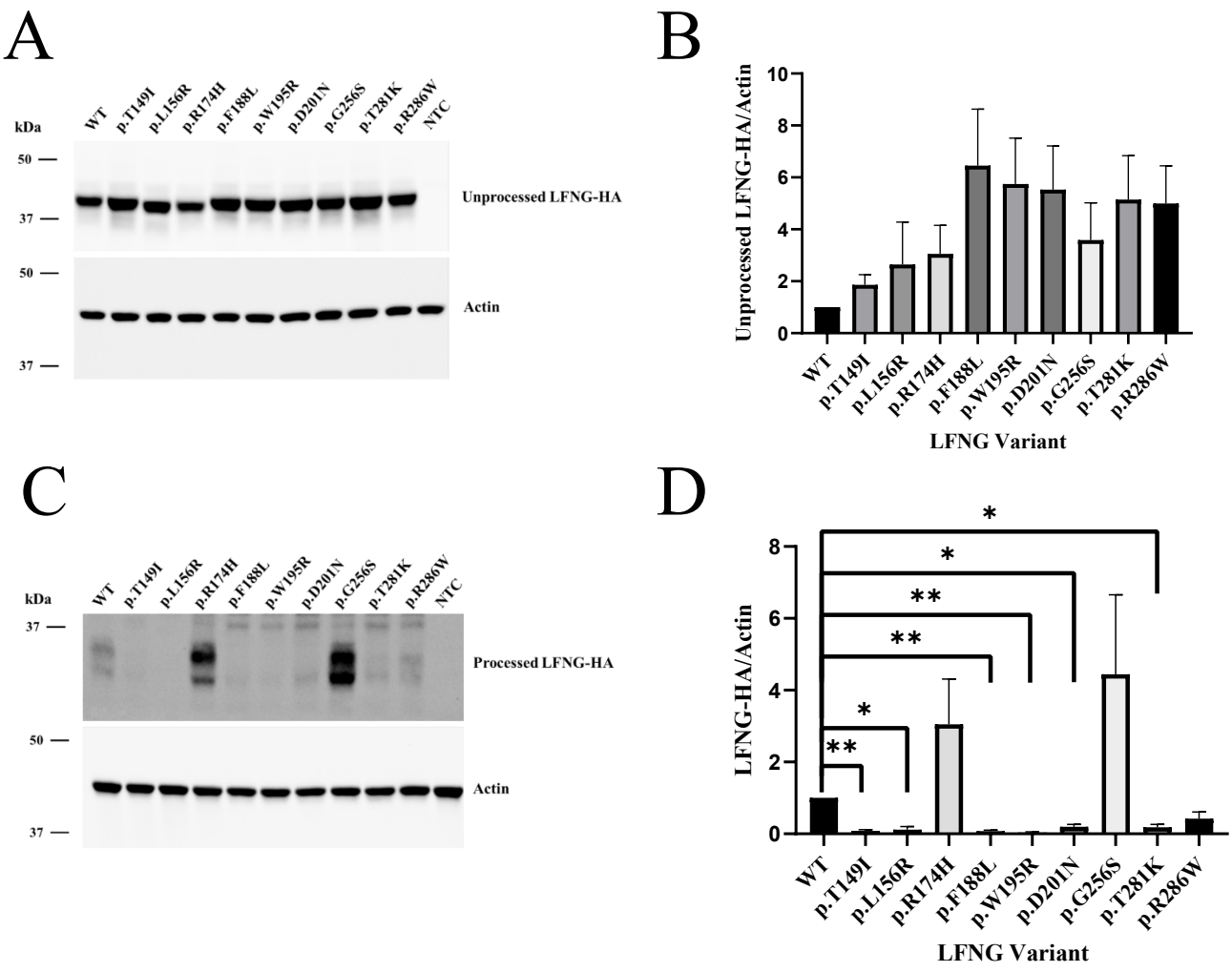


Figure 13: Western Blot of LFNG Variants to Assess Protein Processing. A, C) Western blot of cell lysate (15 µg per sample) from NIH-3T3 cells transiently transfected with *LFNG-HA* constructs. Samples were run on a 10% SDS gel and membranes were incubated with anti-HA primary then anti-Rabbit IgG secondary antibodies, or anti-Actin primary then anti-Mouse IgG secondary antibodies. A) and C) are representative samples from two separate blots. **B, D)** Densitometry of pre-pro-LFNG-HA (B) or processed (D) LFNG-HA bands normalized to actin. B) One-Way ANOVA failed to identify significant differences in the pre-pro-LFNG-HA condition ($F(4, 9) = 1.54$, $p = 1.8 \times 10^{-1}$. D) One-Way ANOVA identified significant differences in the LFNG-HA ($F(4, 9) = 3.6$, $p = 3.9 \times 10^{-3}$) condition. After Bonferroni post-hoc adjustment ($\alpha = 5.56 \times 10^{-3}$), WT band signal was significantly more intense than those of p.T149I, p.L156R, p.F188L, p.W195R, p.D201N, and p.T281K. Bars indicate mean ± S.E of four independent experiments (N = 4) plated in triplicate. *$p < 1.0 \times 10^{-4}$, **$p < 1.0 \times 10^{-7}$.

3.3.2. *LFNG* Variants Which Were Not Processed Were Also Mislocalized

The variants p.R174H and p.G256S have previously been shown to be associated with normal levels of processed LFNG and colocalization within the Golgi (Wengryn et al., 2023). Moreover, it was known that p.F188L does not process LFNG appropriately and mislocalized (Sparrow et al., 2006; Wengryn et al., 2023). As a result, it was hypothesized that protein misplacement inhibits LFNG processing, primarily because SPC6 cleaves LFNG within the Golgi. To test this hypothesis, we predicted that the unprocessed variants (p.T149I, p.L156R, p.W195R, p.D201N, and p.T281K) would be mislocalized whereas the processed variant (p.R286W) would be colocalized with the Golgi when visualized using immunofluorescent microscopy. Colocalization generally appears as a concentrated, overlapping LFNG-Golgi signal with low levels of diffusive intracellular LFNG signal (Shifley & Cole, 2008; Sparrow et al., 2006; Wengryn et al., 2023). In contrast, these works characterized mislocalization as a diffuse intracellular LFNG signal with slight artifactual or no Golgi signal overlap.

In this experiment, p.F188L was employed as a negative control (Figure 14E) whereas p.R174H (Figure 14D) and p.G256S (Figure 14H) were used as positive, localized controls. As has previously been shown, p.R174H and pG256S, as well as the WT (Figure 14A) control localized appropriately while p.F188L (Figure 14E) did not (Sparrow et al., 2006; Wengryn et al., 2023). Five of the remaining six variants, p.T149I (Figure 14B), p.L156R (Figure 14C), p.W195R (Figure 14F), p.D201N (Figure 14G), and p.T281K (Figure 14I), mislocalized. Notably, the ninth variant, p.R286W, only some of the cells (Figure 14J, K) were appropriately localized but did so within a ‘mislocalized’ pattern, although this result correlated with relative protein processing (see, ‘Mislocalization Prevents Processing and Secretion). Our prediction that

abnormal processing of the variants was due to mislocalization was largely supported. Full composites with unmerged images are available in the supplementary information (Figure S2).

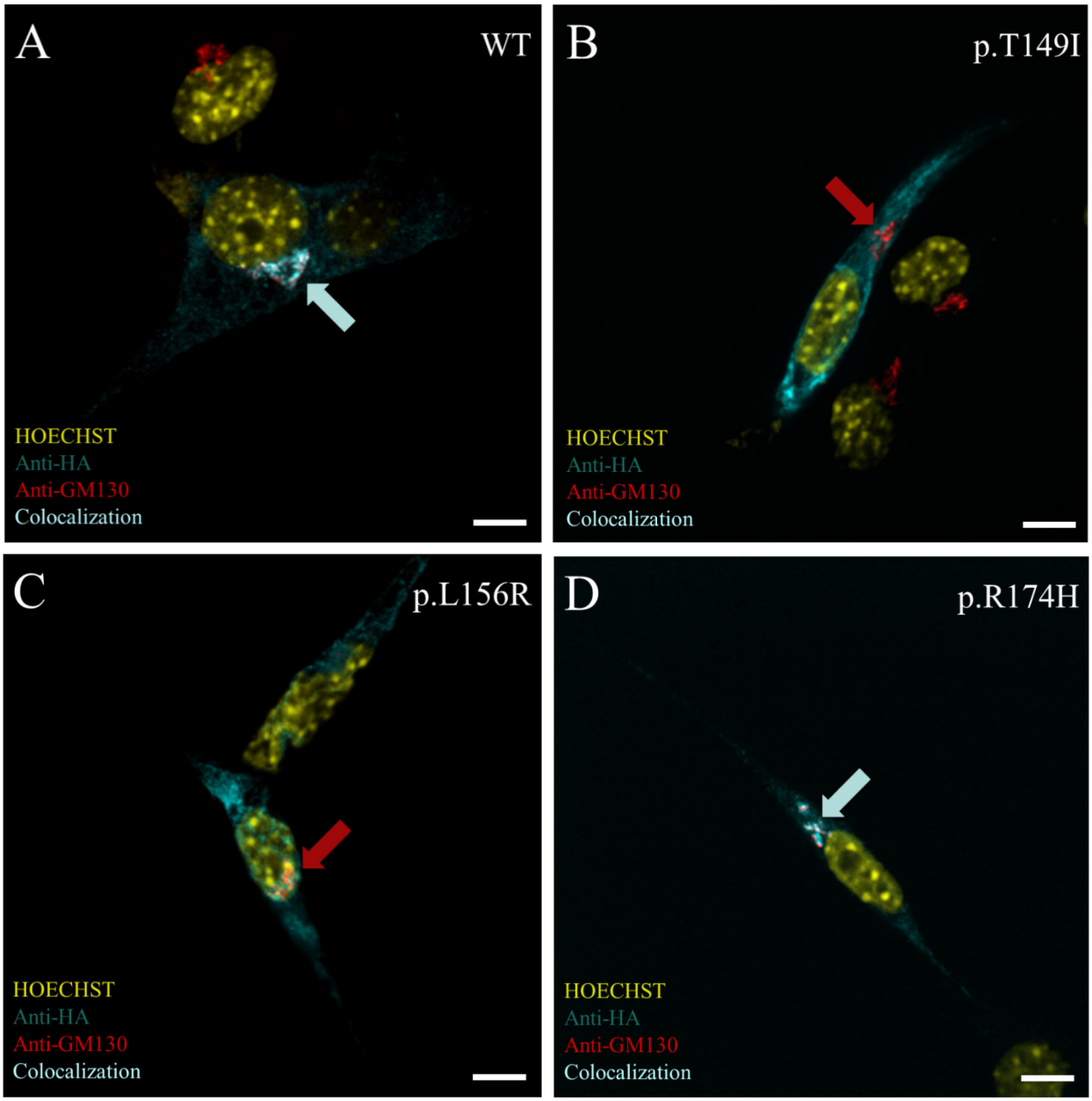


Figure 14: Immunofluorescent Imaging of LFNG Variants.

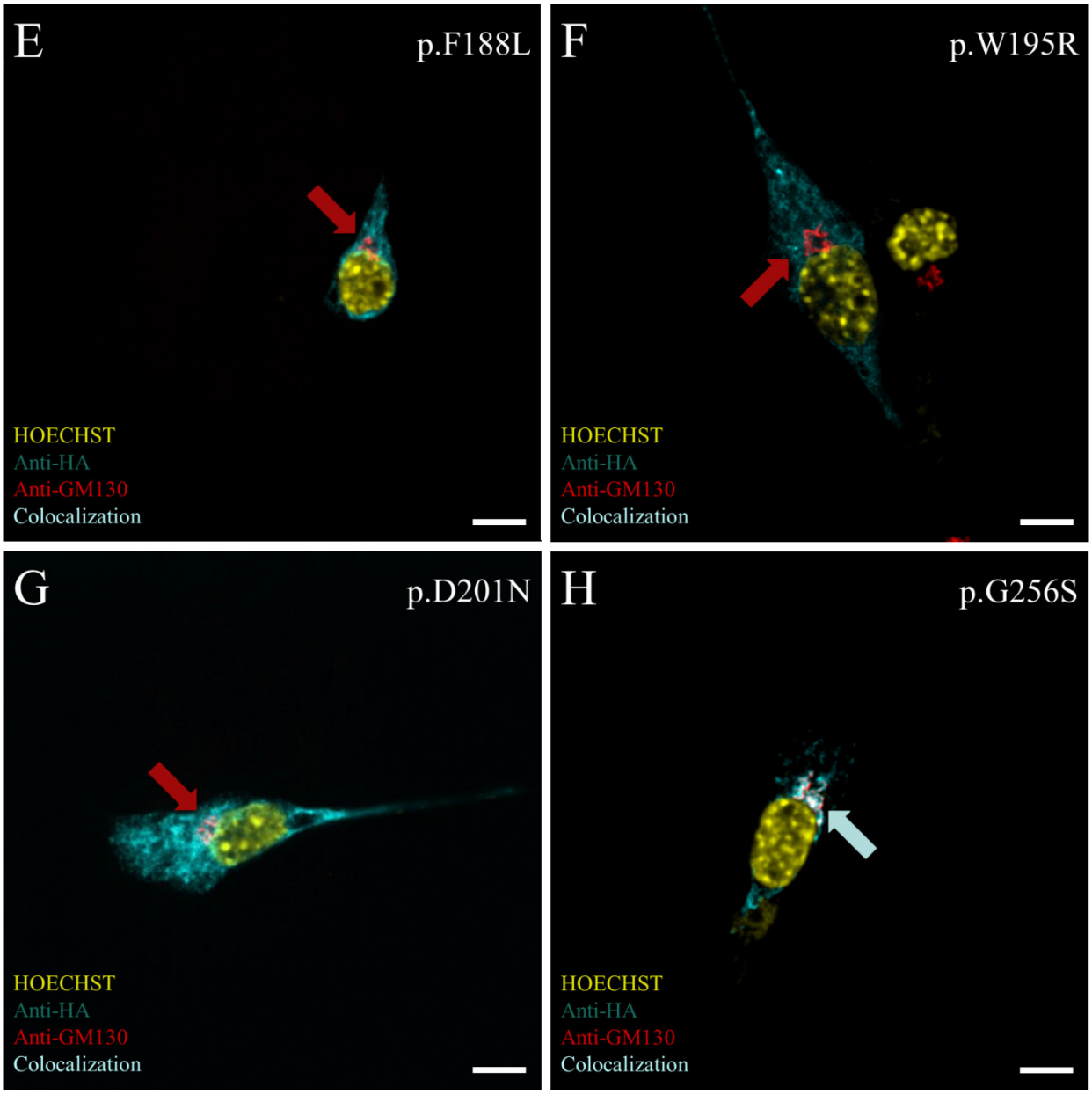


Figure 14: Immunofluorescent Imaging of LFNG Variants

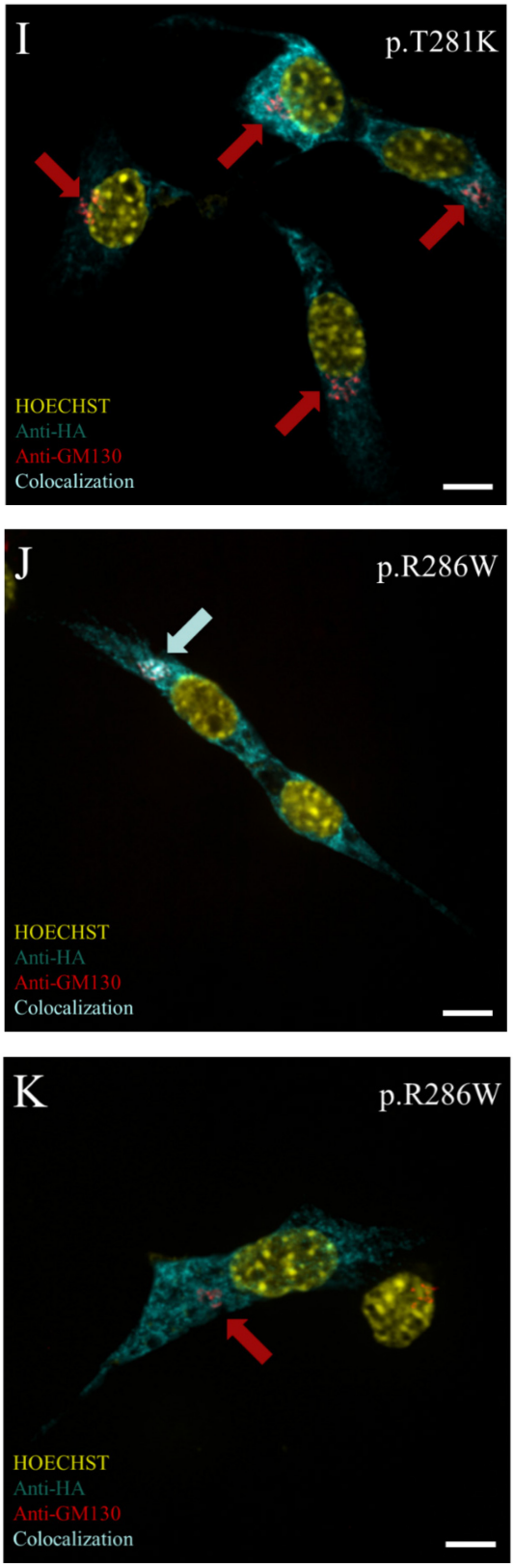


Figure 14: Immunofluorescent Imaging of LFNG Variants. Immunofluorescent microscopy images of NIH-3T3 cells transiently transfected with *LFNG* variant plasmids. Cells were treated with rabbit anti-HA and mouse anti-GM130 primary antibodies, then fluorescent secondary antibodies followed by HOECHST. Turquoise arrows indicate Anti-HA/Anti-GM130 signal colocalization. Red arrows indicate Anti-GM130 signal and thus lack of Anti-HA/Anti-GM130 colocalization. Images are representative samples of 40 cells screened at random across the three independent experiments (N = 3), and values (/40) indicate number of cells which showed the pattern exemplified in the panel. **A)** WT-(40/40), **B)** p.T149I (38/40), **C)** p.L156R (36/40), **D)** p.R174H (40/40), **E)** p.F188L (39/40), **F)** p.W195R (38/40), **G)** p.D201N (37/40), **H)** p.G256S (40/40), **I)** p.T281K (38/40), and p.R286W **(J, K)**. Panel J) shows an example of localization (25/40) whereas panel K) shows mislocalization for the same variant (15/40). Scale bars = 15 µm. The fluorescent signal was artificially altered to be colour-blind friendly.

3.3.3. The p.T149I Substitution Is Functionally Hypomorphic Whereas p.W195R and p.T281K Are Null

Although most of the variants in this study have been previously assessed for glycosyltransferase activity, p.T149I, p.W195R, and p.T281K, have not. We therefore quantitatively assessed the effect of these three variants on Glc-NAc-transferase activity through the use of a previously validated assay (Otomo et al., 2019; Takeda et al., 2018). Only variants which were not previously assessed were tested here as the quantitative nature of this assay consistently demonstrates significant evidence of reliability in previous work (Otomo et al., 2019; Takeda et al., 2018; Wengryn et al., 2023) (Figure 15 Legend; Figure 15). To this end, we employed a *LFNG* plasmid that encoded the soluble FRINGE domain of LFNG with an N-terminal Mouse Anionic Trypsin 2 (PRRS1) pre-pro-trypsin (PPT) leader sequence and 3XFLAG epitope tag in place of the LFNG transmembrane domain and signal peptide (Takeda et al., 2018). This PPT (MSALLILALVGAAVA) was shown to direct 3XFLAG-LFNG to the secretory pathway via the Golgi to allow for highly efficient isolation.

To determine whether 3XFLAG-LFNG was secreted, western blot of protein lysate and media from COS7 cells transiently transfected with each of the *LFNG* variants was undertaken. p.T149I, p.W195R, and p.T281K 3XFLAG-LFNG variant constructs were not detected in the media (data not shown; Table 3) but all were present in the cell lysate (Figure 15A). Therefore, only LFNG from cell lysate was isolated for further analysis. Enzyme activity was assessed using the previously reported UDP-Glo assay (Otomo et al., 2019; Takeda et al., 2018). Briefly, resin-bound 3XFLAG-hLFNG was reacted at 37°C for four hours in the presence of UDP-GlcNAc (the donor) and *p*-nitrophenyl-α-L-fucose (the acceptor). The reaction by product (UDP)

was then employed in a luminescent reaction as a surrogate to quantitatively measure enzyme activity. Therefore, decreased luminescence indicated that the variant had a more deleterious effect on enzyme activity. The data indicated that p.T149I ($p = 2.1 \times 10^{-11}$), p.W195R ($p = 6.8 \times 10^{-13}$), and p.T281K ($p = 3.2 \times 10^{-13}$) had significantly decreased enzymatic activity (Figure 15B). Although there was no statistically significant difference in activity between p.W195R and p.T281K ($p > 9.9 \times 10^{-1}$), p.T149I activity was significantly elevated compared to both p.W195R ($p = 4.3 \times 10^{-2}$) and p.T281K ($p = 8.5 \times 10^{-3}$) (Figure 15B). p.T149I is the second *LFNG* allele associated with SCD3 to be described with hypomorphic activity.

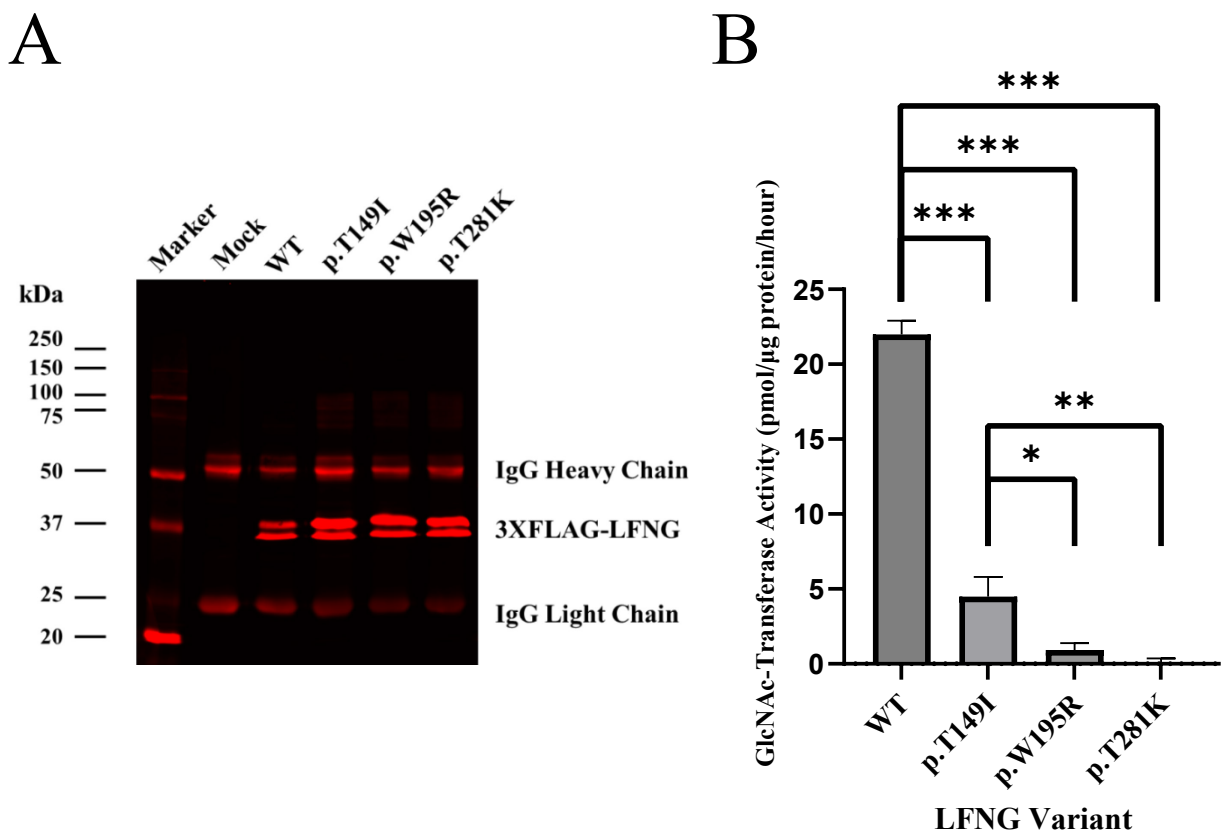


Figure 15: Quantitative Analysis of LFNG Glycosyltransferase Activity. A) Fluorescent western blotting of 3XFLAG-tagged WT, p.T149I, p.W195R, and p.T281K-LFNG proteins. The expression level of recombinant 3XFLAG-LFNG was quantified by the fluorescent intensity using a calibration curve of 3XFLAG-bacterial alkaline phosphatase (Sigma). Two bands of ~25 and 50 kDa were light and heavy chains of IgG, respectively, derived from anti-FLAG antibody-conjugated affinity gel, which was utilized for purification of recombinant LFNG. **B)** Glc-NAc-transferase activity of the recombinant LFNG. The recombinant LFNGs were incubated with UDP-GlcNAc and pNP-Fuc as the donor and acceptor substrates, respectively. The amount of reaction product, UDP, was determined by UDP-Glo™ Glycosyltransferase Assay kit. The variant LFNG's showed significantly reduced enzyme activity. *p < 5.0 x 10-2, **p < 1.0 x 10^{-2}, ***p < 1.0 x 10^{-10} vs. WT enzyme by ANOVA and Bonferroni post-hoc comparisons (N = 6).

3.3.4. In Silico Proteomic and Genetic Analysis Partially Explains Functional Perturbations

To further investigate a potential biochemical mechanism to explain the changes in p.T149I, p.L156R, p.W195R, p.D201N, p.T281K, and p.R286W LFNG variant enzyme, we modelled the variants *in silico* with AlphaFold2 generated structures. The structures of p.F188L, p.R174H, and p.G256S were assessed elsewhere (Sparrow et al., 2006; Wengryn et al., 2023). The p.T149I substitution (RMSD = 0.542) appeared to have led to the disruption of a hydrogen bond between Thr-149 and His-126 and destabilized a region containing two alpha helices between Asp-131 and His-159 (Figure 16A; Table 3). Similarly, the p.L156R (RMSD = 0.521) substitution shifted the alpha helix from Asp-153 towards Thr-160 (Figure 16B). The change induced by p.W195R (RMSD = 0.392) shifted His-219 outwards from the protein core and was predicted to destabilize the Arg-195/Ser-216 loop but not any flanking secondary structures (Figure 16C). Finally, p.T281K (RMSD = 0.476) pushed against Arg-284 and Tyr-295 and appears to have pivoted an entire beta-hairpin from residues 234 to 250 away from the protein core (shifted up to a maximum of 2.6 angstroms) (Figure 16D). The RMSD's of p.D201N (0.318) and p.R286W (0.285) do not predict significant structural changes.

Previous work has demonstrated evidence suggesting substitutions closer to the active site are more likely to affect catalysis directly whereas those farther may affect catalysis through structural changes (Luther et al., 2009). This would predict that substitutions closer to key enzymatic sites would be less likely to affect structure and more likely to affect function. Therefore, we measured predicted α-carbon to α-carbon distances between each variant's substituted amino acid and LFNG's key enzymatic residues, the DxD motif (p.D200-D201-

D202) and the catalytic residue (p.D290) (Table 3; Figure 6). In order of closest to farthest, the predicted α-carbon to α-carbon distances from p.D201(DxD) to each variant were as follows: 1) p.D201 = 0.0 Å, 2) p.G256 = 8.6 Å. 3) p.T149 = 10.0 Å; 4) p.L156 = 16.3 Å; 5) p.R174 = 17.4 Å; 6) p.F188 = 17.8 Å; 7) p.W195 = 19.0 Å; 8) p.R286 = 24.2 Å; 9) p.T281- 24.5 Å. In order of closest to farthest, the predicted α-carbon to α-carbon distances from p.D290 (the catalytic residue) to each variant were as follows: 1) p.G256 = 8.7 Å, 2) p.R174 = 10.1 Å. 3) p.R286 = 11.7 Å; 4) p.R281 = 11.8 Å; 5) p.D201 = 15.3 Å; 6) p.F188 = 16.1 Å; 7) p.W195 = 16.8 Å; 8) p.T149 = 17.3 Å; 9) p.L156 = 26.1 Å.

To begin assessing evolutionary importance of each residue, conservation was estimated with orthologous/paralogous analysis. Across 162 LFNG orthologues, all sequences associated with the substitutions were highly conserved (Figure S3). Specifically, p.T149 (GERP = 3.95), p.L156 (GERP = 4.83), and p.W195 (GERP = 5.32) were ~90% conserved with no variation in amino acid residue (Uniref50; Figure S3B). Residues p.D201 (GERP = 5.32), p.T281 (GERP = 4.70), and p.R286 (GERP = 4.70) were ~100% conserved (Uniref50; Figure S3B). To further examine the hypomorphic p.T149I substitution, we assessed the paralogous conservation of Thr-149. Similar to Arg-174, Thr-149 was only ~80% conserved across three known fringe proteins (LFNG, Radical Fringe (RFNG), and Manic Fringe (MFNG), and their associated isoforms (Figure S3C)).

Finally, REVEL scores were obtained as a quantifiable metric of predicted pathogenicity (Table 3). REVEL scores range from 0.0 (benign) to 1.0 (pathogenic) and are an aggregate of 13 predictive variant pathogenicity algorithms (Ioannidis et al., 2016). In order of most predicted pathogenicity to least, the scores were as follows: 1) p.W195R = 0.9; 2) p.G256S = 0.89; 3)

p.D201N = 0.84; 4) p.L156R = 0.74; 5) p.R286W = 0.73; 6) p.T149I = 0.68; 7) p.T281K = 0.62; 8) p.R174H = 0.52; 9) p.F188L = 0.51.

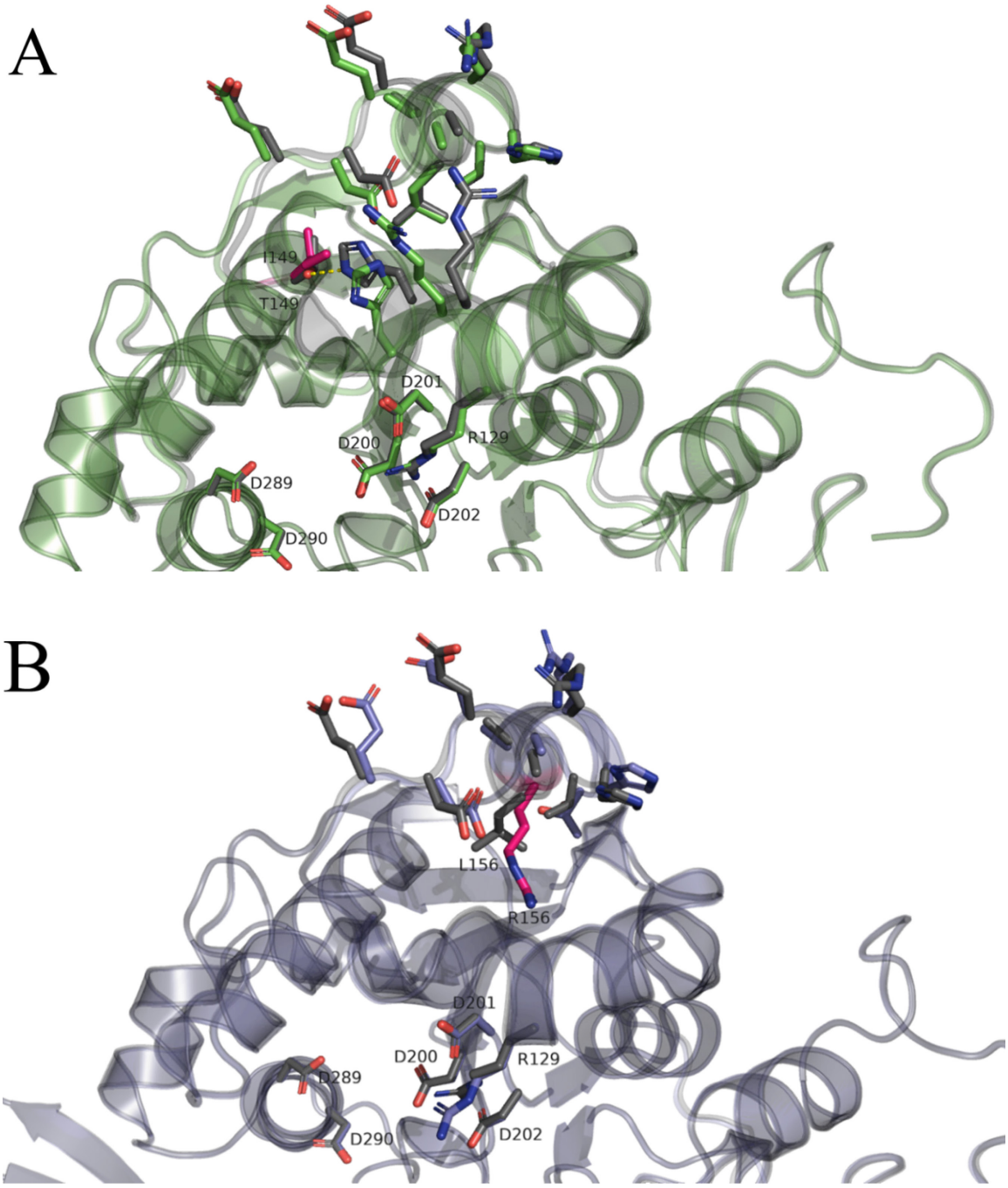


Figure 16: *In Silico* Structure Prediction of LFNG variants.

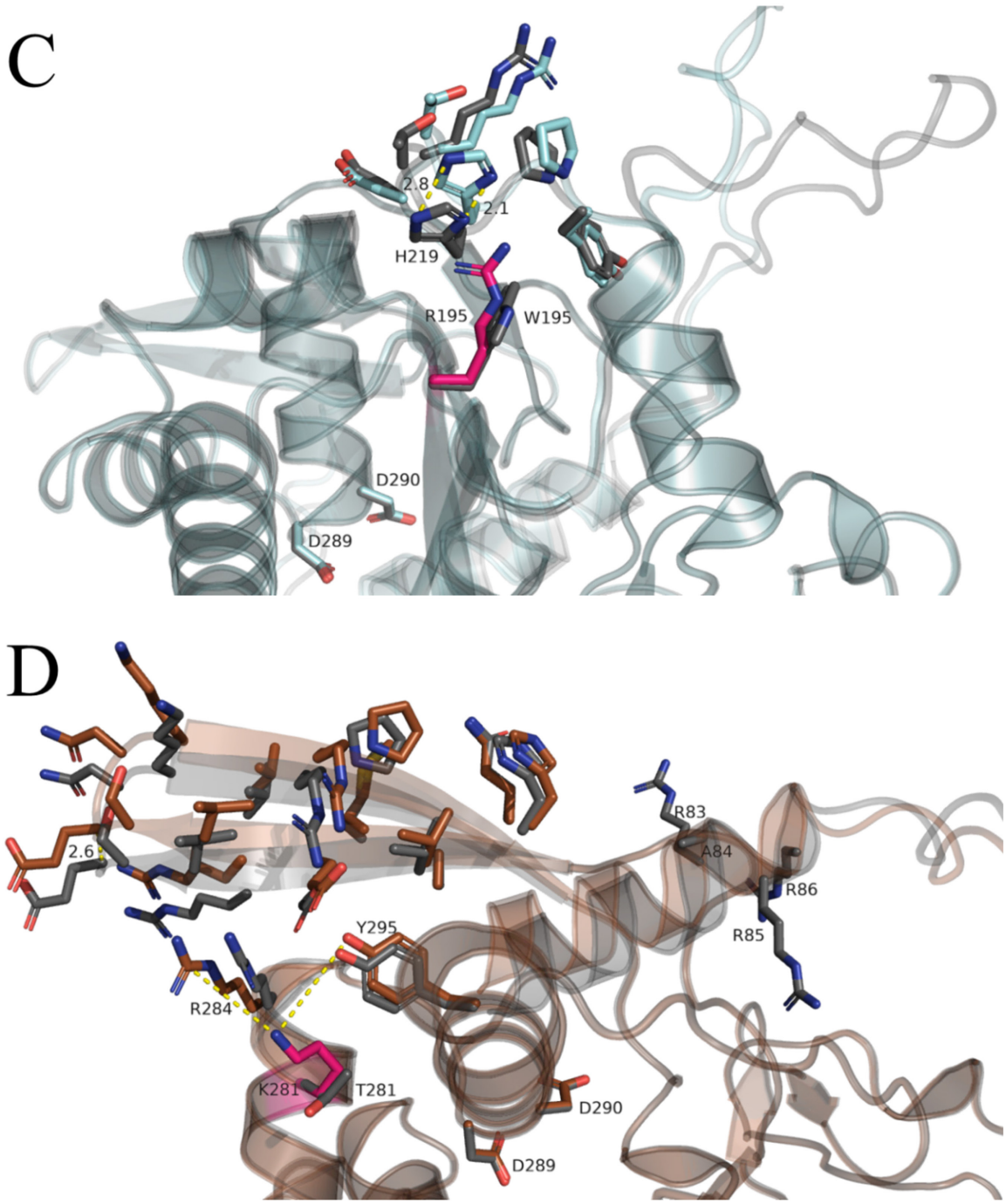


Figure 16: *In Silico* Structure Prediction of LFNG Variants

Figure 16: *In Silico* Structure Prediction of LFNG variants. Overlayed models of WT (grey) and variant (coloured) LFNG tertiary structures. In all cases, the referenced secondary structures are outlined with enlarged side chains. **A)** p.T149I is predicted to disrupt a hydrogen bond between p.H126, leading to destabilization of a region encompassing two alpha helices (residues p.D131 to p.H159). Variant structure overlayed in green. **B)** p.L156R similarly is predicted to affect this region but primarily shifts the alpha helix from p.D153 to p.T160. Variant structure overlayed in purple. **C)** p.W195R is predicted to induce a shift in p.H219 outwards from the protein core, destabilizing a loop (p.R195 to p.S216) without impacting flanking secondary structures. Variant structure overlayed in turquoise. **D)** p.T281K is predicted to exert pressure on p.R284 and p.Y295, leading to the predicted pivoting of an entire beta-hairpin (residues 234-250) away from the protein core, with each position shifted up to a maximum of 2.6 angstroms. Variant structure overlayed in orange. Critical residues: Active site, p.D290. DxD motif (UDP-GlcNAc/Mn^{2+} binding), p.D200-p.D201-p.D202. Mn^{2+} binding, p.H314.

3.4. Discussion

SCD is a monogenic condition caused by loss-of-function variants in six genes critical to somitogenesis: *DLL3* (SCD1), *HES7* (SCD2), *LFNG* (SCD3), *MESP2* (SCD4), *RIPPLY2* (SCD5), and *TBX6* (SCD6) (Nóbrega et al., 2021). Although subtle phenotypic variation exists between subtypes, variation within subtypes remains an area of minimal investigation. Genotype-phenotype correlations could lead to more accurate genetic counselling about prognosis and clinical management plans. Assessing the possibility of genotype-dependant phenotypic variance in SCD3 is particularly useful as *LFNG* variants can lead to diverse levels of functional inhibition (Luther et al., 2009; Wengryn et al., 2023). Therefore, we sought to investigate the functional effect of nine *LFNG* variants which cause SCD3 in the hopes of more deeply understanding the genotype-phenotype relationship. All variants except for c.766G>A (p.G256S) and c.521G>A (p.R174H) are misprocessed and mislocalized, and all variants except for c.521G>A (p.R174H) and c.446C>T (p.T149I) are enzymatically inactive (review Tables 2, 3). The American College of Medical Geneticists (ACMG) uses strict criteria to classify a variant as pathogenic (Brnich et al., 2019; Richards et al., 2015). This data satisfies PS3 (functional evidence via well-established functional assay) criteria for the pathogenicity of c.583T>C (p.W195R), c.842C>A (p.T281K), and c.446C>T (p.T149I), and further validates this distinction for the remaining six variants.

3.4.1. Mislocalization is a Common Etiology of SCD3

In order for LFNG to transfer Glc-NAc to NOTCH receptors it must be located within the Golgi lumen (Evrard et al., 1998; Johnston et al., 1997; Sparrow et al., 2006). Here, it acts alongside other critical glycosylating enzymes to potentiate appropriate NOTCH receptor-ligand

interactions. Without LFNG, the 'wrong' receptor-ligand interactions ensue, and the segmentation clock fails to oscillate (Chu et al., 2019; Kakuda et al., 2020; Kakuda & Haltiwanger, 2017; Nandagopal et al., 2018). Therefore, current understanding suggests that genotype may not explain phenotypic variability between SCD3 cases caused by the equally null, mislocalized variants.

As mislocalization is suspected to necessarily perturb LFNG activity, and thus cause the SCD3 phenotype, it is important to explore factors which might influence a variants effect on subcellular localization. One well known cause of mislocalization is variant-induced structural perturbation . Previous work has suggested that variants distal to LFNG's catalytic site (p.D290) or Mn^{2+} binding residues (DxD, p.D200-p.D201-p.D202) are more likely to structurally alter LFNG compared to those which are proximal (Luther et al., 2009). In support of this, the three variants closest to the catalytic site (1) p.G256S = 8.7 Å; 2) p.R174 = 10.1 Å; 3) p.R286 = 11.7 Å) were localized (p.G256S), localized (p.R174H), and partially localized (p.R286W), respectively (Figures 14, 15; Table 3). Furthermore, each of these localized variants were not predicted to affect structure whereas four of the six mislocalized variants were. Two variants which fail to support this notion include p.F188L and p.D201N. Although both variants were farther from the catalytic site, they were not predicted to affect structure and did not colocalize with the Golgi (Figure 14; Table 3) (Sparrow et al., 2006; Wengryn et al., 2023). There also does not appear to be a correlation between proximity to the DxD motif and localization (Table 3). Critically, the unique biochemical properties of each variant omit reasonably suggesting that variant proximity to the catalytic site alone could be more than a single contributing factor amongst many for mislocalization. Precise structural location, substitution, and local side-chain

interactions, discussed at length here and elsewhere, play very significant modulatory roles (Luther et al., 2009; Otomo et al., 2019; Schuhmann et al., 2021; Sparrow et al., 2006; Takeda et al., 2018; Wengryn et al., 2023). Finally, the inherent predictive limitations of AlphaFold and *in silico* software further limit the strength of this influence. Therefore, although promising, it remains unclear whether proximity to the catalytic site, structural disturbance, and mislocalization are linked. It is hoped that future investigations can utilize the *in vitro* data generated here to assess the possibility of such a contributing factor more thoroughly.

3.4.2. Mislocalization Prevents Processing and Secretion

The examination of protein processing is essential in understanding the impact of LFNG variants on the modulation of SCD3 presentation. This significance is underscored by the fact that impaired protein processing has been associated with the SCD3 phenotype in mouse models (Shifley & Cole, 2008; D. R. Williams et al., 2016). Nonetheless, earlier studies have indicated that aberrant processing might result from the protein being misplaced to an incorrect location, rather than being a separate occurrence. Indeed, data from current work supports this hypothesis as all mislocalized variants were not processed (Figure 13, 14). The basis for this discovery aligns with evidence showing that LFNG undergoes cleavage by the pre-pro-protein convertase SPC6, which is found exclusively within the Golgi (Shifley & Cole, 2008; D. R. Williams et al., 2016). Additionally, since these variant constructs are not released into the media (Table 3), they may not have entered the Golgi secretory pathway. This study provides strong support for the notion that the incapacity to reach the Golgi apparatus obstructs the processing and subsequent release of LFNG. Furthermore, it underscores the concept that the FRINGE domain is the

fundamental reason behind this unsuccessful translocation, given that the transmembrane domain was deliberately excluded from this experiment (Otomo et al., 2019; Takeda et al., 2018).

Importantly, the results associated with p.R286W provide additional confirmation that mislocalization impedes both protein processing and secretion. In our analysis of p.R286W localization, we noticed distinct LFNG-Golgi signal overlap in 24/40 cells (see Figure 14J for example) and clear signs of misplacement 16/40 (see Figure 14K for example). Furthermore, colocalized cells possessed a diffuse LFNG signal characteristic of mislocalized cells (Figure 14B-C, E-F, G, I, K) in contrast to typical localized conditions here (Figure 14A, D, H) and elsewhere (Shifley & Cole, 2008; Sparrow et al., 2006; Wengryn et al., 2023). These observations remained consistent across various experimental replicates. As we delved into the analysis of protein processing, it became clear that p.R286W underwent cleavage at a significantly slower pace compared to p.R174H and p.G256S, with roughly 30 times less processed protein (Figure 13D). Evidently, because there was a smaller population of cells expressing properly localized LFNG (Figure 14J-K), a diminished number of cells were able to produce cleaved protein, which explains both sets of findings (Figure 13D). These findings harmonize with previous research that suggested p.R286W was secreted with both the artificial and endogenous signal peptides (Takeda et al., 2018). In this study, the prolonged multi-day incubation period would have allowed the partially functional localization mechanism enough time to release an adequate quantity of LFNG into the culture medium. However, the mechanism by which partial localization occurs remains unexplained. The possibility of plasmid diversity can be reasonably dismissed, as all plasmid samples underwent sequencing and validation before transfection. Additionally, disproportional variations in cell-to-cell transfection efficiencies are

unlikely to be the root cause, as all other experimental conditions consistently yielded conclusive results across all experiments. Finally, although Z-stacking was unable to completely prevent artifactual colocalization, p.R286W had far more cells associated with both patterns and this overlap did not appear artifactual in origin (Figure 14 legend; Figure S2).

In silico, the distinctive patterns observed in p.R286W hint at nuanced spatial dynamics in a potentially unrecognized region. Examination of the p.R286W variant reveals that it has remarkable proximity to variants p.T281K and p.R174H. Unlike p.T281K, which exhibits complete mislocalization, and p.R174H, which localizes to the Golgi, p.R286W presents a partial colocalization pattern in some cells (Figure 14J) alongside clear displacement in others (Figure 14K). Examining this ambiguous region in more detail may aid in elucidating residues or domains critical to the intracellular translocation of LFNG.

3.4.3. Residual Glycosyltransferase Activity of p.T149I Does Not Modulate Phenotype

Previous research has primarily concentrated on Glc-NAc-transferase activity when evaluating the functional impact of LFNG variants (Otomo et al., 2019; Sparrow et al., 2006; Takeda et al., 2018). This is because, without glycosyltransferase activity, LFNG cannot exert an effect on NOTCH receptors, and consequently, it cannot influence the segmentation clock. We report the sixth and seventh variants associated with complete nullification of LFNG enzymatic activity (c.583T>C [p.W195R]; c.842C>A [p.T281K]) (Figure 15; Table 3). Biochemically, these variants are predicted to exhibit significant structural changes. Protein modeling suggests that p.W195R shifts His-219 outward from the protein core (Figure 16C), whereas in p.T281K, the

beta-hairpin from residues 234-250 undergoes a significant shift (Figure 16D). Both of these changes can reasonably be attributed to the complete nullification of Glc-NAc-transferase activity.

Although many variants lead to complete functional knockout (Table 3), previous work has characterized a variant associated with SCD3 with residual Glc-NAc-transferase activity (p.R174H) (Wengryn et al., 2023). Here, we report the second partially active variant associated with an SCD3 phenotype, c.446C>T (p.T149I) (Figure 15B; Table 3). Although both variants lead to similar levels of functional activity, they differ in structural perturbation and effect on localization. Primarily, *in silico* analysis suggest that the variants affect different regions of the enzyme (Figure 16A; Table 3) as structural changes are only associated with p.T149I and not p.R174H. Secondly, unlike p.R174H, p.T149I is mislocalized and therefore would not be able to act on NOTCH receptors with its residual activity. Indeed, the phenotype of the proband with homozygous c.446C>T variants has very few vertebrae (<11) and ribs (<10 bilaterally), significant kyphoscoliosis, and pronounced pebble beach sign. This suggests that this patient's phenotype may not have been modulated by residual LFNG activity.

3.4.4. Trunk Length may be Sensitive to Residual LFNG Activity

Our findings suggest that six of the seven previously reported probands carry biallelic null *LFNG* variants. Except for p.R174H, all other variants display mislocalization and/or complete loss of enzymatic activity, preventing GlcNAc transfer to NOTCH receptors (See Table 3;Sparrow et al., 2006). Clinical data is incomplete for the six probands (Table S2; Lefebvre et al., 2018; Otomo et al., 2019; Schuhmann et al., 2021; Sparrow et al., 2006; Takeda et al., 2018) but heights trend below the 5th percentile, and vertebral numbers are typically fewer than 17,

with variable yet often severe scoliosis presentation. The seventh proband is compound heterozygous for p.R174H and null p.G256S LFNG alleles, presenting a different axial phenotype with more vertebrae and a longer trunk, perhaps due to relatively increased LFNG activity compared to null genotypes (Wengryn et al., 2023). Supporting this possibility of genotype-dependant changes of the axial skeleton, murine models which utilized *LFNG*-dosage as a proxy for varied LFNG function demonstrated that deceasing dosages correlated with progressively fewer somites and vertebrae (Oginuma et al., 2010; Stauber et al., 2009; D. R. Williams et al., 2014). However, this hypothesis does not explain additional SCD3 features like arachno/camptodactyly and auditory abnormalities (Lecca et al., 2023; Sparrow et al., 2006). It also does not account for individual genomic variation which may influence the genotype-phenotype relationship. Future work assessing the impact of *LFNG* variants with different phenotypic manifestations in iPSC-derived organoid-development models (Matsuda et al., 2020; Miao et al., 2023; Sanaki-Matsumiya et al., 2022) may be able to determine which, if any, factors influence the varied presentation of these additional phenotypes. Correlating this with a quantitative assessment of the variants effect on NOTCH activation in isolation with a previously utilized assay (Sparrow et al., 2006) may allow for further elucidation of factors which contribute to the genotype-phenotype relationship. Crucially, more probands will need to be identified to draw any conclusions regarding the expansion of the SCD3 phenotypic spectrum.

3.4.5. Limitations

Primarily, assessing protein processing and localization via transient and unstable transfections introduces internal reliability concerns, particularly evident in the quantity and standard error of intracellular unprocessed LFNG-HA (Figure 12B). In contrast to previous

research, we observed no significant increase in unprocessed protein levels for p.F188L compared to WT (Figure 12A-B; (Wengryn et al., 2023). Processed protein data (Figure 12C-D) was more internally reliable and therefore appeared dependable in determining whether *LFNG* variants inhibit processing. However, statistical variability was still evident in this experiment, particularly in the localized, processed samples (p.R174H and p.G256S), where there may appear to be a qualitative increase in LFNG-HA compared to WT (Figure 12D). Although not suggested by statistical analysis, future investigations harnessing different methods to further understand LFNG-processing should remain open to the possibility of novel, more nuanced mechanisms.

Secondly, this study was unable to provide evidence regarding the particular subcellular compartment(s) to which LFNG is mislocalized. While mislocalization likely hinders LFNG's ability to support somitogenesis, particularly given that each mislocated variant is enzymatically inactive, it remains unclear to what degree an enzymatically active and mislocalized variant would impact somitogenesis.

Thirdly, evaluating LFNG function under standardized conditions ignores Golgi pH variations that can affect LFNG activity (Rampal et al., 2005). Since pH variability might influence LFNG glycosyltransferase activity *in vivo*, investigating whether pH fluctuations partially rescue or further diminish enzyme activity could provide insights into LFNG's Golgi lumen regions that act on NOTCH. This protocol and others (Luther et al., 2009; Otomo et al., 2019; Takeda et al., 2018; Wengryn et al., 2023) may also cause protein misfolding due to the extended incubations at 37°C (three-day expression period followed by either 30-minute reaction or four-hour reaction). However, previous work did not identify misfolding with other LFNG variants after 72 hours and 30 minutes (Luther et al., 2009), and thus the additional three hours

and 30 minutes may not substantially impact protein folding. Irrespective of likelihood, future work should be aware of this factor when conducting glycosyltransferase assays and interpreting their results.

Finally, inconsistent phenotypic reporting hampers accurate genotype-phenotype interpretations. Gathering comprehensive clinical data including, but not limited to: age, sex, consanguinity, height (absolute and relative), arm-span to height ratio, mid-parental height, vertebral body number, rib number, scoliosis presence/cobb angle, and descriptions of both axial and non-axial characteristics (Table S2) is vital for understanding the SCD3 genotype-phenotype relationship. To this end, a patient registry can aid data collection, fostering collaboration and a deeper disease understanding, enhancing molecular data interpretation across related studies.

3.4.6. Future Directions

A significant gap in LFNG research remains a focussed, detailed investigation of the subcellular processes which govern its subcellular transport. Although we previously reported that LFNG mislocalizes to the cytoplasm due to its diffuse distributive pattern (Wengryn et al., 2023), we have been unable to conclusively illustrate the specific subcellular compartment to which the variants mislocalize. An alterative explanation is that variant LFNG is actually mislocalized to the endoplasmic reticulum (ER), as the diffuse pattern is characteristic of ER-resident proteins (Figure 14B-C, F-G, I-K; (Silveira et al., 2023; Tsai et al., 2021). One explanatory mechanism could be misfolding in response to amino-acid substitutions (Hegde & Zavodszky, 2019). Indeed, previous evidence from disease causing variants in *Glycosyltransferase 8 Domain–Containing Protein 1* (*GLT8D1*), a GT-A fold family, DxD motif-containing glycosyltransferase with a type-II transmembrane domain with similarities to *LFNG*,

has shown misfolding and persistence in the ER (Tsai et al., 2021). Since misfolding can arise from substitutions in various regions throughout the protein, findings corroborating this hypothesis may explain the wide spatial distribution of variants that result in mislocalization (Table 3).

Future work focussing on the underlying principles and mechanisms of LFNG transport and mislocalization may address this gap and enhance current understanding of these critical subcellular processes. Such objectives may involve identifying the compartment(s) to which LFNG mislocalizes, exploring the mechanisms and factors involved in LFNG folding, determining the protein(s) responsible for transporting LFNG to the Golgi apparatus (and where these transport proteins bind LFNG), and investigating the potential role of the N-terminal in translocation. While LFNG mislocalization is recognized as a molecular cause of SCD3, irrespective of its exact location, a more in-depth inquiry could reveal unidentified factors influencing the SCD3 phenotype.

An overarching difficulty from such a project would be determining the extent to which LFNG misfolding is a consequence of transfection protocol and cell-line. To date, transient transfection of NIH-3T3 cells has been the exclusively utilized for the analysis of LFNG localization (Sparrow et al., 2006; Wengryn et al., 2023). Primarily, endogenous folding machinery can become overwhelmed by the transient expression of exogenous plasmid DNA, potentially exaggerating misfolding (Gibson et al., 2013). This could make it difficult to determine a mechanism to explain the subtle mislocalization differences noted here (e.g. p.R286W; Figure 14J, K). Secondly, the protein transport mechanisms, folding capacity, and responses to misfolding unique to each cell line need to be considered and evaluated before

externalizing results (Brnich et al., 2019; Hegde & Zavodszky, 2019). Even with stable transfection, the differences between the effect of LFNG variant from cell-to-cell is unknown. Therefore, culturing physiologically relevant cell-lines, aggregating data from many cell lines and transfection methods, and/or utilizing CRISPR-based mutagenesis of endogenous *LFNG* may help to separate artifactual and genuine signals, improving external viability.

3.5. Conclusions

In this work, we fully or partially characterized the p.T149I, p.L156R, p.W195R, p.D201N, p.T281K, and p.R286W LFNG variants. We also replicated functional analysis for p.R174H, p.F188L, and p.G256S. In this set of samples, we conclude that p.G256S and p.R174H are the only appropriately localized and processed LFNG variants and p.R174H and p.T149I are the only variants with residual glycosyltransferase activity. We also conclude that protein mislocalization resulting in impaired protein processing and secretion is a common disease mechanism causing SCD3. *In silico* LFNG variant modelling was unable to conclusively identify a mechanism for this phenomenon, although proximity to the catalytic residue remains promising. Finally, we hypothesize that the axial phenotype of SCD3 may be closely linked with genotype and not other influences. It is hoped that future work will gather updated clinical data from each of the probands and examine *LFNG* variants in the context of iPSC-derived organoid models. Such work may allow for a more precise investigation of the relationship between SCD3 genotype and phenotype.

3.6. Data Availability

The data that support the findings of this study are available in the methods and/or supplementary material of this article. Materials are available upon request to the corresponding author.

3.7. Conflict of Interest

The authors report no conflicts of interest.

3.8. Ethical Approval

This research did require ethical approval as per TCPS-2 (2022) Article 12.21, the University of Alberta Research Ethics Board.

3.9. Acknowledgements

The authors would primarily like to thank each of the families and patients for their significant and crucial contribution to the study of SCD3. The authors also sincerely thank both reviewers for their expert insight; Reviewer 1 for their detailed suggestions surrounding LFNG localization and Reviewer 2 for their comprehensive perspective that helped unify this work. P.W. would like to thank Dr. Toshifumi Yokota and Dr. Fred Berry for their scientific insight, encouragement, and constructive commentary throughout his graduate studies. P.W. would also like to thank Dr. Darryl Rolfson and Dr. Roger Epp for their approval of his request to continue research

throughout his medical studies. This work was supported by grants from the Women and Children's Health Research Institute at the University of Alberta, the Rare Disease Foundation, the Shriners Scoliosis Research Chair, and in part by a Grant-in-Aid for Scientific Research (C) 23K06142 (S.M.) from the Japan Society for the Promotion of Science, Japan, and a Grant-in-Aid for the Research Center for Pathogenesis of Intractable Diseases from the Research Institute of Meijo University (S.M. and S.Y.). This work was also supported by the Alberta Innovates Graduate Student Scholarship, the Alberta Graduate Excellence Scholarship, and the Graduate Entrance Scholarship.

3.10. Bibliography

Brnich, S. E., Abou Tayoun, A. N., Couch, F. J., Cutting, G. R., Greenblatt, M. S., Heinen, C. D., Kanavy, D. M., Luo, X., McNulty, S. M., Starita, L. M., Tavtigian, S. V., Wright, M. W., Harrison, S. M., Biesecker, L. G., Berg, J. S., Abou Tayoun, A. N., Berg, J. S., Biesecker, L. G., Brenner, S. E., … Topper, S. (2019). Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. *Genome Medicine, 12*(1), 1–12. https://doi.org/10.1186/s13073-019-0690-2

Chu, L. F., Mamott, D., Ni, Z., Bacher, R., Liu, C., Swanson, S., Kendziorski, C., Stewart, R., & Thomson, J. A. (2019). An In Vitro Human Segmentation Clock Model Derived from Embryonic Stem Cells. *Cell Reports, 28*(9), 2247-2255.e5. https://doi.org/10.1016/j.celrep.2019.07.090

Dequéant, M. L., & Pourquié, O. (2008). Segmental patterning of the vertebrate embryonic axis. *Nature Reviews Genetics, 9*(5), 370–382. https://doi.org/10.1038/nrg2320

Dunwoodie, S. L. (2009). Mutation of the fucose-specific β1,3 N-acetylglucosaminyltransferase LFNG results in abnormal formation of the spine. *Biochimica et Biophysica Acta - Molecular Basis of Disease, 1792*(2), 100–111. https://doi.org/10.1016/j.bbadis.2008.11.003

Evrard, Y. A., Lun, Y., Aulehla, A., Gan, L., & Johnson, R. L. (1998). Lunatic fringe is an essential mediator of somite segmentation and patterning. *Nature, 394*(6691), 377–381. https://doi.org/10.1038/28632

Gibson, T. J., Seiler, M., & Veitia, R. A. (2013). The transience of transient overexpression. *Nature Methods, 10*(8), 715–721. https://doi.org/10.1038/nmeth.2534

Hegde, R. S., & Zavodszky, E. (2019). Recognition and degradation of mislocalized proteins in health and disease. *Cold Spring Harbor Perspectives in Biology, 11*. https://doi.org/10.1101/cshperspect.a033902

Ioannidis, N. M., Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., Musolf, A., Li, Q., Holzinger, E., Karyadi, D., Cannon-Albright, L. A., Teerlink, C. C., Stanford, J. L., Isaacs, W. B., Xu, J., Cooney, K. A., Lange, E. M., Schleutker, J., Carpten, J. D., … Sieh, W. (2016). REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *American Journal of Human Genetics, 99*(4), 877–885. https://doi.org/10.1016/j.ajhg.2016.08.016

Johnston, S. H., Rauskolb, C., Wilson, R., Prabhakaran, B., Irvine, K. D., & Vogt, T. F. (1997). A family of mammalian Fringe genes implicated in boundary determination and the Notch pathway. *Development, 124*(11), 2245–2254. https://doi.org/https://doi.org/10.1242/dev.124.11.2245

Kakuda, S., & Haltiwanger, R. S. (2017). Deciphering the Fringe-mediated Notch Code: Identification of activating and inhibiting sites allowing discrimination between ligands. *Developmental Cell, 40*(2), 193–201. https://doi.org/https://doi.org/10.1016/j.devcel.2016.12.013

Kakuda, S., LoPilato, R. K., Ito, A., & Haltiwanger, R. S. (2020). Canonical notch ligands and fringes have distinct effects on NOTCH1 and NOTCH2. *Journal of Biological Chemistry, 295*(43), 14710–14722. https://doi.org/10.1074/jbc.RA120.014407

Lecca, M., Bedeschi, M. F., Izzi, C., Dordoni, C., Rinaldi, B., Peluso, F., Caraffi, S. G., Prefumo, F., Signorelli, M., Zanzucchi, M., Bione, S., Ghigna, C., Sassi, S., Novelli, A., Valente, E. M., Superti-Furga, A., Garavelli, L., & Errichiello, E. (2023). Identification of bi-allelic LFNG variants in three patients and further clinical and molecular refinement of spondylocostal dysostosis 3. *Clinical Genetics, 2023*, 1–8. https://doi.org/10.1111/cge.14336

Lefebvre, M., Dieux-Coeslier, A., Baujat, G., Schaefer, E., Judith, S. O., Bazin, A., Pinson, L., Attie-Bitach, T., Baumann, C., Fradin, M., Pierquin, G., Julia, S., Quélin, C., Doray, B., Berg, S., Vincent-Delorme, C., Lambert, L., Bachmann, N., Lacombe, D., … Thevenon, J. (2018). Diagnostic strategy in segmentation defect of the vertebrae: a retrospective study of 73 patients. *Journal of Medical Genetics, 55*(6), 422–429. https://doi.org/10.1136/jmedgenet-2017-104939

Luther, K. B., Schindelin, H., & Haltiwanger, R. S. (2009). Structural and mechanistic insights into lunatic fringe from a kinetic analysis of enzyme mutants. *Journal of Biological Chemistry, 284(*5), 3294–3305. https://doi.org/10.1074/jbc.M805502200

Matsuda, M., Yamanaka, Y., Uemura, M., Osawa, M., Saito, M. K., Nagahashi, A., Nishio, M., Guo, L., Ikegawa, S., Sakurai, S., Kihara, S., Maurissen, T. L., Nakamura, M., Matsumoto, T., Yoshitomi, H., Ikeya, M., Kawakami, N., Yamamoto, T., Woltjen, K., …

Alev, C. (2020). Recapitulating the human segmentation clock with pluripotent stem cells. *Nature, 580*(7801), 124–129. https://doi.org/10.1038/s41586-020-2144-9

Matsumoto, K., Luther, K. B., & Haltiwanger, R. S. (2021). Diseases related to Notch glycosylation. *Molecular Aspects of Medicine, 79*. https://doi.org/https://doi.org/10.1016/j.mam.2020.100938

Miao, Y., Djeffal, Y., De Simone, A., Zhu, K., Lee, J. G., Lu, Z., Silberfeld, A., Rao, J., Tarazona, O. A., Mongera, A., Rigoni, P., Diaz-Cuadros, M., Song, L. M. S., Di Talia, S., & Pourquié, O. (2023). Reconstruction and deconstruction of human somitogenesis in vitro. *Nature, 614*, 500–508. https://doi.org/10.1038/s41586-022-05655-4

Moloney, D. J., Panin, V. M., Johnston, S. H., Chen, J., Shao, L., Wilson, R., Wang#, Y., Stanley, P., Irvine, K. D., Haltiwanger, R. S., & Vogt, T. F. (2000). Fringe is a glycosyltransferase that modifies Notch. *Nature, 406*, 369–375. https://doi.org/https://doi.org/10.1038/35019000

Musumeci, G., Castrogiovanni, P., Coleman, R., Szychlinska, M. A., Salvatorelli, L., Parenti, R., Magro, G., & Imbesi, R. (2015). Somitogenesis: From somite to skeletal muscle. *Acta Histochemica, 117*, 313–328. https://doi.org/10.1016/j.acthis.2015.02.011

Nandagopal, N., Santat, L. A., LeBon, L., Sprinzak, D., Bronner, M. E., & Elowitz, M. B. (2018). Dynamic Ligand Discrimination in the Notch Signaling Pathway. *Cell, 172*(4), 869-880.e19. https://doi.org/10.1016/j.cell.2018.01.002

Nóbrega, A., Maia-Fernandes, A. C., & Andrade, R. P. (2021). Altered Cogs of the Clock: Insights into the Embryonic Etiology of Spondylocostal Dysostosis. *Journal of Developmental Biology, 9*(5), 1–14. https://doi.org/10.3390/jdb9010005

Oginuma, M., Takahashi, Y., Kitajima, S., Kiso, M., Kanno, J., Kimura, A., & Saga, Y. (2010). The oscillation of Notch activation, but not its boundary, is required for somite border formation and rostral-caudal patterning within a somite. *Development, 137*(9), 1515–1522. https://doi.org/10.1242/dev.044545

Otomo, N., Mizumoto, S., Lu, H. F., Takeda, K., Campos-Xavier, B., Mittaz-Crettol, L., Guo, L., Takikawa, K., Nakamura, M., Yamada, S., Matsumoto, M., Watanabe, K., & Ikegawa, S. (2019). Identification of novel LFNG mutations in spondylocostal dysostosis. *Journal of Human Genetics, 64*(3), 261–264. https://doi.org/10.1038/s10038-018-0548-2

Rampal, R., Li, A. S. Y., Moloney, D. J., Georgiou, S. A., Luther, K. B., Nita-Lazar, A., & Haltiwanger, R. S. (2005). Lunatic fringe, manic fringe, and radical fringe recognize similar specificity determinants in O-fucosylated epidermal growth factor-like repeats. *Journal of Biological Chemistry, 280*(51), 42454–42463. https://doi.org/10.1074/jbc.M509552200

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., & Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine, 17*(5), 405–424. https://doi.org/10.1038/gim.2015.30

Sanaki-Matsumiya, M., Matsuda, M., Gritti, N., Nakaki, F., Sharpe, J., Trivedi, V., & Ebisuya, M. (2022). Periodic formation of epithelial somites from human pluripotent stem cells. *Nature Communications, 13*(2325), 1–14. https://doi.org/10.1038/s41467-022-29967-1

Schuhmann, S., Koller, H., Sticht, H., Kraus, C., Krumbiegel, M., Uebe, S., Ekici, A. B., Reis, A., & Thiel, C. T. (2021). Clinical and molecular delineation of spondylocostal dysostosis type 3. *Clinical Genetics, 99*(6), 851–852. https://doi.org/10.1111/cge.13952

Shifley, E. T., & Cole, S. E. (2008). Lunatic fringe protein processing by proprotein convertases may contribute to the short protein half-life in the segmentation clock. Biochimica et *Biophysica Acta - Molecular Cell Research, 1783*(12), 2384–2390. https://doi.org/10.1016/j.bbamcr.2008.07.009

Shifley, E. T., VanHorn, K. M., Perez-Balaguer, A., Franklin, J. D., Weinstein, M., & Cole, S. E. (2008). Oscillatory lunatic fringe activity is crucial for segmentation of the anterior but not posterior skeleton. *Development, 135*(5), 899–908. https://doi.org/10.1242/dev.006742

Silveira, K. C., Fonseca, I. C., Oborn, C., Wengryn, P., Ghafoor, S., Beke, A., Dreseris, E. S., Wong, C., Iacovone, A., Soltys, C. L., Babul-Hirji, R., Artigalas, O., Antolini-Tavares, A., Gingras, A. C., Campos, E., Cavalcanti, D. P., & Kannu, P. (2023). CYP26B1-related disorder: expanding the ends of the spectrum through clinical and molecular evidence. *Human Genetics, 142*(11), 1571–1586. https://doi.org/10.1007/s00439-023-02598-2

Sparrow, D. B., Chapman, G., Wouters, M. A., Whittock, N. V., Ellard, S., Fatkin, D., Turnpenny, P. D., Kusumi, K., Sillence, D., & Dunwoodie, S. L. (2006). Mutation of the Lunatic

Fringe gene in humans causes spondylocostal dysostosis with a severe vertebral phenotype. *American Journal of Human Genetics, 78*(1), 28–37. https://doi.org/10.1086/498879

Stauber, M., Sachidanandan, C., Morgenstern, C., & Ish-Horowicz, D. (2009). Differential axial requirements for Lunatic fringe and Hes7 transcription during mouse somitogenesis. *PLoS ONE, 4*(11). https://doi.org/10.1371/journal.pone.0007996

Takeda, K., Kou, I., Mizumoto, S., Yamada, S., Kawakami, N., Nakajima, M., Otomo, N., Ogura, Y., Miyake, N., Matsumoto, N., Kotani, T., Sudo, H., Yonezawa, I., Uno, K., Taneichi, H., Watanabe, K., Shigematsu, H., Sugawara, R., Taniguchi, Y., … Ikegawa, S. (2018). Screening of known disease genes in congenital scoliosis. Molecular Genetics and *Genomic Medicine, 6*(6), 966–974. https://doi.org/10.1002/mgg3.466

Tsai, P. C., Jih, K. Y., Shen, T. Y., Liu, Y. H., Lin, K. P., Liao, Y. C., & Lee, Y. C. (2021). Genetic and Functional Analysis of Glycosyltransferase 8 Domain–Containing Protein 1 in Taiwanese Patients with Amyotrophic Lateral Sclerosis. *Neurology: Genetics, 7*(6), e627. https://doi.org/10.1212/NXG.0000000000000627

Venzin, O. F., & Oates, A. C. (2020). What are you synching about? Emerging complexity of Notch signaling in the segmentation clock. *Developmental Biology, 460*(1), 40–54. https://doi.org/10.1016/J.YDBIO.2019.06.024

Wengryn, P., Silveira, K. da C., Oborn, C., Soltys, C.-L., Beke, A., Chacon-Fonseca, I., Damseh, N., Rodriguez, M. Q., Badilla-Porras, R., & Kannu, P. (2023). Functional

Characterization of Novel Lunatic Fringe Variants in Spondylocostal Dysostosis Type-III with Scoliosis. *Human Mutation, 2023*, 1–12. https://doi.org/10.1155/2023/5989733

Williams, D. R., Shifley, E. T., Braunreiter, K. M., & Cole, S. E. (2016). Disruption of somitogenesis by a novel dominant allele of Lfng suggests important roles for protein processing and secretion. *Development, 143*, 822–830. https://doi.org/10.1242/dev.128538

Williams, D. R., Shifley, E. T., Lather, J. D., & Cole, S. E. (2014). Posterior skeletal development and the segmentation clock period are sensitive to Lfng dosage during somitogenesis. *Developmental Biology, 388*(2), 159–169. https://doi.org/10.1016/j.ydbio.2014.02.006

Williams, S., Alkhatib, B., & Serra, R. (2019). Development of the axial skeleton and intervertebral disc. In *Current topics in developmental biology* (pp. 50–82). Elsevier. https://doi.org/10.1016/bs.ctdb.2018.11.018

Yoshioka-Kobayashi, K., Matsumiya, M., Niino, Y., Isomura, A., Kori, H., Miyawaki, A., & Kageyama, R. (2020). Coupling delay controls synchronized oscillation in the segmentation clock. Nature, 580(7801), 119–123. https://doi.org/10.1038/s41586-019

Chapter 4: Conclusions

The symmetrical organization of the vertebral column is critical for walking, standing, load bearing, and breathing, amongst many other vital physiological features. In disrupting this anatomy, scoliosis can lead to poor quality of life, pain, and even mortality (Hawes & Weinstein, 2003; Kebaish et al., 2011; Kwon et al., 2021; Sung et al., 2021). Although 1-3% of the population suffers from this condition, most cases are of idiopathic etiology, hindering diagnosis, prognosis, and treatment. Studying diseases which affect vertebral development *in utero*, like SCD3, provides an opportunity to reveal new genetic factors that could contribute to scoliosis etiology. Therefore, the goal of this thesis was to investigate *LFNG* variant(s) and SCD3 presentation in the hopes of deriving a more complete relationship between genotype and phenotype. Specific conclusions will be reviewed first and their implications will follow.

The first goal of this work was to determine whether the *LFNG* variants studied here satisfy PS3 ACMG criteria. I conclude that the novel variants identified in Chapter 2, c.521G>A (p.R174H) and c.766G>A (p.R174H) (Wengryn et al., 2023), fulfill PS3 ACMG criteria due to hypomorphic and null glycosyltransferase activity, respectfully. This elevates these variants to ACMG "likely pathogenic" status, allowing for these variants to be utilized during reproductive risk counselling and any further clinical management (Richards et al., 2015). However, skepticism should be maintained for the pathogenicity of c.521G>A as it was identified in a compound heterozygous proband. It is unknown whether homozygous c.521G>A variants would lead to SCD3. In Chapter 3, I conclude that the previously described c.583T>C (p.W195R) (Lefebvre et al., 2018), c.842C>A (p.T281K) (Lefebvre et al., 2018), and c.446C>T (p.T149I) (Schuhmann et al., 2021) variants also satisfy PS3 ACMG criteria due to null glycosyltransferase activity, null glycosyltransferase activity, and mislocalization, respectfully. However, these

variants were not classified under ACMG guidelines, and there lacks enough clinical data (Table S2) to categorize them here. It is therefore unknown whether the PS3 data would affect their classification. I hypothesize that these three variants are pathogenic or likely pathogenic due to the pronounced SCD3 phenotype and the novel molecular data.

The second goal of this thesis aimed to elucidate the mechanism(s) that each *LFNG* variant inhibits function by assessing glycosyltransferase activity, protein processing, and localization. Primarily, I conclude that all variants tested here, except for c.446T (p.T149I) (Schuhmann et al., 2021) and c.521G>A (p.R174H) (Wengryn et al., 2023) completely inhibit glycosyltransferase activity (Table 3). These two alleles, which retain partial activity, are the first to be associated with human SCD3 (the significance of which is discussed below). Secondly, I conclude that all variants tested here, except for c.521G>A (p.174H) (Wengryn et al., 2023), c.766G>C (p.G256S) (Wengryn et al., 2023), and c.856C>T (p.R256W) (Takeda et al., 2018), are not processed. Furthermore, the same three variants are mislocalized from the Golgi whereas the remaining variants are localized to the Golgi (Table 3). This correlation suggests that the processing and localization experiments are different methods which evaluate LFNG trafficking. Specifically, testing for cleavage with western blot is an indirect measure (i.e. the result of inappropriate trafficking) whereas immunofluorescence microscopy is a direct measure (i.e. the cause of null processing). The c.856C>T (p.R256W) (Takeda et al., 2018) variant provides further support for this connection as it is partially localized and borderline unprocessed.

The third goal of this work was to determine whether there is a link between variant location and type of functional alteration. The data generated here suggests that there may be a relationship as variants proximal to the DxD motif (p.D201) are more likely to disturb

glycosyltransferase activity but not localization, whereas those distal from the DxD motif perturb activity and localization (Table 3). Based on work here and elsewhere (Luther et al., 2009; Tsai et al., 2021), I propose the following hypotheses: 1) Pathogenic variants distal to the active site cause LFNG to fold abnormally. Misfolding acts to prevent the biochemical interactions necessary for active site function and promotes ER-retention. 2) Pathogenic variants proximal to the active site do not change the overall structure of LFNG; instead, they cause changes to the active site, abrogating normal catalytic activity. Combined, these hypotheses postulate that variant location predicts functional alteration, and in so doing, links variant location and *in silico* modelling to glycosyltransferase activity and LFNG trafficking. Future work will be necessary to test this hypothesis further.

The fourth aim of this work was to investigate whether a genotype-phenotype relationship exists for SCD3. I conclude that a link between *LFNG* variant and SCD3 presentation is likely; specifically, that the number of vertebral bodies (and thus the length of the trunk) may be modulated by the relative amount of gene activity conferred by specific variants. Of the seven SCD3 probands studied here, one has a hypomorphic *LFNG* genotype (conferred by the c.521G>A allele) (Wengryn et al., 2023) whereas the other six have null genotypes. Identified in Chapter 2, this proband is taller, possesses more vertebral bodies, and has a lower arm-span-to-height ratio compared to the others. This finding aligns with previous studies that found dosage-dependant trunk length increases in mouse models(Oginuma et al., 2010; Shifley et al., 2008; Stauber et al., 2009; D. R. Williams et al., 2014). Furthermore, *in vivo* and *in vitro* evidence strongly suggests that a wide, dosage-dependant range of LFNG activity leads to various degrees of segmentation defects (Cole et al., 2002; Riley et al., 2013; Shifley et al., 2008;

Shifley & Cole, 2008; Sparrow et al., 2012; D. R. Williams et al., 2014, 2016). Although many more probands and *LFNG* variants will need to be studied to test this further, it appears likely that within a set range of activity, different hypomorphic *LFNG* variants could lead to different vertebral numbers. The data here does not support that genotype correlates with additional abnormalities such as camptodactyly or hearing loss. Further investigation of *LFNG*'s role in internal ear and hand development is necessary to adequately examine the genotype-phenotype relationship in those structures.

It is hoped that the work conducted in this thesis leads to the beginning of a discussion surrounding the possibility that hypomorphic alleles could contribute to certain idiopathic scoliosis etiologies. Specifically, cases of scoliosis caused by a single SDV or slight changes to normal vertebral anatomy. The work conducted in this thesis provides molecular evidence of the first two *LFNG* variants which retain partial enzymatic activity (c.521G>A, c.446C>T) (Schuhmann et al., 2021; Wengryn et al., 2023) and the first which is partially localized (c.856C>T) (Takeda et al., 2018). Furthermore, this work provides a link between what appears to be a truly hypomorphic allele (c.521G>A) and an SCD3 proband with different vertebral anatomy compared to others with null genotypes. Therefore, and in the context of a NOTCH gene dosage-dependant range of segmentation deficiencies (Oginuma et al., 2010; Stauber et al., 2009; D. R. Williams et al., 2014), genotypes with threshold activity could contribute to some cases of idiopathic scoliosis.

The conclusions provided here are subject to a variety of limitations (sections 2.4 and 3.4), the most significant of which is the utilization of a small sample size as the foundation for generalizable conclusions. Although more data has been published since the initiation of this

thesis, only 14 SCD3 probands have been published with radiographic and genetic findings as of 2024 (Lecca et al., 2023; Lefebvre et al., 2018; Otomo et al., 2019; Schuhmann et al., 2021; Sparrow et al., 2006; Takeda et al., 2018; Wengryn et al., 2023). Until a normal curve can be established (typically N = 30), it is difficult to derive a reliable and viable relationship between genotype and phenotype.

In addition to the low sample size, data gathering is inconsistent (Table 3, Table S2). Chapter 3 addressed many molecular gaps by performing *in vitro* analysis for all of the *LFNG* variants published until 2021, although there still lacks uniform presentations of clinical data. Without knowing each phenotypic characteristic (see Table S2), the molecular results cannot be generalized with significant reliability. The creation of a database where each finding can be deposited in an organized fashion may help to increase the reliability and viability of future projects. This would also likely allow for the categorization of these variants under ACMG criteria and a more thorough understanding of how LFNG activity relates to phenotype (Richards et al., 2015). Such work could lead to a more accurate prognosis of SCD3 severity, potentially enhancing reproductive risk counselling and supportive management.

Another important limitation is the lack of LFNG trafficking research. It is currently unknown how LFNG is transported from the ER to the Golgi, why some variants lead to mislocalization, and which factors maintain LFNG's Golgi compartmentalization. Future work should first extend findings here and assess the subcellular compartment to which mislocalized LFNG is sequestered. IF microscopy of ER (e.g. anti-calnexin antibody), lysosomal (e.g. LysoLive/LysoSensor, ABCAM 253380), and cytosolic (e.g. anti-vimentin antibody) markers may help define these location(s) (Albrecht et al., 2020; Silveira et al., 2023). I hypothesize that

the ER is the most likely candidate, and thus western blot of ER stress markers such as binding immunoglobulin protein, C/EBP homologous protein, and spliced X-box–binding protein could test this hypothesis further in the same way that it was for GLT8D1 (Tsai et al., 2021; see Section 3.4). Although elevated stress would lend credence to this hypothesis, it is important to recognize the damaging effect transient transfection can have on the ER (Gibson et al., 2013; Tsai et al., 2021). To elucidate how LFNG is trafficked to and within the Golgi, a variety of methods such as brefeldin A treatments with washout, Synchronized Release of Cargo Proteins Using the Retention Using Selective Hooks (RUSH), and subcellular fractionation may be of value (Li et al., 2023). These tests can identify the mechanism of ER-Golgi anterograde transport, the extent of intra-Golgi anterograde transport, and the Golgi sub-compartment in which LFNG resides, respectively (Li et al., 2023). Furthermore, enzyme-linked immunosorbent assay (i.e. ELISA) could help screen for binding partners to evaluate during western blot of LFNG coimmunoprecipitate. Although other methodologies are also viable to address these questions, it is hoped that this brief overview provides direction to those who choose to study LFNG trafficking in future work.

Crucially, clinical action and reproductive risk management require that VUS are reclassified, often through the addition of BS3/PS3 *in vitro* functional evidence (Kanavy et al., 2019; Richards et al., 2015). However, in this work, functional testing of *LFNG* variants consumed significant amounts of time and resources. For example, 18 months were required to prepare reagents (e.g. plasmid cloning and cell culture) and troubleshoot whereas only 10 were required to collect data. Accounting for publication time and project overlap, this averages 13 months of work per variant in Chapter 2 and 1.77 months of work per variant in Chapter 3. Time

is of the essence in variant classification projects, and each moment spent functionally assessing variants is another that the proband and their families do not have access to optimal treatment. To address this problem, novel methods of *in vitro* functional analysis should be investigated. One promising direction is saturation prime editing (Erwood et al., 2022). In this work, over 700 variants were functionally assessed at an approximate rate of 0.03 months per variant. A brief overview of the methodology is as follows: cells of interest are cultured *in vitro* and transfected with numerous pegRNA derivatives of the same template region and edited with CRISPR-Prime. In this way, each cell receives one of all possible point variants within the guide region. After editing, the gene function of each cell is quantified with a verified fluorescent readout where deleterious changes lead to one level of intensity (i.e. high) whereas benign changes are the opposite (i.e. low). Fluorescence-activated cell sorting then organizes each cell into pools, and variants are determined by next-generation sequencing. Interestingly, once gating parameters are established and validated, hypomorphic alleles of various gene activities can also be identified, allowing for a nuanced evaluation of gene activity. Although a reliable fluorescent readout for LFNG has yet to be established, it may be interesting to evaluate whether a signal intensity difference exists between localized and mislocalized variants. Although enzymatically dysfunctional variants could not be ruled out with this method, variants which impede trafficking are necessarily pathogenic and thus identified variants would satisfy PS3 criteria.

Finally, an important future direction will be whole exome sequencing or genome-wide association studies with a specific focus on Clock and Wavefront gene polymorphisms in AIS patients. Examples include those presented here (e.g. *MESP2*, *LFNG*, *Ripply2*, etc.) and those not directly addressed (wavefront genes including *WNT1/2*/etc., *CYP26A2*, *FGF2*, etc.). This

direction may hint at new monogenic and polygenic etiological contributions to scoliosis. The rationale for this project is as follows: First, the WES approach has demonstrated viability as monogenic causes for scoliosis have been identified by screening pathways specific to vertebral formation (Takeda et al., 2018; Terhune et al., 2021). Secondly, more monogenic causes likely exist, but their contribution cannot yet be elucidated as the current genomic and AIS pathogenesis knowledge has been surpassed by contemporary sequencing technology (Man et al., 2019; Terhune et al., 2021, 2022). This is evidenced by the identification of many AIS-associated SNPs in non-coding DNA or other regions of the genome that are not yet understood. Thirdly, as polygenicity is suggested to play a role in AIS, such work could also account for the combinatorial effect of many variants in vertebral structure pathways. Although such a study would require significant investment as well as a large body of genomic and phenotypic data, it appears to have the potential to identify novel AIS contributors.

In sum, this thesis has upgraded the variant classification of two novel *LFNG* variants, provided PS3 evidence for three others, discovered the first two human SCD3-associated *LFNG* alleles with partial enzyme activity, discovered the first *LFNG* allele to partially traffic to the Golgi, and contributed to elucidating the effect of all nine *LFNG* variants studied here on glycosyltransferase activity, processing, and localization. With these discoveries, this work has provided support for a loci-dependant pattern of LFNG functional perturbation, a clearer perspective of LFNG trafficking and its evaluation, and associated an SCD3 proband's larger vertebral number with the first clinically identified hypomorphic *LFNG* allele. Based on this data, I conclude that it is likely that the vertebral number of SCD3 probands is modulated in a variant-specific fashion. Therefore, there does appear to be a genotype-phenotype relationship in

SCD3. It is hoped that this work and others like it can help future researchers identify more scoliosis etiologies, and in so doing create a better life for affected patients and the family and friends who support them.

Bibliography

Adzhubei, I., Jordan, D. M., & Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Current Protocols in Human Genetics*, *SUPPL.76*. https://doi.org/10.1002/0471142905.hg0720s76

Albrecht, L. V., Tejeda-Muñoz, N., & De Robertis, E. M. (2020). Protocol for Probing Regulated Lysosomal Activity and Function in Living Cells. *STAR Protocols*, *1*(3). https://doi.org/10.1016/j.xpro.2020.100132

Asher, M. A., & Burton, D. C. (2006). Adolescent idiopathic scoliosis: Natural history and long term treatment effects. *Scoliosis*, *1*(1), 1–10. https://doi.org/10.1186/1748-7161-1-2

Aulehla, A., Wiegraebe, W., Baubet, V., Wahl, M. B., Deng, C., Taketo, M., Lewandoski, M., & Pourquié, O. (2008). A β-catenin gradient links the clock and wavefront systems in mouse embryo segmentation. *Nature Cell Biology*, *10*(2), 186–193. https://doi.org/10.1038/ncb1679

Babaee, T., Moradi, V., Hashemi, H., Shariat, A., Anastasio, A. T., Khosravi, M., & Bagheripour, B. (2023). Does Bracing Control the Progression of Adolescent Idiopathic Scoliosis in Curves Higher Than 40°? A Systematic Review and Meta-analysis. *Asian Spine Journal*, *17*(1), 203–212. https://doi.org/10.31616/asj.2022.0162

Bagnall, K. M. (2008). Using a synthesis of the research literature related to the aetiology of adolescent idiopathic scoliosis to provide ideas on future directions for success. *Scoliosis*, *3*(1). https://doi.org/10.1186/1748-7161-3-5

Berdon, W. E., Lampl, B. S., Cornier, A. S., Ramirez, N., Turnpenny, P. D., Vitale, M. G., Seimon, L. P., & Cowles, R. A. (2011). Clinical and radiological distinction between spondylothoracic dysostosis (Lavy-Moseley syndrome) and spondylocostal dysostosis (Jarcho-Levin syndrome). *Pediatric Radiology*, *41*(3), 384–388. https://doi.org/10.1007/s00247-010-1928-8

Bessho, Y., Hirata, H., Masamizu, Y., & Kageyama, R. (2003). Periodic repression by the bHLH factor Hes7 is an essential mechanism for the somite segmentation clock. *Genes and Development*, *17*(12), 1451–1456. https://doi.org/10.1101/gad.1092303

Bessho, Y., Miyoshi, G., Sakata, R., & Kageyama, R. (2001). Hes7: a bHLH-type repressor gene regulated by Notch and expressed in the presomitic mesoderm. *Genes to Cells*, *6*(2), 175–185.

Boareto, M., Jolly, M. K., Lu, M., Onuchic, J. N., Clementi, C., & Ben-Jacob, E. (2015). Jagged-delta asymmetry in Notch signaling can give rise to a sender/receiver hybrid phenotype. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(5), E402–E409. https://doi.org/10.1073/pnas.1416287112

Bochter, M. S., Servello, D., Kakuda, S., D'Amico, R., Ebetino, M. F., Haltiwanger, R. S., & Cole, S. E. (2022). Lfng and Dll3 cooperate to modulate protein interactions in cis and coordinate oscillatory Notch pathway activation in the segmentation clock. *Developmental Biology*, *487*, 42–56. https://doi.org/10.1016/j.ydbio.2022.04.004

Bochukova, E. G., Huang, N., Keogh, J., Henning, E., Purmann, C., Blaszczyk, K., Saeed, S., Hamilton-Shield, J., Clayton-Smith, J., O'Rahilly, S., Hurles, M. E., & Farooqi, I. S. (2010).

Large, rare chromosomal deletions associated with severe early-onset obesity. *Nature*, *463*(7281), 666–670. https://doi.org/10.1038/nature08689

Bozzio, A. E., Hu, X., & Lieberman, I. H. (2019). Cost and clinical outcome of adolescent idiopathic scoliosis surgeries—experience from a nonprofit community hospital. *International Journal of Spine Surgery*, *13*(5), 474–478. https://doi.org/10.14444/6063

Brnich, S. E., Abou Tayoun, A. N., Couch, F. J., Cutting, G. R., Greenblatt, M. S., Heinen, C. D., Kanavy, D. M., Luo, X., McNulty, S. M., Starita, L. M., Tavtigian, S. V., Wright, M. W., Harrison, S. M., Biesecker, L. G., Berg, J. S., Abou Tayoun, A. N., Berg, J. S., Biesecker, L. G., Brenner, S. E., … Topper, S. (2019). Recommendations for application of the functional evidence PS3/BS3 criterion using the ACMG/AMP sequence variant interpretation framework. *Genome Medicine*, *12*(1), 1–12. https://doi.org/10.1186/s13073-019-0690-2

Chang, W. P., Lin, Y., Huang, H. L., Lu, H. F., Wang, S. T., Chi, Y. C., Hung, K. S., & Chen, H. Y. (2016). Scoliosis and the subsequent risk of depression. *Spine*, *41*(3), 253–258. https://doi.org/10.1097/BRS.0000000000001187

Chu, L. F., Mamott, D., Ni, Z., Bacher, R., Liu, C., Swanson, S., Kendziorski, C., Stewart, R., & Thomson, J. A. (2019). An In Vitro Human Segmentation Clock Model Derived from Embryonic Stem Cells. *Cell Reports*, *28*(9), 2247-2255.e5. https://doi.org/10.1016/j.celrep.2019.07.090

Clark, E. (2021). Time and space in segmentation. *Interface Focus*, *11*(3), 27–37. https://doi.org/10.1098/rsfs.2020.0049

Cole, S. E., Levorse, J. M., Tilghman, S. M., & Vogt, T. F. (2002). Clock Regulatory Elements Control Cyclic Expression of Lunatic fringe during Somitogenesis. *Developmental Cell*, *3*(1), 75–84. https://doi.org/https://doi.org/10.1016/S1534-5807(02)00212-5

Cooker, J., & Zeeman, E. C. (1976). A Clock and Wavefront Model for Control of the Number of Repeated Structures during Animal Morphogenesis. In *J. theor. Biol* (Vol. 58).

De Salvatore, S., Ruzzini, L., Longo, U. G., Marino, M., Greco, A., Piergentili, I., Costici, P. F., & Denaro, V. (2022). Exploring the association between specific genes and the onset of idiopathic scoliosis: a systematic review. *BMC Medical Genomics*, *15*(1). https://doi.org/10.1186/s12920-022-01272-2

Dequéant, M. L., & Pourquié, O. (2008). Segmental patterning of the vertebrate embryonic axis. *Nature Reviews Genetics*, *9*(5), 370–382. https://doi.org/10.1038/nrg2320

Di Felice, F., Zaina, F., Donzelli, S., & Negrini, S. (2018). The Natural History of Idiopathic Scoliosis during Growth: A Meta-Analysis. *American Journal of Physical Medicine and Rehabilitation*, *97*(5), 346–356. https://doi.org/10.1097/PHM.0000000000000861

Diaz-Cuadros, M., Miettinen, T. P., Skinner, O. S., Sheedy, D., Díaz-García, C. M., Gapon, S., Hubaud, A., Yellen, G., Manalis, S. R., Oldham, W. M., & Pourquié, O. (2023). Metabolic regulation of species-specific developmental rates. *Nature*, *613*(7944), 550–557. https://doi.org/10.1038/s41586-022-05574-4

Diaz-Cuadros, M., Wagner, D. E., Budjan, C., Hubaud, A., Tarazona, O. A., Donelly, S., Michaut, A., Al Tanoury, Z., Yoshioka-Kobayashi, K., Niino, Y., Kageyama, R., Miyawaki,

A., Touboul, J., & Pourquié, O. (2020). In vitro characterization of the human segmentation clock. *Nature*, *580*(7801), 113–118. https://doi.org/10.1038/s41586-019-1885-9

Dubrulle, J., McGrew, M. J., & Pourquié, O. (2001). FGF signaling controls somite boundary position and regulates segmentation clock control of spatiotemporal Hox gene activation. *Cell*, *106*(2), 219–232. https://doi.org/10.1016/S0092-8674(01)00437-8

Dunwoodie, S. L. (2009). Mutation of the fucose-specific β1,3 N-acetylglucosaminyltransferase LFNG results in abnormal formation of the spine. *Biochimica et Biophysica Acta - Molecular Basis of Disease*, *1792*(2), 100–111. https://doi.org/10.1016/j.bbadis.2008.11.003

Erwood, S., Bily, T. M. I., Lequyer, J., Yan, J., Gulati, N., Brewer, R. A., Zhou, L., Pelletier, L., Ivakine, E. A., & Cohn, R. D. (2022). Saturation variant interpretation using CRISPR prime editing. *Nature Biotechnology 2022*, 1–11. https://doi.org/10.1038/s41587-021-01201-1

Erwood, S., I Bily, T. M., Lequyer, J., Yan, J., Gulati, N., Brewer, R. A., Zhou, L., Pelletier, L., Ivakine, E. A., & Cohn, R. D. (2022). Saturation variant interpretation using CRISPR prime editing. *Nature Biotechnology*, *40*(6), 885–895. https://doi.org/10.1038/s41587-021-01201-1

Evrard, Y. A., Lun, Y., Aulehla, A., Gan, L., & Johnson, R. L. (1998). Lunatic fringe is an essential mediator of somite segmentation and patterning. *Nature*, *394*(6691), 377–381. https://doi.org/10.1038/28632

Finnerty, J. R. (2003). The origins of axial patterning in the metazoa: how old is bilateral symmetry? In *Int. J. Dev. Biol* (Vol. 47). www.ijdb.ehu.es

Fleming, A., Kishida, M. G., Kimmel, C. B., & Keynes, R. J. (2015). Building the backbone: The development and evolution of vertebral patterning. In *Development (Cambridge)* (Vol. 142, Issue 10, pp. 1733–1744). Company of Biologists Ltd. https://doi.org/10.1242/dev.118950

Friedmann, H. (1960). The Avian Embryo. Structural and functional development. *Science*, *131*(3408), 1219–1219. https://doi.org/10.1126/science.131.3408.1219.b

Gallant, J. N., Morgan, C. D., Stoklosa, J. B., Gannon, S. R., Shannon, C. N., & Bonfield, C. M. (2018). Psychosocial Difficulties in Adolescent Idiopathic Scoliosis: Body Image, Eating Behaviors, and Mood Disorders. *World Neurosurgery*, *116*, 421-432.e1. https://doi.org/10.1016/j.wneu.2018.05.104

Giampietro, P. F. (2012). Genetic Aspects of Congenital and Idiopathic Scoliosis. *Scientifica*, *2012*(Cvm), 1–15. https://doi.org/10.6064/2012/152365

Gibson, T. J., Seiler, M., & Veitia, R. A. (2013). The transience of transient overexpression. *Nature Methods*, *10*(8), 715–721. https://doi.org/10.1038/nmeth.2534

Gomez, C., Özbudak, E. M., Wunderlich, J., Baumann, D., Lewis, J., & Pourquié, O. (2008). Control of segment number in vertebrate embryos. *Nature*, *454*(7202), 335–339. https://doi.org/10.1038/nature07020

Gomez, C., & Pourquie, O. (2009). Developmental control of segment numbers in vertebrates. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, *312*(6), 533–544. https://doi.org/10.1002/jez.b.21305

Hawes, M. C., & Weinstein, S. L. (2003). Health and Function of Patients with Untreated Idiopathic Scoliosis [2] (multiple letters). *Journal of the American Medical Association*, *289*(20), 2644–2645. https://doi.org/10.1001/jama.289.20.2644-a

Hegde, R. S., & Zavodszky, E. (2019). Recognition and degradation of mislocalized proteins in health and disease. *Cold Spring Harbor Perspectives in Biology*, *11*. https://doi.org/10.1101/cshperspect.a033902

Henrique, D., & Schweisguth, F. (2019). Mechanisms of Notch signaling: a simple logic deployed in time and space. *Development*, *146*, 1–11. https://doi.org/10.1242/dev.172148

Hou, D., Kang, N., Yin, P., & Hai, Y. (2018). Abnormalities associated with congenital scoliosis in high-altitude geographic regions. *International Orthopaedics*, *42*(3), 575–581. https://doi.org/10.1007/s00264-018-3805-2

Hubaud, A., & Pourquié, O. (2014). Signalling dynamics in vertebrate segmentation. *Nature Reviews Molecular Cell Biology*, *15*(11), 709–721. https://doi.org/10.1038/nrm3891

Hussain, A., & Burns, B. (2024). Anatomy, Thorax, Wall. In *StatPearls [Internet]*. StatPearls Publishing.

Ioannidis, N. M., Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., Musolf, A., Li, Q., Holzinger, E., Karyadi, D., Cannon-Albright, L. A., Teerlink, C. C., Stanford, J. L., Isaacs, W. B., Xu, J., Cooney, K. A., Lange, E. M., Schleutker, J., Carpten, J. D., … Sieh, W. (2016). REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *American Journal of Human Genetics*, *99*(4), 877–885. https://doi.org/10.1016/j.ajhg.2016.08.016

Irvine, K. D., Helfand, S. L., & Hognesst, D. S. (1991). The large upstream control region of the Drosophila homeotic gene Ultrabithorax. In *Development* (Vol. 111).

Irvine, K. D., & Okajima, T. (2002). Regulation of Notch Signaling by O-Linked Fucose The biochemical function of FNG raises the question. *Cell*, *111*, 893–904. https://doi.org/https://doi.org/10.1016/S0092-8674(02)01114-5

Irvine, K. D., & Wieschaus, E. (1994). fringe, a boundary-specific signaling molecule, mediates interactions between dorsal and ventral cells during Drosophila wing development. *Cell*, *79*(4), 595–606. https://doi.org/10.1016/0092-8674(94)90545-2

Jackson, M., Marks, L., May, G. H. W., & Wilson, J. B. (2018). The genetic basis of disease. *Essays in Biochemistry*, *62*(5), 643–723. https://doi.org/10.1042/EBC20170053

Johnston, S. H., Rauskolb, C., Wilson, R., Prabhakaran, B., Irvine, K. D., & Vogt, T. F. (1997). A family of mammalian Fringe genes implicated in boundary determination and the Notch pathway. *Development*, *124*(11), 2245–2254. https://doi.org/https://doi.org/10.1242/dev.124.11.2245

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kakuda, S., & Haltiwanger, R. S. (2017). Deciphering the Fringe-mediated Notch Code: Identification of activating and inhibiting sites allowing discrimination between ligands.

Developmental Cell, *40*(2), 193–201. https://doi.org/https://doi.org/10.1016/j.devcel.2016.12.013

Kakuda, S., LoPilato, R. K., Ito, A., & Haltiwanger, R. S. (2020). Canonical notch ligands and fringes have distinct effects on NOTCH1 and NOTCH2. *Journal of Biological Chemistry*, *295*(43), 14710–14722. https://doi.org/10.1074/jbc.RA120.014407

Kanavy, D. M., McNulty, S. M., Jairath, M. K., Brnich, S. E., Bizon, C., Powell, B. C., & Berg, J. S. (2019). Comparative analysis of functional assay evidence use by ClinGen Variant Curation Expert Panels. *Genome Medicine*, *11*(1). https://doi.org/10.1186/s13073-019-0683-1

Kebaish, K. M., Neubauer, P. R., Voros, G. D., Khoshnevisan, M. A., & Skolasky, R. L. (2011). Scoliosis in adults aged forty years and older: Prevalence and relationship to age, race, and gender. *Spine*, *36*(9), 731–736. https://doi.org/10.1097/BRS.0b013e3181e9f120

Kelly, A., Younus, A., & Lekgwara, P. (2020). Adult degenerative scoliosis – A literature review. In *Interdisciplinary Neurosurgery: Advanced Techniques and Case Management* (Vol. 20). Elsevier B.V. https://doi.org/10.1016/j.inat.2019.100661

Keynes, R. (2018). Patterning spinal nerves and vertebral bones. In *Journal of Anatomy* (Vol. 232, Issue 4, pp. 534–539). Blackwell Publishing Ltd. https://doi.org/10.1111/joa.12714

Kikanloo, S. R., Tarpada, S. P., & Cho, W. (2019). Etiology of adolescent idiopathic scoliosis: A literature review. *Asian Spine Journal*, *13*(3), 519–526. https://doi.org/10.31616/asj.2018.0096

Kopan, R., & Ilagan, M. X. G. (2009). The Canonical Notch Signaling Pathway: Unfolding the Activation Mechanism. *Cell*, *137*(2), 216–233. https://doi.org/10.1016/j.cell.2009.03.045

Kwon, J. W., Chae, H. W., Lee, H. S., Kim, S., Sung, S., Lee, S. Bin, Moon, S. H., Lee, H. M., & Lee, B. H. (2021). Incidence rate of congenital scoliosis estimated from a nationwide health insurance database. *Scientific Reports*, *11*(1), 1–11. https://doi.org/10.1038/s41598-021-85088-7

Lecca, M., Bedeschi, M. F., Izzi, C., Dordoni, C., Rinaldi, B., Peluso, F., Caraffi, S. G., Prefumo, F., Signorelli, M., Zanzucchi, M., Bione, S., Ghigna, C., Sassi, S., Novelli, A., Valente, E. M., Superti-Furga, A., Garavelli, L., & Errichiello, E. (2023). Identification of bi-allelic LFNG variants in three patients and further clinical and molecular refinement of spondylocostal dysostosis 3. *Clinical Genetics*, *2023*, 1–8. https://doi.org/10.1111/cge.14336

Lefebvre, M., Dieux-Coeslier, A., Baujat, G., Schaefer, E., Judith, S. O., Bazin, A., Pinson, L., Attie-Bitach, T., Baumann, C., Fradin, M., Pierquin, G., Julia, S., Quélin, C., Doray, B., Berg, S., Vincent-Delorme, C., Lambert, L., Bachmann, N., Lacombe, D., … Thevenon, J. (2018). Diagnostic strategy in segmentation defect of the vertebrae: a retrospective study of 73 patients. *Journal of Medical Genetics*, *55*(6), 422–429. https://doi.org/10.1136/jmedgenet-2017-104939

Li, J., Zhang, J., Bui, S., Ahat, E., Kolli, D., Reid, W., Xing, L., & Wang, Y. (2023). Common Assays in Mammalian Golgi Studies. In V. V. and G. T. R. Wang Yanzhuang and Lupashin (Ed.), *Golgi: Methods and Protocols* (Vol. 2557, pp. 303–332). Springer US. https://doi.org/10.1007/978-1-0716-2639-9_20

Liebsch, C., & Wilkie, H.-J. (2018). Basic Biomechanics of the Thoracic Spine and Rib Cage. In F. Galbusera & H.-J. Wilke (Eds.), *Biomechanics of the Spine* (1st ed., pp. 35–50). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-812851-0.00001-X

Lin, T., Meng, Y., Ji, Z., Jiang, H., Shao, W., Gao, R., & Zhou, X. (2019). Extent of Depression in Juvenile and Adolescent Patients with Idiopathic Scoliosis During Treatment with Braces. *World Neurosurgery*, *126*, e27–e32. https://doi.org/10.1016/j.wneu.2019.01.095

Liu, Y., Yang, Q., & Zhao, F. (2021). Synonymous but Not Silent: The Codon Usage Code for Gene Expression and Protein Folding. *Annual Review of Biochemistry*, *90*, 375–401. https://doi.org/10.1146/annurev-biochem-071320-112701

Luther, K. B., Schindelin, H., & Haltiwanger, R. S. (2009). Structural and mechanistic insights into lunatic fringe from a kinetic analysis of enzyme mutants. *Journal of Biological Chemistry*, *284*(5), 3294–3305. https://doi.org/10.1074/jbc.M805502200

Man, G. C. W., Tang, N. L. S., Chan, T. F., Lam, T. P., Li, J. W., Ng, B. K. W., Zhu, Z., Qiu, Y., & Cheng, J. C. Y. (2019). Replication Study for the Association of GWAS-associated Loci with Adolescent Idiopathic Scoliosis Susceptibility and Curve Progression in a Chinese Population. *Spine*, *44*(7), 464–471. https://doi.org/10.1097/BRS.0000000000002866

Marini, J. C., Forlino, A., Bächinger, H. P., Bishop, N. J., Byers, P. H., De Paepe, A., Fassier, F., Fratzl-Zelman, N., Kozloff, K. M., Krakow, D., Montpetit, K., & Semler, O. (2017). Osteogenesis imperfecta. In *Nature Reviews Disease Primers* (Vol. 3). Nature Publishing Group. https://doi.org/10.1038/nrdp.2017.52

Martinez-Frias, M. L., Bermejo, E., Paisan, L., Martin, M., Egues, J., Lopez -, J. A., Martinez, S., Orbea, C., Cucalon, F., Gairi, J. M., Urioste, M., & De la Cruz, M. A. (1994). Severe spondylocostal dysostosis associated with other congenital anomalies: A clinical/epidemiologic analysis and description of ten cases from the Spanish registry. *American Journal of Medical Genetics*, *51*(3), 203–212. https://doi.org/10.1002/ajmg.1320510306

Matsuda, M., Yamanaka, Y., Uemura, M., Osawa, M., Saito, M. K., Nagahashi, A., Nishio, M., Guo, L., Ikegawa, S., Sakurai, S., Kihara, S., Maurissen, T. L., Nakamura, M., Matsumoto, T., Yoshitomi, H., Ikeya, M., Kawakami, N., Yamamoto, T., Woltjen, K., … Alev, C. (2020). Recapitulating the human segmentation clock with pluripotent stem cells. *Nature*, *580*(7801), 124–129. https://doi.org/10.1038/s41586-020-2144-9

Matsumoto, K., Luther, K. B., & Haltiwanger, R. S. (2021). Diseases related to Notch glycosylation. *Molecular Aspects of Medicine*, *79*. https://doi.org/https://doi.org/10.1016/j.mam.2020.100938

McIntyre, B., Asahara, T., & Alev, C. (2020). Overview of Basic Mechanisms of Notch Signaling in Development and Disease. *Advances in Experimental Medicine and Biology*, *1227*, 9–27. https://doi.org/10.1007/978-3-030-36422-9_2

Menger, R. P., & Sin, A. H. (2023). Adolescent Idiopathic Scoliosis. In *Stat Pearls [Internet]*. StatPearls Publishing LLC. https://www.ncbi.nlm.nih.gov/books/NBK499908/

Miao, Y., Djeffal, Y., De Simone, A., Zhu, K., Lee, J. G., Lu, Z., Silberfeld, A., Rao, J., Tarazona, O. A., Mongera, A., Rigoni, P., Diaz-Cuadros, M., Song, L. M. S., Di Talia, S., & Pourquié,

O. (2023). Reconstruction and deconstruction of human somitogenesis in vitro. *Nature*, *614*, 500–508. https://doi.org/10.1038/s41586-022-05655-4

Miao, Y., & Pourquié, O. (2024). Cellular and molecular control of vertebrate somitogenesis. In *Nature Reviews Molecular Cell Biology*. Nature Research. https://doi.org/10.1038/s41580-024-00709-z

Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., & Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nature Methods*, *19*(6), 679–682. https://doi.org/10.1038/s41592-022-01488-1

Moloney, D. J., Panin, V. M., Johnston, S. H., Chen, J., Shao, L., Wilson, R., Wang#, Y., Stanley, P., Irvine, K. D., Haltiwanger, R. S., & Vogt, T. F. (2000). Fringe is a glycosyltransferase that modifies Notch. *Nature*, *406*, 369–375. https://doi.org/https://doi.org/10.1038/35019000

Moore, K. L., Dalley, A. F., & Agur, A. M. R. (2014). Thorax. In C. Taylor (Ed.), *Clinically Oriented Anatomy* (7th ed., pp. 72–128). Lippincott Williams & Wilkins.

Muller, H. J. (1932). Further studies on the nature and causes of gene mutations. *Proceedings of the 6th International Congress of Genetics*, 213–255. https://api.semanticscholar.org/CorpusID:204159630

Musumeci, G., Castrogiovanni, P., Coleman, R., Szychlinska, M. A., Salvatorelli, L., Parenti, R., Magro, G., & Imbesi, R. (2015). Somitogenesis: From somite to skeletal muscle. *Acta Histochemica*, *117*, 313–328. https://doi.org/10.1016/j.acthis.2015.02.011

Naganathan, S. R., & Oates, A. C. (2020). Patterning and mechanics of somite boundaries in zebrafish embryos. *Seminars in Cell and Developmental Biology*, *107*, 170–178. https://doi.org/10.1016/j.semcdb.2020.04.014

Nandagopal, N., Santat, L. A., LeBon, L., Sprinzak, D., Bronner, M. E., & Elowitz, M. B. (2018). Dynamic Ligand Discrimination in the Notch Signaling Pathway. *Cell*, *172*(4), 869-880.e19. https://doi.org/10.1016/j.cell.2018.01.002

Ng, P. C., & Henikoff, S. (2001). Predicting deleterious amino acid substitutions. *Genome Research*, *11*(5), 863–874. https://doi.org/10.1101/gr.176601

Nitanda, Y., Matsui, T., Matta, T., Higami, A., Kohno, K., Nakahata, Y., & Bessho, Y. (2014). 3′-UTR-dependent regulation of mRNA turnover is critical for differential distribution patterns of cyclic gene mRNAs. *FEBS Journal*, *281*(1), 146–156. https://doi.org/10.1111/febs.12582

Nóbrega, A., Maia-Fernandes, A. C., & Andrade, R. P. (2021). Altered Cogs of the Clock: Insights into the Embryonic Etiology of Spondylocostal Dysostosis. *Journal of Developmental Biology*, *9*(5), 1–14. https://doi.org/10.3390/jdb9010005

Oginuma, M., Niwa, Y., Chapman, D. L., & Saga, Y. (2008). Mesp2 and Tbx6 cooperatively create periodic patterns coupled with the clock machinery during mouse somitogenesis. *Development*, *135*(15), 2555–2562. https://doi.org/10.1242/dev.019877

Oginuma, M., Takahashi, Y., Kitajima, S., Kiso, M., Kanno, J., Kimura, A., & Saga, Y. (2010). The oscillation of Notch activation, but not its boundary, is required for somite border formation and rostral-caudal patterning within a somite. *Development*, *137*(9), 1515–1522. https://doi.org/10.1242/dev.044545

Okajima, T., Xu, A., Lei, L., & Irvine, K. D. (2005). Roles of O-Fucosyltransferase 1 and O-Linked Fucose in Notch Receptor Function. *Science*, *307*(5715), 1599–1603.

Otomo, N., Mizumoto, S., Lu, H. F., Takeda, K., Campos-Xavier, B., Mittaz-Crettol, L., Guo, L., Takikawa, K., Nakamura, M., Yamada, S., Matsumoto, M., Watanabe, K., & Ikegawa, S. (2019). Identification of novel LFNG mutations in spondylocostal dysostosis. *Journal of Human Genetics*, *64*(3), 261–264. https://doi.org/10.1038/s10038-018-0548-2

Palmeirim, I., Henrique, D., Ish-Horowicz, D., & Pourquie, O. (1997). Avian hairy Gene Expression Identifies a Molecular Clock Linked to Vertebrate Segmentation and Somitogenesis. *Cell*, *91*, 639–648. https://doi.org/https://doi.org/10.1016/S0092-8674(00)80451-1

Palmeirim, I., Pais-de-Azevedo, T., Magno, R., & Duarte, I. (2018). Recent advances in understanding vertebrate segmentation. *F1000Research*, *7*. https://doi.org/10.12688/f1000research.12369.1

Pourquié, O. (2022). A brief history of the segmentation clock. *Developmental Biology*, *485*, 24–36. https://doi.org/10.1016/j.ydbio.2022.02.011

Rampal, R., Li, A. S. Y., Moloney, D. J., Georgiou, S. A., Luther, K. B., Nita-Lazar, A., & Haltiwanger, R. S. (2005). Lunatic fringe, manic fringe, and radical fringe recognize similar specificity determinants in O-fucosylated epidermal growth factor-like repeats. *Journal of Biological Chemistry*, *280*(51), 42454–42463. https://doi.org/10.1074/jbc.M509552200

Rentzsch, P., Witten, D., Cooper, G. M., Shendure, J., & Kircher, M. (2019). CADD: Predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Research*, *47*(D1), D886–D894. https://doi.org/10.1093/nar/gky1016

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., & Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: A joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, *17*(5), 405–424. https://doi.org/10.1038/gim.2015.30

Riley, M. F., Bochter, M. S., Wahi, K., Nuovo, G. J., & Cole, S. E. (2013). Mir-125a-5p-Mediated Regulation of Lfng Is Essential for the Avian Segmentation Clock. *Developmental Cell*, *24*(5), 554–561. https://doi.org/10.1016/j.devcel.2013.01.024

Rimon, D. L., Fletcher, B. D., & McKusick, V. A. (1968). Spondylocostal Dysplasia* A dominantly inherited form of short-trunked dwarfism. *American Journal of Medicine*, *45*, 948–953.

Rini, J. M., Moremen, K. W., Davis, B. G., & Esko, J. D. (2022). Glycosyltransferases and Glycan-Processing Enzymes. In A. Varki (Ed.), *Essentials of Glycobiology [Internet]* (4th ed.). Cold Spring Harbor. https://doi.org/https://doi:10.1101/glycobiology.4e.6.

Rowe, D. E., Bernstein, S. M., Riddick, M. F., Adler, F., Emans, J. B., & Gardner-Bonneau, D. (1997). A Meta-Analysis of the Efficacy of Non-Operative Treatments for Idiopathic Scoliosis. *The Journal of Bone and Joint Surgery*, *79*(5).

https://journals.lww.com/jbjsjournal/fulltext/1997/05000/a_meta_analysis_of_the_efficacy_of_non_operative.5.aspx

Sadler, B., Haller, G., Antunes, L., Bledsoe, X., Morcuende, J., Giampietro, P., Raggio, C., Miller, N., Kidane, Y., Wise, C. A., Amarillo, I., Walton, N., Seeley, M., Johnson, D., Jenkins, C., Jenkins, T., Oetjens, M., Tong, R. S., Druley, T. E., … Gurnett, C. A. (2019). Distal chromosome 16p11.2 duplications containing SH2B1 in patients with scoliosis. *Journal of Medical Genetics*, *56*(7), 427–433. https://doi.org/10.1136/jmedgenet-2018-105877

Saga, Y. (2012). The mechanism of somite formation in mice. *Current Opinion in Genetics & Development*, *22*(4), 331–338. https://doi.org/https://doi.org/10.1016/j.gde.2012.05.004

Sanaki-Matsumiya, M., & Ebisuya, M. (2022). *Human embryonic organoids to recapitulate periodic somitogenesis in vitro*. https://doi.org/10.21203/rs.3.pex-1852/v1

Sanaki-Matsumiya, M., Matsuda, M., Gritti, N., Nakaki, F., Sharpe, J., Trivedi, V., & Ebisuya, M. (2022). Periodic formation of epithelial somites from human pluripotent stem cells. *Nature Communications*, *13*(2325), 1–14. https://doi.org/10.1038/s41467-022-29967-1

Sawada, A., Shinya, M., Jiang, J.-Y., Kawakami, A., Kuroiwa, A., & Takeda, H. (2001). Fgf/MAPK Signalling is a Crucial Positional Cue in Somite Boundary Formation. *Development*, *128*, 4873–4880.

Schröter, C., Herrgen, L., Cardona, A., Brouhard, G. J., Feldman, B., & Oates, A. C. (2008). Dynamics of zebrafish somitogenesis. *Developmental Dynamics*, *237*(3), 545–553. https://doi.org/10.1002/dvdy.21458

Schuhmann, S., Koller, H., Sticht, H., Kraus, C., Krumbiegel, M., Uebe, S., Ekici, A. B., Reis, A., & Thiel, C. T. (2021). Clinical and molecular delineation of spondylocostal dysostosis type 3. *Clinical Genetics*, *99*(6), 851–852. https://doi.org/10.1111/cge.13952

Sebaaly, A., Daher, M., Salameh, B., Ghoul, A., George, S., & Roukoz, S. (2022). Congenital scoliosis: A narrative review and proposal of a treatment algorithm. *EFORT Open Reviews*, *7*(5), 318–327. https://doi.org/10.1530/EOR-21-0121

Shi, S., & Stanley, P. (2002). Protein O-fucosyltransferase 1 is an essential component of Notch signaling pathways. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(9), 5234–5239. https://doi.org/https://doi/10.1073/pnas.0831126100

Shifley, E. T., & Cole, S. E. (2008). Lunatic fringe protein processing by proprotein convertases may contribute to the short protein half-life in the segmentation clock. *Biochimica et Biophysica Acta - Molecular Cell Research*, *1783*(12), 2384–2390. https://doi.org/10.1016/j.bbamcr.2008.07.009

Shifley, E. T., VanHorn, K. M., Perez-Balaguer, A., Franklin, J. D., Weinstein, M., & Cole, S. E. (2008). Oscillatory lunatic fringe activity is crucial for segmentation of the anterior but not posterior skeleton. *Development*, *135*(5), 899–908. https://doi.org/10.1242/dev.006742

Silveira, K. C., Fonseca, I. C., Oborn, C., Wengryn, P., Ghafoor, S., Beke, A., Dreseris, E. S., Wong, C., Iacovone, A., Soltys, C. L., Babul-Hirji, R., Artigalas, O., Antolini-Tavares, A., Gingras, A. C., Campos, E., Cavalcanti, D. P., & Kannu, P. (2023). CYP26B1-related disorder: expanding the ends of the spectrum through clinical and molecular evidence. *Human Genetics*, *142*(11), 1571–1586. https://doi.org/10.1007/s00439-023-02598-2

Sinnott, M. L. (1990). Catalytic mechanism of enzymic glycosyl transfer. *Chemical Reviews*, *90*, 1171–1202. https://doi.org/https://doi:10.1021/CR00105A006

Sparrow, D. B., Chapman, G., Smith, A. J., Mattar, M. Z., Major, J. A., O'Reilly, V. C., Saga, Y., Zackai, E. H., Dormans, J. P., Alman, B. A., McGregor, L., Kageyama, R., Kusumi, K., & Dunwoodie, S. L. (2012). A mechanism for gene-environment interaction in the etiology of congenital scoliosis. *Cell*, *149*(2), 295–306. https://doi.org/10.1016/j.cell.2012.02.054

Sparrow, D. B., Chapman, G., Wouters, M. A., Whittock, N. V., Ellard, S., Fatkin, D., Turnpenny, P. D., Kusumi, K., Sillence, D., & Dunwoodie, S. L. (2006). Mutation of the Lunatic Fringe gene in humans causes spondylocostal dysostosis with a severe vertebral phenotype. *American Journal of Human Genetics*, *78*(1), 28–37. https://doi.org/10.1086/498879

Stauber, M., Sachidanandan, C., Morgenstern, C., & Ish-Horowicz, D. (2009). Differential axial requirements for Lunatic fringe and Hes7 transcription during mouse somitogenesis. *PLoS ONE*, *4*(11). https://doi.org/10.1371/journal.pone.0007996

Sung, S., Chae, H. W., Lee, H. S., Kim, S., Kwon, J. W., Lee, S. Bin, Moon, S. H., Lee, H. M., & Lee, B. H. (2021). Incidence and surgery rate of idiopathic scoliosis: A nationwide database study. *International Journal of Environmental Research and Public Health*, *18*(15). https://doi.org/10.3390/ijerph18158152

Takeda, K., Kou, I., Mizumoto, S., Yamada, S., Kawakami, N., Nakajima, M., Otomo, N., Ogura, Y., Miyake, N., Matsumoto, N., Kotani, T., Sudo, H., Yonezawa, I., Uno, K., Taneichi, H., Watanabe, K., Shigematsu, H., Sugawara, R., Taniguchi, Y., … Ikegawa, S. (2018).

Screening of known disease genes in congenital scoliosis. *Molecular Genetics and Genomic Medicine*, *6*(6), 966–974. https://doi.org/10.1002/mgg3.466

Tam, P. P. L. (1981). The control of somitogenesis in mouse embryos. *The Journal Embryology and Experimental Morphology*, *65*, 103–131. https://doi.org/https://doi-org.login.ezproxy.library.ualberta.ca/10.1242/dev.65.Supplement.103

Tan, K.-J., Moe, M. M., Vaithinathan, R., & Wong, H.-K. (2009). Curve Progression in Idiopathic Scoliosis Follow-up Study to Skeletal Maturity. *Spine*, *34*(7), 697–700.

Tavtigian, S. v., Deffenbaugh, A. M., Yin, L., Judkins, T., Scholl, T., Samollow, P. B., de Silva, D., Zharkikh, A., & Thomas, A. (2006). Comprehensive statistical study of 452 BRCA1 missense substitutions with classification of eight recurrent substitutions as neutral. *Journal of Medical Genetics*, *43*(4), 295–305. https://doi.org/10.1136/jmg.2005.033878

Terhune, E. A., Heyn, P. C., Piper, C. R., & Hadley-Miller, N. (2022). Genetic variants associated with the occurrence and progression of adolescent idiopathic scoliosis: a systematic review protocol. *Systematic Reviews*, *11*(1). https://doi.org/10.1186/s13643-022-01991-8

Terhune, E. A., Wethey, C. I., Cuevas, M. T., Monley, A. M., Baschal, E. E., Bland, M. R., Baschal, R., Devon Trahan, G., Taylor, M. R. G., Jones, K. L., & Miller, N. H. (2021). Whole exome sequencing of 23 multigeneration idiopathic scoliosis families reveals enrichments in cytoskeletal variants, suggests highly polygenic disease. *Genes*, *12*(6). https://doi.org/10.3390/genes12060922

Tsai, P. C., Jih, K. Y., Shen, T. Y., Liu, Y. H., Lin, K. P., Liao, Y. C., & Lee, Y. C. (2021). Genetic and Functional Analysis of Glycosyltransferase 8 Domain–Containing Protein 1 in

Taiwanese Patients with Amyotrophic Lateral Sclerosis. *Neurology: Genetics*, *7*(6), e627. https://doi.org/10.1212/NXG.0000000000000627

Turnpenny, P. D., Sloman, M., & Dunwoodie, S. (2009). Spondylocostal Dysostosis, Autosomal Recessive Synonyms: Costovertebral Dysplasia, Spondylocostal Dysplasia. In M. P. Adam, J. Feldman, & G. M. Mirzaa (Eds.), *GeneReviews* (pp. 1993–2024). US National Library of Medicine. https://www.ncbi.nlm.nih.gov/books/

Turnpenny, P. D., Whittock, N., Duncan, J., Dunwoodie, S., Kusumi, K., & Ellard, S. (2003). Novel mutations in DLL3, a somitogenesis gene encoding a ligand for the Notch signalling pathway, cause a consistent pattern of abnormal vertebral segmentation in spondylocostal dysostosis. *Journal of Medical Genetics*, *40*(5), 333–339. https://doi.org/10.1136/jmg.40.5.333

UniProtKB. (2021, April 7). *LFNG - Beta-1,3-N-acetylglucosaminyltransferase lunatic fringe - Homo sapiens (Human) - LFNG gene & protein*. UniProt Consortium. https://www.uniprot.org/uniprot/Q8NES3

Ünligil, U. M., & Rini, J. M. (2000). Glycosyltransferase structure and mechanism. *Current Opinion in Structural Biology*, *10*(5), 510–517. https://doi.org/https://doi.org/10.1016/S0959-440X(00)00124-X

Venkatachalam, K., Hofmann, T., & Montell, C. (2006). Lysosomal localization of TRPML3 depends on TRPML2 and the mucolipidosis-associated protein TRPML1. *Journal of Biological Chemistry*, *281*(25), 17517–17527. https://doi.org/10.1074/jbc.M600807200

Venzin, O. F., & Oates, A. C. (2020). What are you synching about? Emerging complexity of Notch signaling in the segmentation clock. *Developmental Biology*, *460*(1), 40–54. https://doi.org/10.1016/J.YDBIO.2019.06.024

Vermot, J., Llamas, J. G., Fraulob, V., Niederreither, K., Chambon, P., & Dollé, P. (2005). Retinoic Acid Controls the Bilateral Symmetry of Somite Formation in the Mouse Embryo. *Science*, *308*(5721), 563–566. https://doi.org/10.1126/science.1108363

Wahi, K., Friesen, S., Coppola, V., & Cole, S. E. (2017). Putative binding sites for mir-125 family miRNAs in the mouse Lfng 3′UTR affect transcript expression in the segmentation clock, but mir-125a-5p is dispensable for normal somitogenesis. *Developmental Dynamics*, *246*(10), 740–748. https://doi.org/10.1002/dvdy.24552

Wahl, M. B., Deng, C., Lewandowski, M., & Pourquié, O. (2007). FGF signaling acts upstream of the NOTCH and WNT signaling pathways to control segmentation clock oscillations in mouse somitogenesis. *Development*, *134*(22), 4033–4041. https://doi.org/10.1242/dev.009167

Waxenbaum, J. A., Reddy, V., & Futterman, B. (2024). Anatomy, Back, Thoracic Vertebrae. In *StatPearls*. StatPearls Publishing.

Weinstein, S. L. (1986). Idiopathic Scoliosis: Natural History. *Spine*, *11*(8). https://journals.lww.com/spinejournal/fulltext/1986/10000/idiopathic_scoliosis__natural_history.6.aspx

Wengryn, P., Silveira, K. da C., Oborn, C., Soltys, C.-L., Beke, A., Chacon-Fonseca, I., Damseh, N., Rodriguez, M. Q., Badilla-Porras, R., & Kannu, P. (2023). Functional Characterization

of Novel Lunatic Fringe Variants in Spondylocostal Dysostosis Type-III with Scoliosis. *Human Mutation*, *2023*, 1–12. https://doi.org/10.1155/2023/5989733

White, K. K., White, K. K., Bober, M. B., Cho, T. J., Goldberg, M. J., Goldberg, M. J., Hoover-Fong, J., Irving, M., Kamps, S. E., Kamps, S. E., Mackenzie, W. G., Raggio, C., Spencer, S. A., Bompadre, V., & Savarirayan, R. (2020). Best practice guidelines for management of spinal disorders in skeletal dysplasia. *Orphanet Journal of Rare Diseases*, *15*(1), 1–11. https://doi.org/10.1186/s13023-020-01415-7

Williams, D. R., Shifley, E. T., Braunreiter, K. M., & Cole, S. E. (2016). Disruption of somitogenesis by a novel dominant allele of Lfng suggests important roles for protein processing and secretion. *Development*, *143*, 822–830. https://doi.org/10.1242/dev.128538

Williams, D. R., Shifley, E. T., Lather, J. D., & Cole, S. E. (2014). Posterior skeletal development and the segmentation clock period are sensitive to Lfng dosage during somitogenesis. *Developmental Biology*, *388*(2), 159–169. https://doi.org/10.1016/j.ydbio.2014.02.006

Williams, S., Alkhatib, B., & Serra, R. (2019). Development of the axial skeleton and intervertebral disc. In *Current topics in developmental biology* (pp. 50–82). Elsevier. https://doi.org/10.1016/bs.ctdb.2018.11.018

Wise, C. A. (2015). The Genetic Architecture of Idiopathic Scoliosis. In J. J. Wise Carol A. and Rios (Ed.), *Molecular Genetics of Pediatric Orthopaedic Disorders* (pp. 71–89). Springer New York. https://doi.org/10.1007/978-1-4939-2169-0_5

Wishart, B. D., & Kivlehan, E. (2021). Neuromuscular Scoliosis: When, Who, Why and Outcomes. In *Physical Medicine and Rehabilitation Clinics of North America* (Vol. 32, Issue 3, pp. 547–556). W.B. Saunders. https://doi.org/10.1016/j.pmr.2021.02.007

Wu, M. M., Llopis, J., Adams, S., Mccaffery, J. M., Kulomaa, M. S., Machen, T. E., Moore, H.-P. H., & Tsien, R. Y. (2000). Organelle pH studies using targeted avidin and fluorescein-biotin. *Chemistry & Biology, 7(3)*, *7*(3), 197–209.

Yasuhiko, Y., Haraguchi, S., Kitajima, S., Takahashi, Y., Kanno, J., & Saga, Y. (2006). Tbx6-mediated Notch signaling controls somite-specific Mesp2 expression. *Proceedings of the Proceedings of the National Academy of Sciences of the United States of America*, *103*(10), 3651–3656. https://doi.org/10.1073/pnas.0508238103

Yoshioka-Kobayashi, K., Matsumiya, M., Niino, Y., Isomura, A., Kori, H., Miyawaki, A., & Kageyama, R. (2020). Coupling delay controls synchronized oscillation in the segmentation clock. *Nature*, *580*(7801), 119–123. https://doi.org/10.1038/s41586-019-1882-z

Zeiser, S., Liebscher, H. V., Tiedemann, H., Rubio-Aliaga, I., Przemeck, G. K., Hrabé De Angelis, M., & Winkler, G. (2006). Number of active transcription factor binding sites is essential for the Hes7 oscillator. *Theoretical Biology and Medical Modelling*, *3*(11). https://doi.org/10.1186/1742-4682-3-11

Zhang, N., Norton, C. R., & Gridley, T. (2002). Segmentation defects of Notch pathway mutants and absence of a synergistic phenotype in lunatic fringe/radical fringe double mutant mice. *Genesis*, *33*(1), 21–28. https://doi.org/10.1002/gene.10081

Zhang, Z., Miteva, M. A., Wang, L., & Alexov, E. (2012). Analyzing effects of naturally occurring missense mutations. *Computational and Mathematical Methods in Medicine*, *2012*. https://doi.org/10.1155/2012/805827

Zhao, W., Ajima, R., Ninomiya, Y., & Saga, Y. (2015). Segmental border is defined by Ripply2-mediated Tbx6 repression independent of Mesp2. *Developmental Biology*, *400*(1), 105–117. https://doi.org/10.1016/j.ydbio.2015.01.020

Zhou, J., Wang, Y., Xie, J., Zhao, Z., Shi, Z., Li, T., Zhang, Y., Zhang, L., Zhu, T., Zhao, W., Yang, X., Bi, N., & Li, Q. (2023). Scoliosis school screening of 139,922 multi-ethnic children in Dali, southwestern China: A large epidemiological study. *IScience*, *26*(12). https://doi.org/10.1016/j.isci.2023.108305

Zhou, M., Yan, J., Ma, Z., Zhou, Y., Abbood, N. N., Liu, J., Su, L., Jia, H., & Guo, A. Y. (2012). Comparative and evolutionary analysis of the HES/HEY gene family reveal exon/intron loss and teleost specific duplication events. *PLoS ONE*, *7*(7). https://doi.org/10.1371/journal.pone.0040649

Appendices

Supplementary Tables

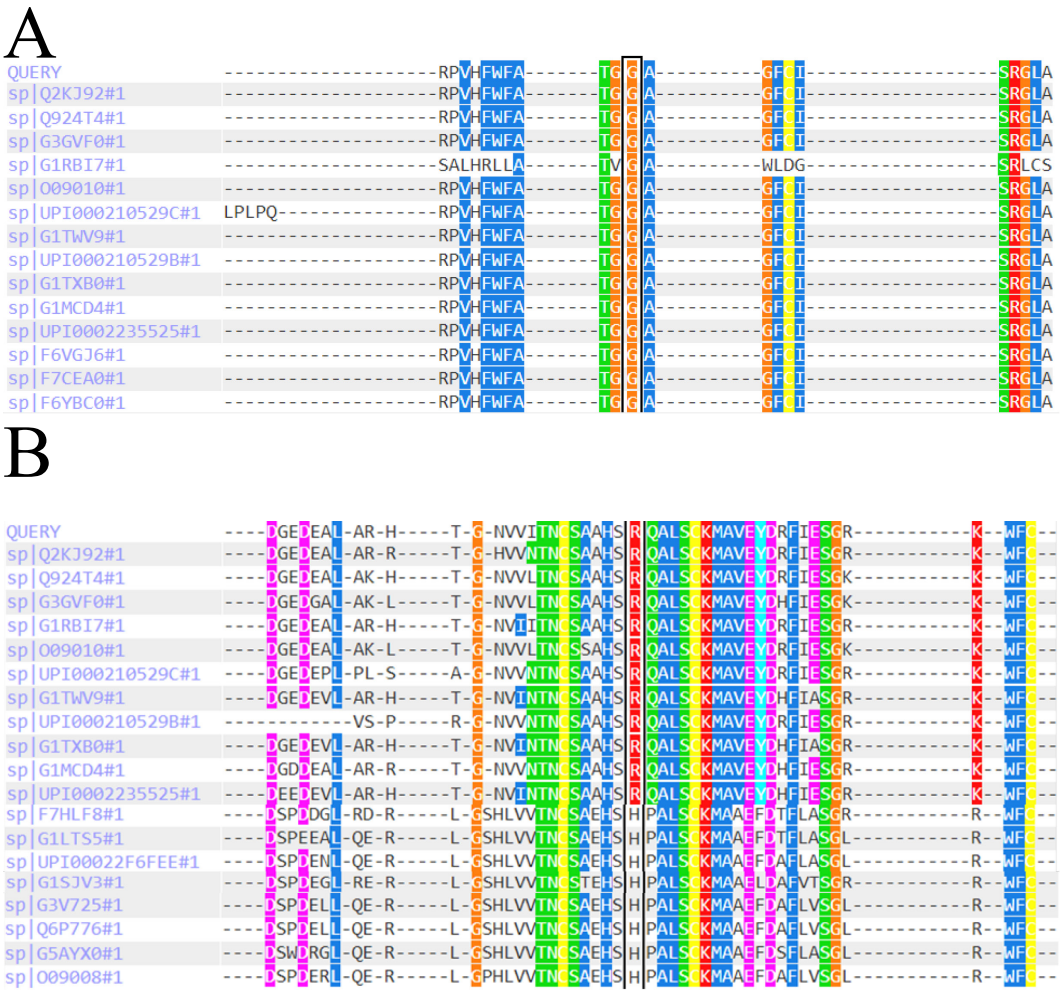
LFNG Variant	Forward (5' to 3')	Reverse (5' to 3')
c.446C>T	TCTTCCCCGTCAATGAAGATGAACGTCATCTCCTTG	CAAGGAGATGACGTTCATCTTCATTGACGGGGAAGA
c.467T>G	CGTGTGCCTGGCCCGGGCCTCATCTTC	GAAGATGAGGCCCGGGCCAGGCACACG
c.583T>C	ACGTGGCAGAACCGCTTCCTGCCGGAC	GTCCGGCAGGAAGCGGTTCTGCCACGT
c.601G>A	GTTGACGTAGTTGTCATTGTCCACGTGGCAGAACC	GGTTCTGCCACGTGGACAATGACAACTACGTCAAC
c.842C>A	CCGCTCAGCCTTATTCATGAAGTGACCCCC	GGGGGTCACTTCATGAATAAGGCTGAGCGG
c.856C>T	TCATCAGGCAGCCAGATCCGCTCAGCC	GGCTGAGCGGATCTGGCTGCCTGATGA

Supplementary Table 1: Primers for Site-Directed Mutagenesis.

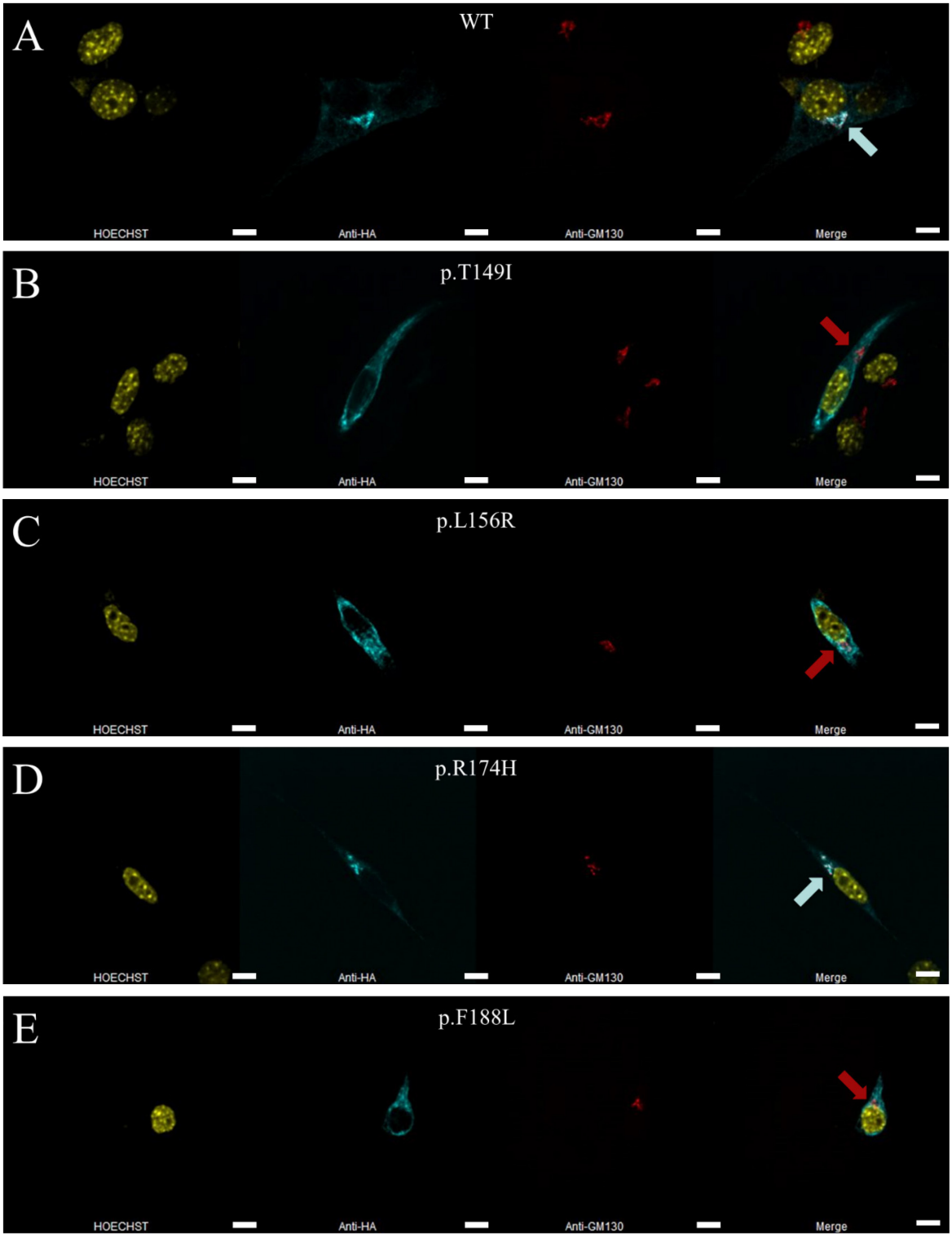
Proband	Age	Sex	Consanguinity	Height/ SD	AS:H Ratio	Mid-Parental Height	Vertebral Body Number	Rib Number	Scoliosis/ Cobb Angle	Axial Abnormally Description	Additional Findings Described
Sparrow et al. 2006	N/A	M	N/A	155cm/ N/A	1.20	N/A	N/A†	N/A†	Scoliosis/ N/A	Discussed	Discussed, present
Lefebvre et al. 2018 (1)	N/A	N/A	No	N/A/ N/A	N/A	N/A	N/A	N/A	N/A†/ N/A	Minimally Discussed	Not Discussed
Lefebvre et al. 2018 (2)	N/A	N/A	No	N/A/ N/A	N/A	N/A	N/A	N/A	N/A†/ N/A	Minimally Discussed	Not Discussed
Takeda et al. 2018	16 y	M	No	N/A/ -2.1 SD	N/A	N/A	N/A†	N/A†	Scoliosis/ 70° (T9-L3)	Discussed	Discussed, none
Otomo et al. 2019	9 m	M	No	N/A/ -2.5 SD	N/A	N/A	N/A†	N/A†	No scoliosis	Discussed	Discussed, none
Schuhmann et al. 2021	17 y	N/A	No	N/A/ N/A	N/A	N/A	N/A†	N/A†	Scoliosis, N/A	Minimally Discussed	Discussed, present
Wengryn et al. 2023	2 y 8 m	M	No	N/A/ -0.67 to -1.28 SD	1.08	179 cm	>20	10 left, 9 right	Scoliosis/ 38° (L1-S1)	Discussed	Discussed, none

Supplementary Table 2: Clinical Data Recorded for Each SCD3 Proband. N/A: Not Available. N/A†: Radiographs present but did not assess. AS:H Ratio: Arm-Span to Height Ratio. M: Male. m: Months. SD: Standard Deviation. y: years.

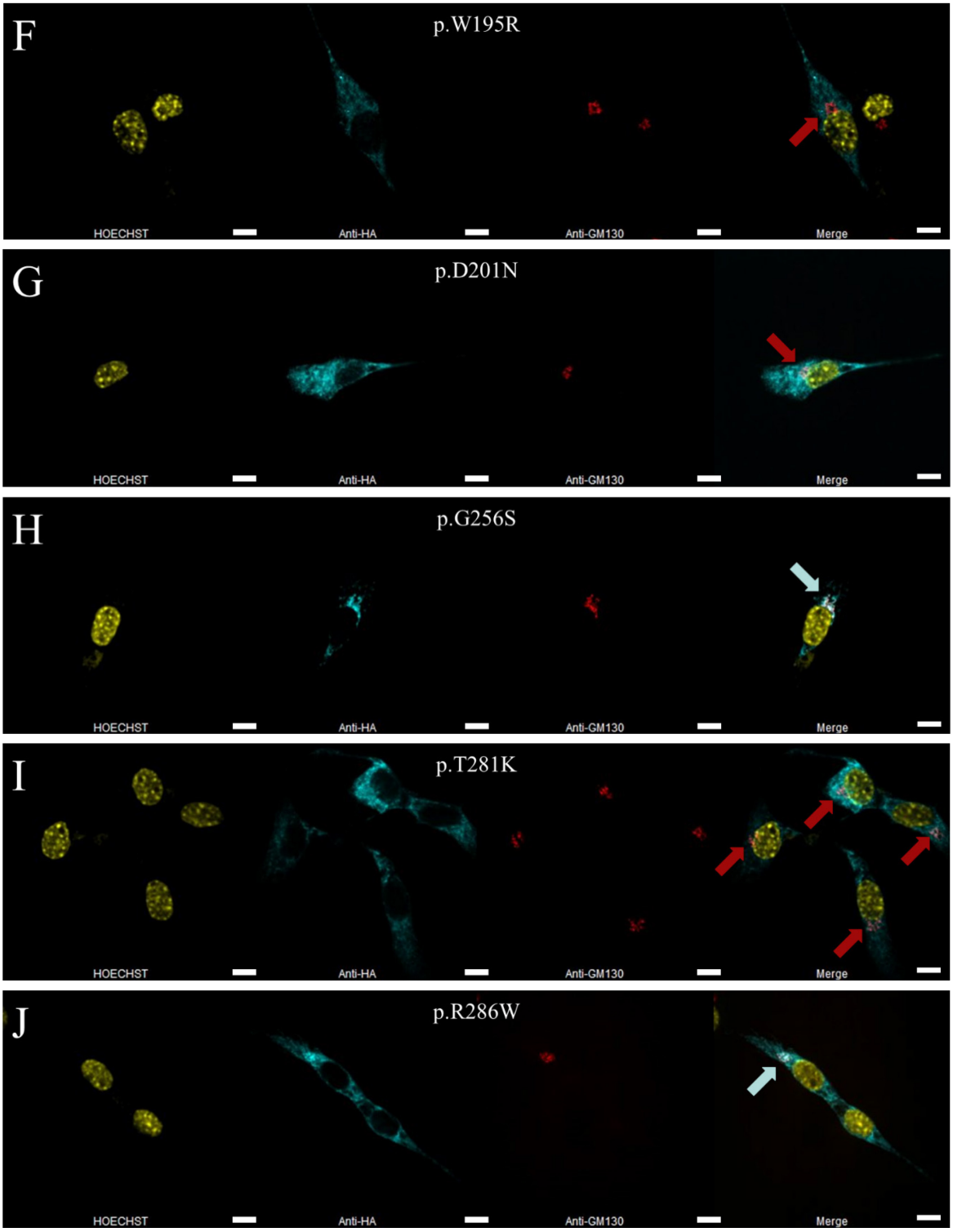
Supplementary Figures



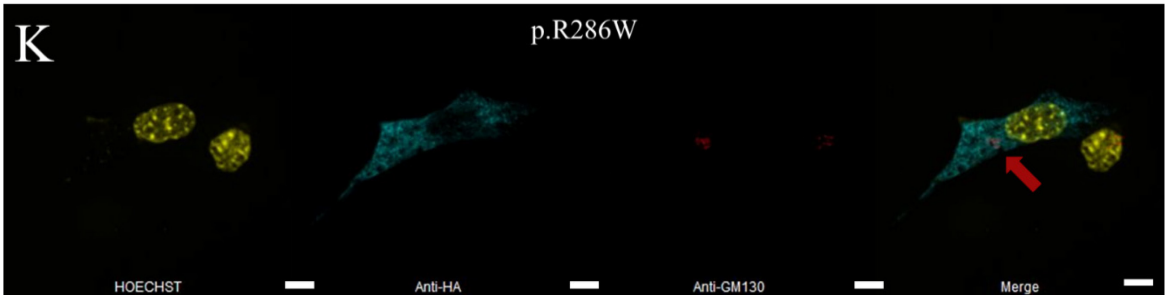
Supplementary Figure 1: ClustalX multiple sequence alignment of all clustered UniRef50 accessions related to LFNG. Displayed are 75 amino acids flanking the positions of interest labeled with a black column. The query sequence is the WT LFNG in both cases. **A)** The MSA conserved residues surrounding position 256. Note the conservation of glycine throughout alignments. **B)** The MSA conserved residues surrounding position 174. The bottom eight residues transition from LFNG sequences to MFNG sequences.



Supplementary Figure 2: Composites of Immunofluorescent Imaging of LFNG Variants.



Supplementary Figure 2: Composites of Immunofluorescent Imaging of LFNG Variants.

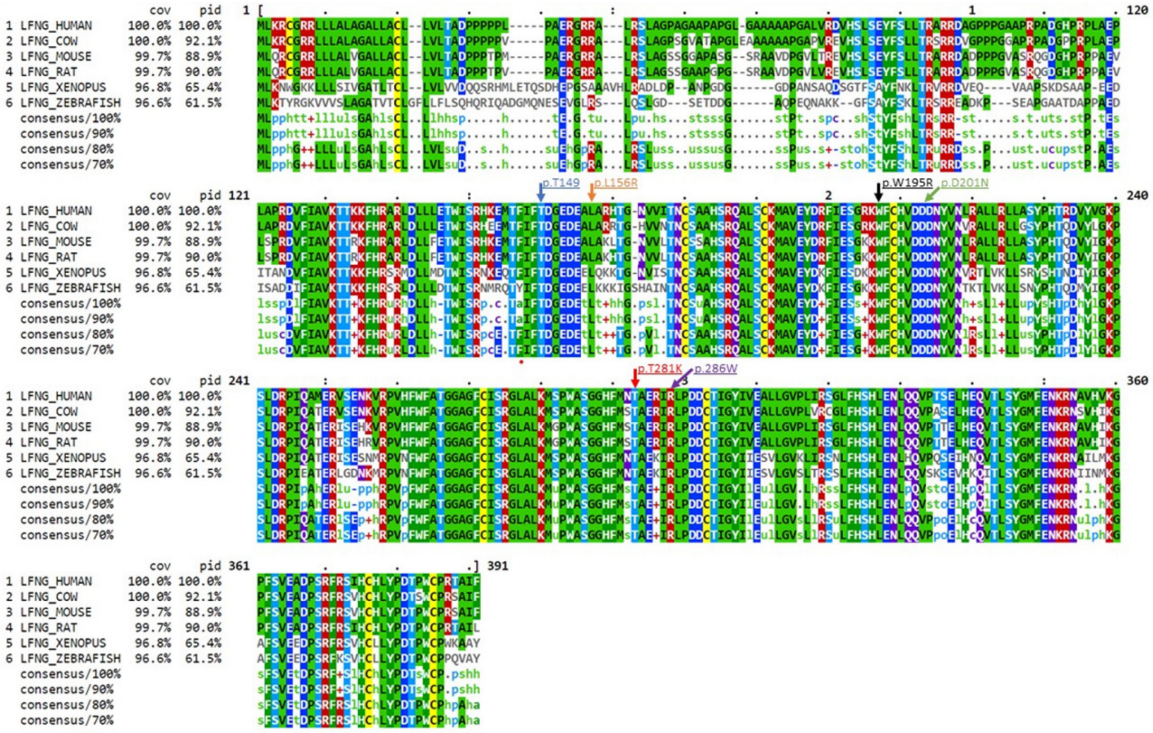


Supplementary Figure 2: Composites of Immunofluorescent Imaging of LFNG Variants. Immunofluorescent microscopy images of NIH-3T3 cells transiently transfected with *LFNG* variant plasmids. Cells were treated with rabbit anti-HA and mouse anti-GM130 primary antibodies, then fluorescent secondary antibodies followed by HOECHST. Turquoise arrows indicate Anti-HA/Anti-GM130 signal colocalization. Red arrows indicate Anti-GM130 signal and thus lack of Anti-HA/Anti-GM130 colocalization. Images are representative samples of 40 cells screened at random across the three independent experiments (N = 3), and values (/40) indicate number of cells which showed the pattern exemplified in the panel. **A)** WT- (40/40), **B)** p.T149I (38/40), **C)** p.L156R (36/40), **D)** p.R174H (40/40), **E)** p.F188L (39/40), **F)** p.W195R (38/40), **G)** p.D201N (37/40), **H)** p.G256S (40/40), **I)** p.T281K (38/40), and p.R286W **(J, K)**. Panel J) shows an example of localization (25/40) whereas panel K) shows mislocalization for the same variant (15/40). Scale bars = 15 μm.

A

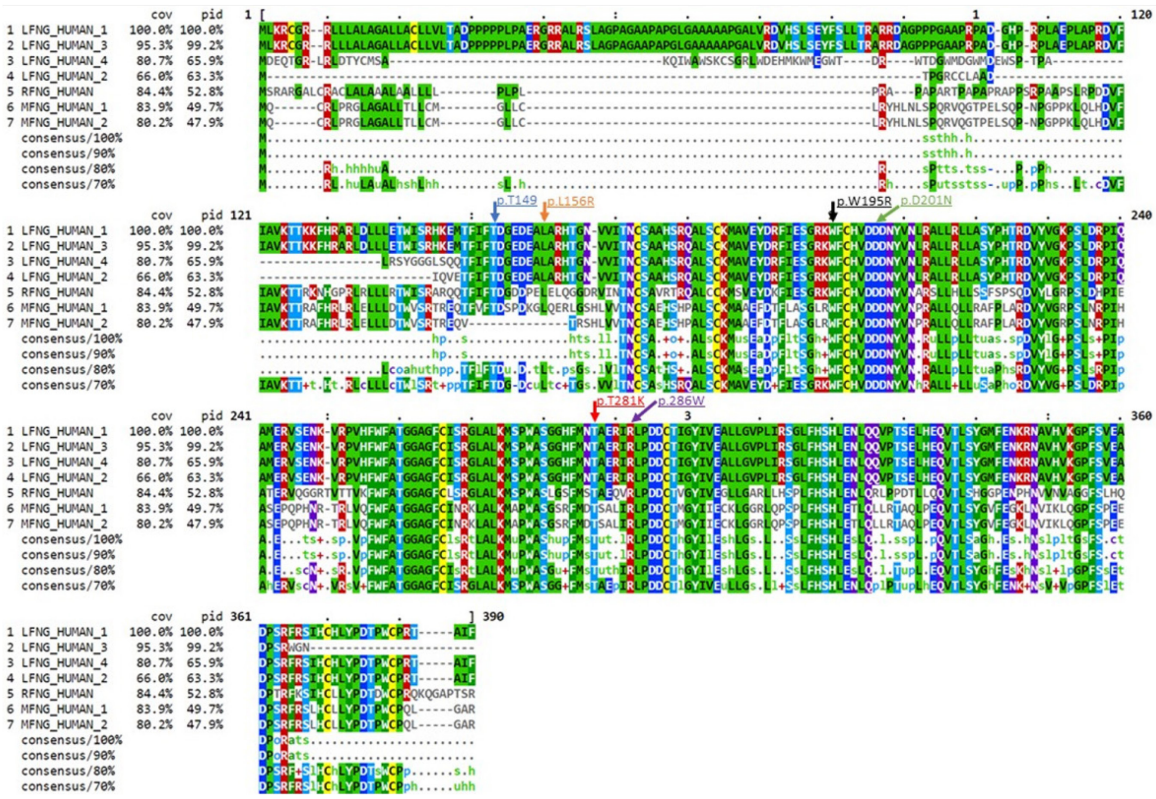
Original Fasta File	Sequence Name	Original Fasta File	Sequence Name
ORTHOLOGOUS SEQUENCES		PARALOGOUS SEQUENCES	
SP | Q8NES3-1	LFNG_HUMAN(_1)	SP | Q8NES3-2	LFNG_HUMAN_2
SP | Q2KJ92	LFNG_COW	SP | Q8NES3-3	LFNG_HUMAN_3
SP | O09010	LFNG_MOUSE	SP | Q8NES3-4	LFNG_HUMAN_4
SP | Q924T4	LFNG_RAT	SP | Q9Y644	RFNG_HUMAN
SP | P79948	LFNG_XENOPUS	SP | O00587-1	MFNG_HUMAN_1
SP | Q8JHF2	LFNG_ZEBRAFISH	SP | O00587-2	MFNG_HUMAN_2

B



Supplementary Figure 3: Orthologous and Paralogous Analysis of LFNG Variants.

C



Supplementary Figure 3: Orthologous and Paralogous Analysis of LFNG Variants. **A)** Original Swiss-prot codes derived from UniprotKB, with associated sequence names to facilitate analysis during alignment. **B)** Through orthologous comparison of 162 LFNG sequences, all residues of interest (p.T149; light blue p.L156, p.W195, p.D201, p.T281, p.R286) appear evolutionarily conserved through a minimum sequence conservation of 90%. **C)** Paralogous comparison of human Fringe proteins LFNG, RFNG, and MFNG indicates sequence conservation of relevant residues with a conservation between sequences of 80-100%.