Identification of the *mtk50f* mutation in *Vangl2* in zebrafish

15 August 2023

ADVISORS

University Advisor, Biological Expertise : Dr. Terry Van Raay

University Advisor, Informatics Expertise : Dr. Lewis Lukens

PREPARED BY

Student name: Olusegun Odumosu

Student number: 1251709

For BINF*6999: Bioinformatics Master's Project

TABLE OF CONTENTS

Abstract…… III

Acknowledgements…… IV

List of Tables…… V

List of Figures…… VI

List of Abbreviations…… VII

Supporting documentation…… VIII

1 Introduction…… 1

2 Methods…… 4

2.1 Data Collection…… 4

2.2 Data analysis…… 4

2.2.1 Overview…… 4

2.2.2 Alignment - Bwa …… 5

2.2.3 Variant calling – BCFtools mpileup and freebayes…… 6

2.2.4 Variant annotation - SnpEff……6

2.2.5 Integrated Genome Viewer (IGV) – Broad Institute …… 6

3 Results…… 7

4 Discussion and Conclusion…… 12

4.1 Alignment …… 12

4.2 Variant calling …… 13

4.3 The *Vangl2*^{mtk50f} mutant allele …… 14

4.4 Limitations …… 15

4.5 Conclusion …… 15

5 References…… 17

Abstract

In early embryo development, precise morphological changes shape tissues, organs, and the gut, orchestrated by signaling pathways like Wnt, governing migration, polarity, and neural patterning. Wnt dysregulation connects to disorders, including skeletal issues, cancers, and neural tube defects (NTDs). The Wnt/Planar cell polarity (Wnt/PCP) pathway, part of non-canonical Wnt signaling, is vital for epithelial polarity and convergent extension. Van Gogh-like protein 2 (Vangl2) is pivotal in the Wnt/PCP pathway, scaffolding core components. *Vangl2* mutants, notably *Vangl2*^{mtk50f}, *Vangl2*^{m209}, and *Vangl2*^{m747} exhibit variable sensitivities to canonical Wnt ligand Wnt8a. While the *m209* and *m747* alleles have been identified, the genetic basis of the *mtk50f* allele remains unknown and is essential for understanding pathway interactions. This study probes the *Vangl2*^{mtk50f} mutation's genetics. Results unveil a multi-nucleotide polymorphism (MNP) in intron 5 that is predicted to generate a variation of the polyadenylation signal and cause premature termination. The mutation causes abnormal transcripts without exons 6-8 that retain part of intron 5. Though constrained by sample size and theory-based methods, this study lays groundwork for future *Vangl* gene research. Crucially, *Vangl* mutations uniquely induce NTDs, underlining research significance. By shedding light on conditions like spina bifida and anencephaly, this analysis advances comprehension of Wnt signaling-related human disorders.

Keywords: Vangl2, Transcription, Wnt signaling, neural tube defect (NTD)

Acknowledgements

I am deeply grateful to Dr. Terry Van Raay and Dr. Lewis Lukens for their invaluable guidance, support, and expertise throughout my research project. Their dedication and insightful advice significantly contributed to the successful completion of this endeavor.

Dr. Terry Van Raay's extensive knowledge and experience in biology provided me with a strong foundation for understanding the intricacies of the research subject. His mentorship not only enriched my scientific understanding but also fostered critical thinking and problem-solving skills.

I extend my sincere appreciation to Dr. Lewis Lukens for his guidance in the computational aspects of the project. His expertise in data analysis and computational methods was instrumental in unraveling complex patterns within the research data. His patient guidance and willingness to address queries greatly enhanced the quality of my work.

The combined mentorship of Dr. Terry Van Raay and Dr. Lewis Lukens has been invaluable in shaping my research journey. I am fortunate to have had the opportunity to learn from their wisdom and experience. My gratitude to both supervisors is beyond measure.

I would also like to thank the University of Guelph for providing me with the tools and skills necessary to complete this project. I am ever grateful for the relationships I made through out the program that led to my academic, professional, and personal development. Of note, I would like to thank Nishita Sharif, Daniel Amoako, Omar Khan, Amjad Osman, Jesse Wolf, and Kasra Foroughi.

List of Tables

Table 1. Percentage alignment rates of sequences to reference genome…… 8

Table 2. Variant calling statistics of bcftools and freebayes…… 9

List of Figures

Figure 1. Different mutant *Vangl2* alleles have differing sensitivities to exogenous Wnt8……. 3

Figure 2. Overview of data analysis pipeline.…… 5

Figure 3. Read coverage in IGV for mutant and wildtype sequences aligned against reference zebrafish genome ……. 10

Figure 4. A closure look at exon 5 and partly retained intron 5 …… 11

Figure 5. Schematic representation of the muti-nucleotide polymorphism (MNP)…… 12

List of Abbreviations

BWA	Burrows-Wheeler Aligner
BWT	Burrows–Wheeler Transform
CE	Convergent extension
IGV	Integrated Genome Viewer
MNP	Multi-nucleotide polymorphisms
NGS	Next-generation sequencing
NTD	Neural tube defects
PAS	Polyadenylation site
SNP	Single nucleotide polymorphism
Vangl2	Van Gogh-like 2
WGS	Whole-genome sequencing
Wnt/PCP	Wnt/Planar cell polarity

Supporting documentation

1. Githuib link to code

https://github.com/odumosuo/zebrafish_mutation

1. Introduction

During early development, the embryo undergoes specific morphological changes that are crucial to the generation of highly organized tissues, organs, and gut. These processes are controlled by several signalling pathways that work in synergy. Among these are the Wnt signaling pathways that are involved in cell migration, cell polarity, cell fate specification, and neural patterning (Komiya & Habas, 2008). The expression of Wnt proteins is spatiotemporally regulated during development and dysregulation of Wnt signalling has been shown to be associated in human diseases such as skeletal defects, several cancers, and neural tube defects (NTDs) (Komiya & Habas, 2008). The Wnt/Planar cell polarity (Wnt/PCP) pathway is a branch of the non canonical Wnt signalling pathway that is pivotal in regulating the polarity of epithelial sheets and convergent extension (CE) to help shape the central nervous system and drive elongation of the body axis (Miles et al., 2017; Shindo et al., 2019). The Van Gogh-like protein 2 (Vangl2) is a transmembrane protein that is a member of the Wnt/PCP signalling pathway and has been found to be necessary for the pathway and serves as a scaffold for the other core components of the pathway to assemble (Dreyer et al., 2022). This protein is encoded by the *Vangl2* gene.

The canonical Wnt/β-catenin signaling pathway is another branch of the Wnt signalling pathway that is responsible for regulation of cell proliferation and gene transcription (Liu et al., 2022). It has been established that in vertebrates the Wnt/PCP pathway can be an antagonist to the canonical Wnt/β-catenin signaling pathway (Torres et al., 1996). *Vangl2* mutants have been found to be hypersensitive to canonical Wnt/β-catenin signaling although they do not show modification in the pathway itself (Angonin & Van Raay, 2013). Among these *Vangl2* mutations,

some are more sensitive to the overexpression of the canonical Wnt/β-catenin ligand Wnt8a as seen in Figure 1 (Horn, 2023). The three mutant homozygous alleles *m209, m747*, and *mtk50f* exhibit the classic phenotypes seen in *Vangl2* mutants with inhibited convergent extension resulting in shorter embryonic axis and dorsal flexure (Solnica-Krezel et al., 2002). However, when the same dose of exogenous Wnt8a is injected into the mutant embryos, the mutant *Vangl2*^{mkt50f} zebrafish show less exacerbated phenotypes than *Vangl2*^{m209} and *Vangl2*^{m747} (Horn, 2023).

The *Vangl2*^{m747} allele is cause by a nonsense mutation at Serine 427 producing a shorter length Vangl2 by 93 amino acids (Figure 1) (Jessen et al., 2002). The *Vangl2*^{m209} allele carries a 13-base pair insertion of intronic sequence causing a frameshift at Alanine 441 and early termination of translation. The cause of the *Vangl2*^{mkt50f} mutant allele is yet to be characterized. Jessen et al. (2002) were unable to find mutant transcripts for *Vangl2*^{mkt50f} homozygous mutant embryos but concluded that at least part of the coding sequence was deleted based on their genomic PCR analysis. This finding suggests that the mutation causes nonsense-mediated decay (Horn, 2023). The evident difference in sensitivity of the mutant *Vangl2*^{mkt50f} allele to canonical Wnt/β-catenin signaling underscores the importance of understanding the genetic basis for this mutation which could further our understanding of the relationship between the non canonical Wnt/PCP pathway and the canonical Wnt/β-catenin signaling, as well as our understanding of the Vangl2 protein.

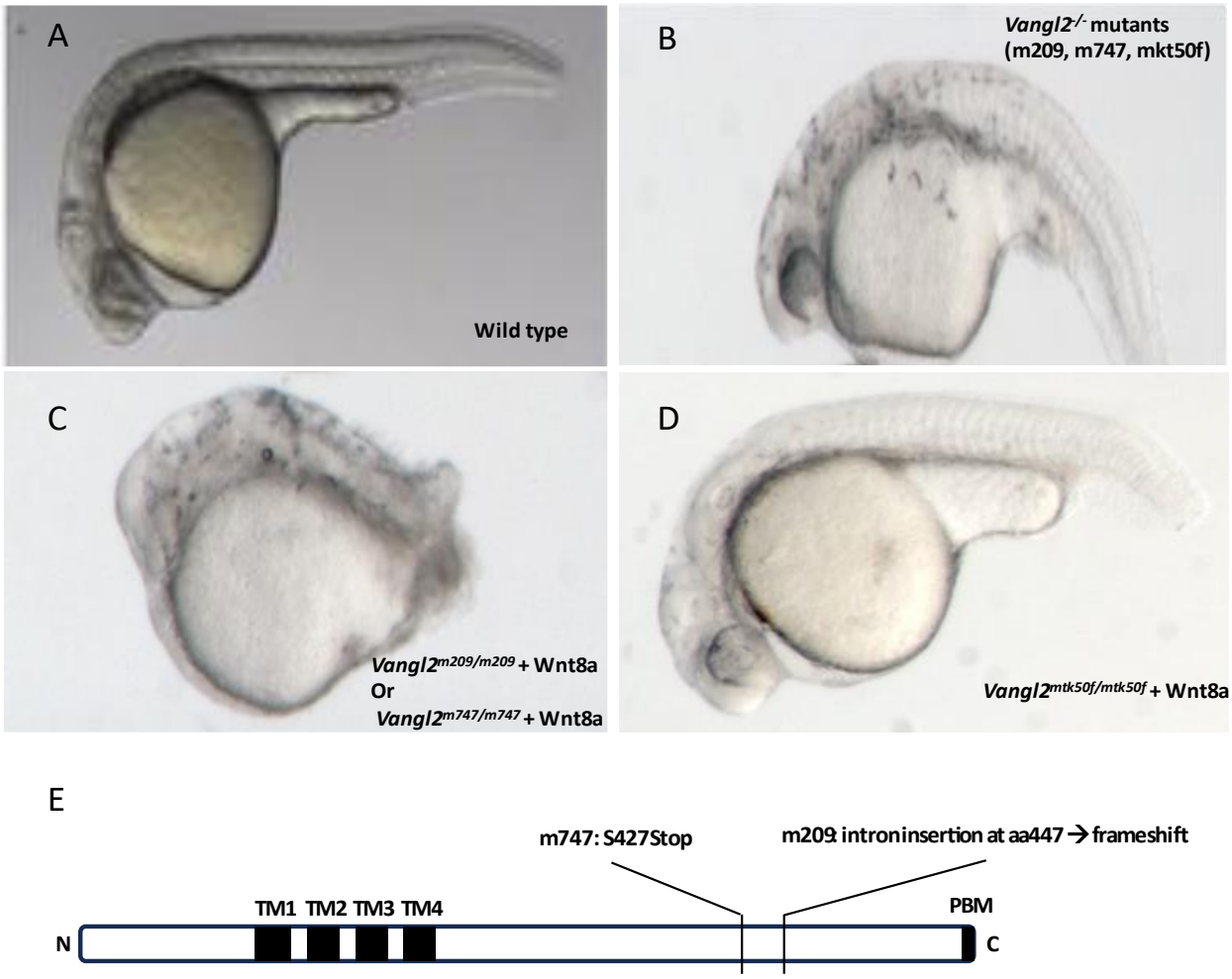

Figure 1. Different mutant *Vangl2* alleles have differing sensitivities to exogenous Wnt8. **(A)** wildtype zebrafish embryos. (**B**) *Vangl2*$^{-/-}$ mutant embryos (*m209, m747, mtk50f*). (**C**) *Vangl2*$^{m747/m747}$ or *Vangl2*$^{m209/m209}$ zebrafish embryos injected with Wnt8a. (**D**) *Vangl2*$^{mtk50f/mtk50f}$ zebrafish embryos injected with Wnt8a. (**E**) A schematic representation of Vangl2 showing mutations in *Vangl2*^{m747} and *Vangl2*^{m209} alleles. (TM, putative transmembrane domains; PDZ-DBM, putative PDZ-domain binding motif). Figure is adapted from Angonin and Van Raay (2013)

2. Methods

2.1 Data Collection

Data used for analysis was gotten from the Van Raay lab. Two heterozygous *Vangl2*$^{+/mtk50f}$ were crossed together to create a filial generation of homozygous wildtypes, heterozygotes, and homozygous mutants for the *mtk50f* allele. 10 embryos with the wildtype phenotype (homozygous wildtypes and heterozygotes) were pooled together and 10 embryos with mutant phenotype (homozygous mutants) were pooled for sequencing. The two pooled samples (wildtype and mutant) were subsampled among four lanes each in the flow cell of Illumina sequencing to obtain mRNA sequences from paired-end reads of 200 base pairs.

2.2 Data analysis

2.2.1 Overview

The fastq files containing the 200 base pair paired-end reads were aligned to the GRCz11 zebrafish reference genome using Burrows-Wheeler Aligner (BWA). Sequence differences between the samples and reference genome were identified using two variant calling software: BCFtools mpileup and freebayes. The variant sites were filtered to only include records with a quality score of 20 or more. A quality score of 20 means that there is a 99% chance that a variant is found at the site (*Qual, QD, and GQ formulation*. Illumina. , n.d.). The variant cites were then annotated using SnpEff. An Integrated Genome Viewer (IGV) was used to visually inspect the alignment and compare the wildtype and mutant samples. A general overview of the data analysis pipeline can be seen in Figure 2.

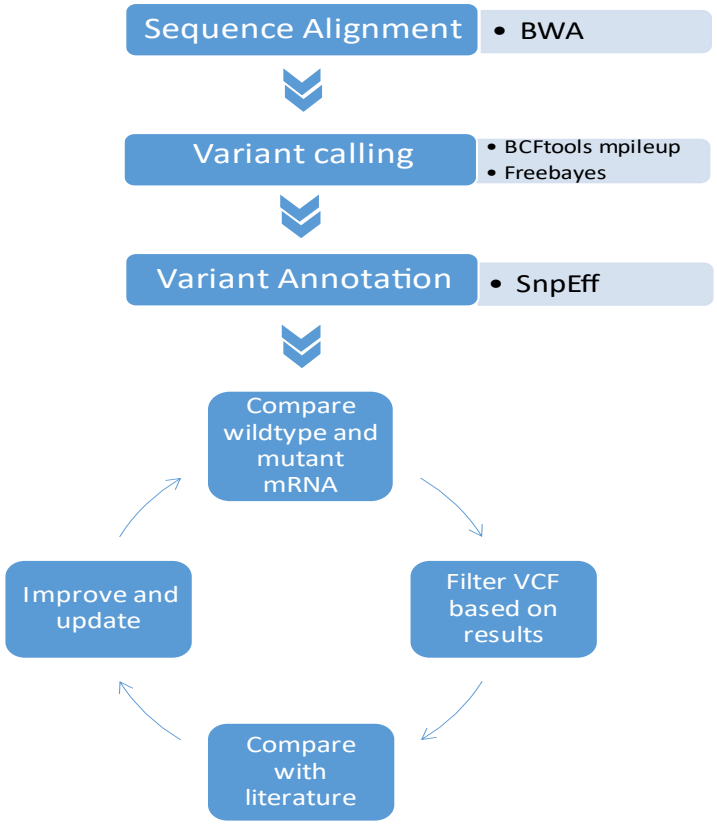

Figure 2. Overview of data analysis pipeline. The figure shows a schematic of the various steps involved in the data analysis and the various tools used at each step.

2.2.2 Alignment - Bwa

The choice of alignment tool is frequently determined by the type of next-generation sequencing (NGS) application. The BWA tool is an alignment package that uses Burrows–Wheeler Transform (BWT) to align short sequencing reads against a reference sequence (Li & Durbin, 2009). The tool supports gapped alignment unlike some other tools like MAQ making it suitable for our analysis. (Li & Durbin, 2009).

2.2.3 Variant calling – BCFtools mpileup and freebayes

Both bcftools and freebayes use a Bayesian statistical framework (Liu et al., 2022). Bcftools is one of the most used single nucleotide polymorphism (SNP) and indel calling pipelines and was found to be the best choice to detect SNPs from chicken NGS data when compared with six other SNP calling pipelines (Liu et al., 2022). Freebayes is designed to find SNPs, indels, multi-nucleotide polymorphisms (MNPs), and composite indel and substitution events (Garrison & Marth, 2012; Liu et al., 2022). We employed both tools in conjunction to attempt to capture a multitude of genetic variations.

2.2.4 Variant annotation - SnpEff

SnpEff is a computer program that annotates variants based on their genomic locations and predicts coding effects on proteins (Cingolani et al., 2012). The program uses predetermined variants in a data file that has the nucleotide change. The program provides annotations for the variants according to their position in the genome, which could be intronic, untranslated regions (5’ UTR or 3’ UTR), at splice sites and so on. It also predicts potential impacts on coding sequences such as synonymous or non-synonymous amino acid substitutions, frame shifts, or gains and losses of start and stop codons. It is important to note that the tool annotates variants independently without considering potential effects of grouped variations (Degalez et al., 2021).

2.2.5 Integrated Genome Viewer (IGV) – Broad Institute

The IGV by Broad Institute was used to visually explore and analyze the genomic data. Humans have been found to be adept in recognizing patterns, a task that is relatively simple for

the brain but computationally expensive (Korteling, 2021). By utilizing both computational and human resources, we were better able to perform our analysis quantitatively and qualitatively.

3. Results

For all samples, the alignment rate with the BWA software was greater than 99% as seen in Table 1. The wildtype samples had a higher alignment rate with a mean of 99.38% where the mutant samples had a mean of 99.28%. The difference was found to be statistically significant $t(6)= 5.9409$, $p =0.000508$. The two variant callers BCFtools and freebayes both performed relatively well. From Table 2 we can see that neither software outperformed the other on all types of variations. BCFtools had a higher number of variant calls, SNPs, indels and multiallelic allele sites, whereas freebayes had a higher number of MNPs, multiallelic sites, and other types of variations not already mentioned. Both software had 0 calls for sites with no alternative alleles (no-ALTs) because the variant calling step was done with an argument restricting calls to loci that had variant sites.

Viewing the alignment from the IGV, we see that most reads align to the expected exon positions in the *Vangl2* gene as can be seen in Figure 3 (Exons - Danio_rerio - Ensembl genome browser 110, n.d.). Exons one through eight are present in the wildtype samples, however, exons six, seven, and eight are missing in the mutant samples. Additionally, we see that intron five is partly retained in both the wildtype and mutant samples (Figure 3B and Figure 4A). At the 3’ end of the retained intron there is an MNP that includes three substitutions and a three-base deletion that gives rise to AAGAAATTTAAT rather than the AAAATCGATTACAT sequence seen in the reference genome (Figure 4A-C & Figure 5). This MNP is found consistently in all reads in the

intron (Figure 4B) and results in a variant polyadenylation signal sequence AAGAAA (Figure 4C & Figure 5).

Table 1

Percentage alignment rates of sequences to reference genome

Sample	Alignment rate (%)
Mutant_L001	99.23
Mutant_L002	99.29
Mutant_L003	99.30
Mutant_L004	99.28
Wildtype_L001	99.36
Wildtype_L002	99.37
Wildtype_L003	99.39
Wildtype_L004	99.38

Table 1 above shows the percentage alignment rates using BWA of the two pooled samples (wildtype and mutant) subsampled among four lanes each in the flow cell of Illumina sequencing.

Table 2

Variant calling statistics of bcftools mpileup and freebayes

Statistic	BCFtools mpileup	Freebayes
number of samples	8	8
number of records	**953014**	789435
number of no-ALTs	0	0
number of SNPs	**858987**	638800
number of MNPs	0	**75177**
number of indels:	**94027**	66166
number of others	0	**21106**
number of multiallelic sites	4947	**17481**
number of multiallelic SNP sites	**1718**	1529

Table 2 above shows the statistics for the two variant software tools used: BCFtools mpileup and freebayes. Bolded numbers are types of variation that were higher between the two software Where no-ALTs is no alternative allele, SNPs is single nucleotide polymorphisms, and MNPs is multiple nucleotide polymorphism

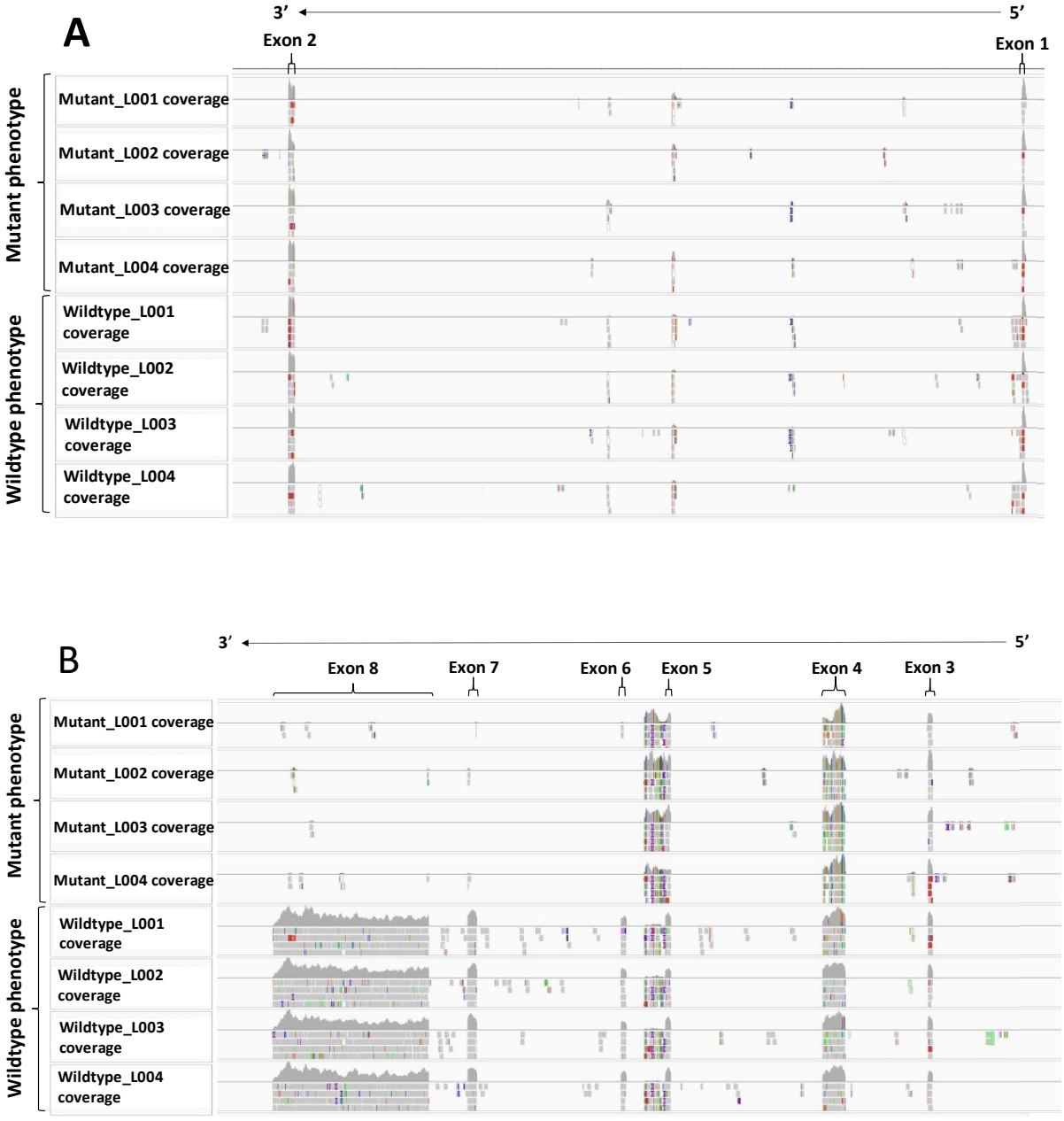

Figure 3. Read coverage in IGV for mutant and wildtype sequences aligned against reference zebrafish genome. Exon positions were obtained from the ensemble database (Exons - Danio_rerio - Ensembl genome browser 110, n.d.) **(A)** Figure showing the positions of exons 1 and 2. **(B)** Figure showing the positions of exon 3-8. Exons 6-8 are missing in mutant samples.

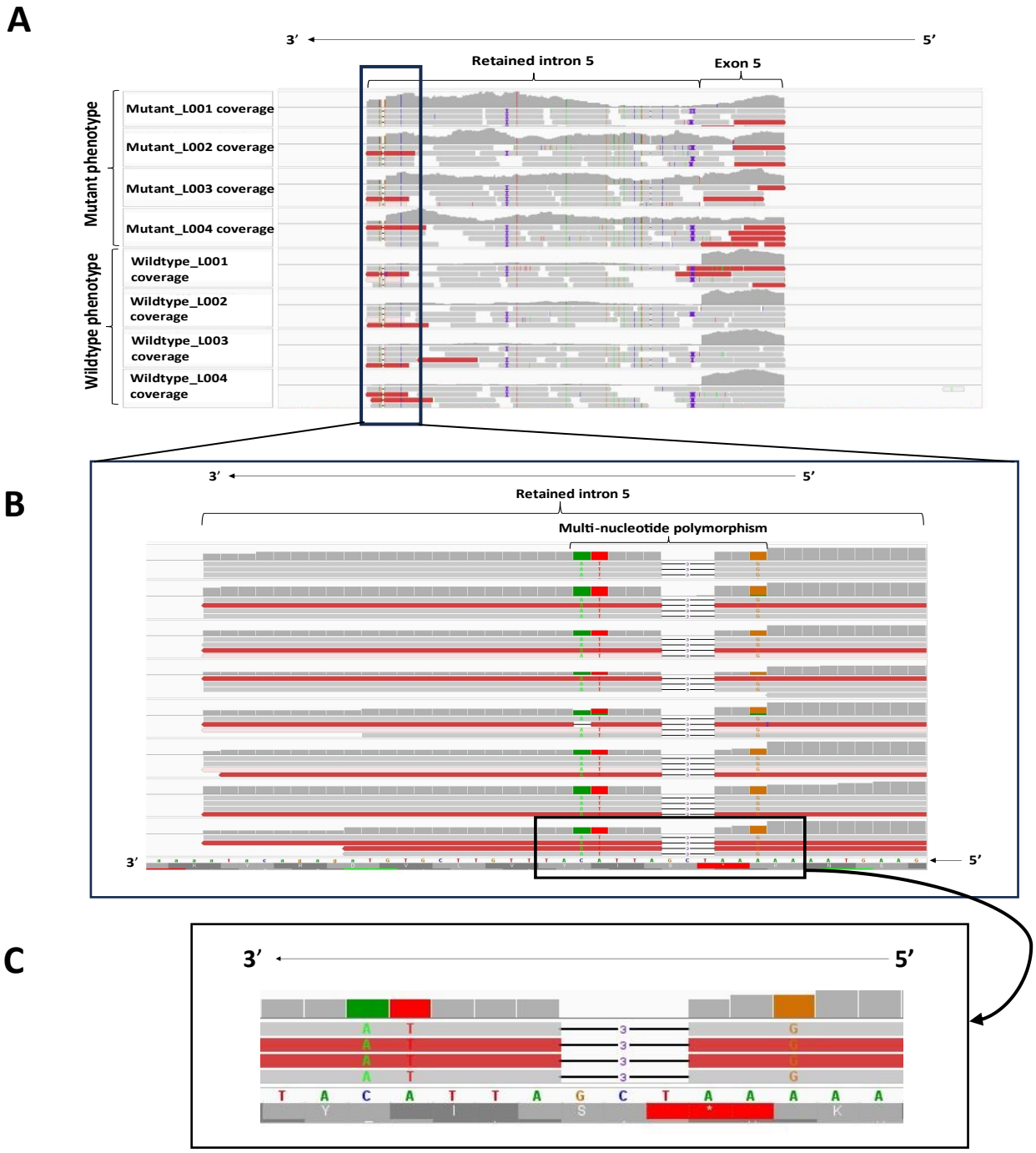

Figure 4. A closure look at exon 5 and partly retained intron 5. (A) Figure shows the read coverage for mutant and wildtype samples within exon 5 and retained intron 5. **(B)** Figure shows the 3′ end of partly retained intron 5. It shows the multi-nucleotide polymorphism (MNP). The MNP includes three substitutions and a three-base deletion. **(C)** A close up of part of the MNP

that results in the variant polyadenylation signal AAGAAA. Read 5’ to 3’ the reference sequence is AAAATCGATTACAT and the mutant sequence is AAGAAATTTAAT.

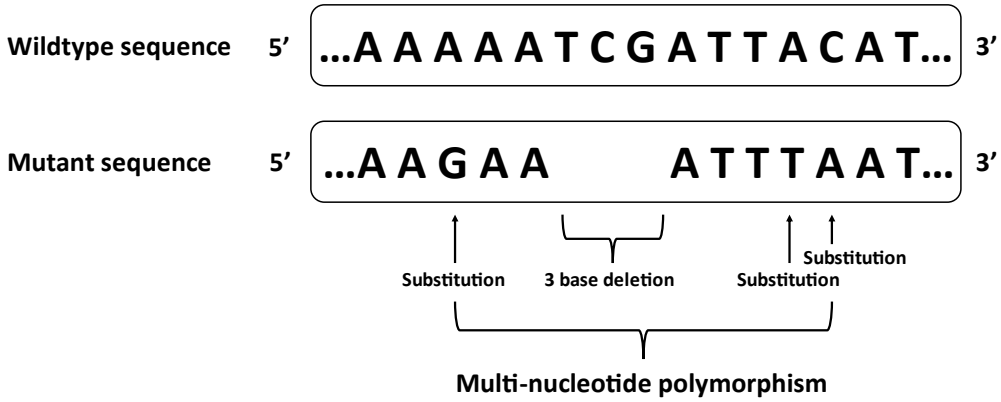

Figure 5. Schematic representation of the muti-nucleotide polymorphism (MNP). The MNP includes a three base deletion and 3 substitution events. Part of the MNP results in the variant polyadenylation signal AAGAAA.

4. Discussion

4.1 Alignment

The accuracy of alignment can be affected by the quality of the reference genome, sequencing technology, sequence quality, and alignment parameters. The latest zebrafish assembly, GRCz11’s quality is high enough that it allowed researchers annotate 26,000 protein-coding genes, making it the biggest gene set of any vertebrate sequenced at the time of it’s release (Howe et al., 2013). Although there was a significant difference between the alignment

rate of the wildtype and mutant samples, this difference was of no consequence for our analysis because all samples had alignment rates greater than 99%.

4.2 Variant calling

BCFtools mpileup called more variants than freebayes, however it was unable to detect MNPs and called 0 MNPs, whereas freebayes called 75177. An MNP is a set multiple adjacent SNPs. Liu et al. (2022) reported that BCFtools calls variants such as SNPs and indels while freebayes call SNPs, indels, MNPs and complex events. The MNP in Figure 4B at position 5334069 on chromosome 7 was called by freebayes but BCFtools failed to call any variants in that region. It has been reported that most variant callers and annotation tools fail to consider whether neighbouring variants belong to the same haplotype (Srinivasan et al., 2021). It has been found that there can be significant discrepancies between different variant calling tools and it is advisable to compare the outcomes of several variant callers (O'Rawe et al., 2013)

4.3 The *Vangl2*^{mtk50f} mutant allele

The mutant samples are missing exons six, seven, and eight. Additionally, we see that these samples have a partial retention of intron five, inserting 568 bases of this intron in the mRNA transcript. Figure 4B shows that this partial intron retention is in both the mutant phenotype and wildtype phenotype samples. We expect that homozygous wildtype individuals would have mRNA transcripts that splice normally, including all expected exons and excluding all introns. We propose that that the reads in the wildtype sample with intronic sequences are from the heterozygous individuals that possess the *Vangl2*^{mtk50f} mutant allele. This proposition is corroborated by the fact that the ratio of the amount intronic sequence reads to the amount

of reads in exon five is much smaller in the wildtype sample than in the mutant sample as can be seen in Figure 4A. We see that the mutant sample has a relatively steady amount of reads in exon five through intron five however, the intronic reads in the wildtype sample is minute compared to reads exon five for the same sample (Figure 4A). These findings imply that the normal splicing process is disrupted in *Vangl2*^{mtk50} mutants. This disruption could potentially result from either premature transcription termination or a combination of both factors.

Transcription termination of mRNA-coding genes in eukaryotes is coupled to RNA 3'-end processing (Xie et al., 2023). After the pre-mRNA is cleaved in the 3'-untranslated region a poly(A) tail is added in the process of polyadenylation. Both cleavage and polyadenylation take place at the polyadenylation site (PAS) (Zhang et al., 2021). The most important signal of the PAS is the polyadenylation signal (Zhang et al., 2021). The polyadenylation signal sequence is a consensus hexamer of AAUAAA or AUUAAA (Beaudoing et al., 2000). However different variants of the site have been reported in literature. Beaudoing et al. (2000) reported that ten single-base variants of AAUAAA accounted for 14.9% of polyadenylation signals. The MNP found in *Vangl2*^{mtk50f} mutants at the 3' end of intron 5 give rise to a variation of the polyadenylation signal sequence reported in literature – AAGAAA (Figure 4B-C & Figure 5) (Beaudoing et al., 2000). This signal sequence is found in 1.1% of polyadenylation sites and is on average 8-30 base pairs from the cleavage site (Beaudoing et al., 2000). Our analysis shows that the site is 25 base pairs upstream of the end of the transcript in *Vangl2*^{mtk50f} mutant alleles (Figure 4B). This premature transcription termination would explain why exons 6-8 are missing in the mRNA of mutant samples because they were never transcribed. It would also explain why part of intron 5 is retained because the 5' splice donor site of intron 5 did not have a corresponding 3' splice

acceptor site to form the needed lariat to excise the intron (Gehring & Roignant, 2021). It has also been reported that this specific variation of the polyadenylation signal sequence, AAGAAA inhibited in vitro splicing of upstream introns in vertebrates (Niwa & Berget, 1991)

4.4 Limitations

The study is comprised of 20 individual zebrafish embryos: 10 phenotypically wildtype samples and 10 phenotypically mutant samples. The study would benefit from a larger sample size that will be a better representation of the populations for more generalizable and accurate results. The analysis is also limited by the fact that only mRNA sequencing data is used. The mRNA data allowed us to investigate the functional effects of the *mtk50f* mutation but restricts the extent of the analysis in that large portions of the genome are missing. Whole-genome sequencing (WGS) would allow for a more in-depth analysis of the mutation. Additionally, WGS would allow for the exploration of the U/GU sequence that is usually found downstream of the cleavage site (Varani & Pérez Cañadillas, 2003). Finally, replicating the hypothesized basis of the mutation experimentally to test whether the MNP is sufficient to give rise to the phenotypes seen in mutant *Vangl2*^{mtk50f} zebrafish would be conclusive evidence of the cause of the mutation.

4.5 Conclusion

Mutant *Vangl2*^{mtk50f} zebrafish have been found to be less sensitive to canonical Wnt/β-catenin signaling compared to *Vangl2*^{m209} and *Vangl2*^{m747} mutants. It has been established that in vertebrates the Wnt/PCP pathway can be an antagonist to the canonical Wnt/β-catenin signaling pathway (Torres et al., 1996). Understanding the genetic basis of this mutation could

further our understanding of the Wnt signalling pathways. The MNP in intron 5 gave rise to a variation of the polyadenylation signal and caused premature termination. This premature termination resulted in transcripts of mutant *Vangl2*^{mtk50f} that lacked exons 6-8 and partly retained intron 5. The study is limited by sample size, lack of whole genomic sequence of the line used to create the mutant, and a solely theory-based approach. However, it lays the groundwork for further research pertaining to the *Vangl* gene. The importance of this research should not be understated as it is purported that *Vangl* genes are the only genes where mutations cause NTDs (Iliescu et al., 2011). This analysis will contribute to research pertaining to human conditions including spina bifida, anencephaly and other NTDs.

5. References

Angonin, D., & Van Raay, T. J. (2013). Nkd1 functions as a passive antagonist of Wnt signaling. *PloS One*, *8*(8), e74666–e74666. https://doi.org/10.1371/journal.pone.0074666

Beaudoing, E., Freier, S., Wyatt, J. R., Claverie, J. M., & Gautheret, D. (2000). Patterns of variant polyadenylation signal usage in human genes. *Genome Research*, *10*(7), 1001–1010. https://doi.org/10.1101/gr.10.7.1001

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., Land, S. J., Lu, X., & Ruden, D. M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w 1118 ; iso-2; iso-3. *Fly (Austin, Tex.)*, *6*(2), 80–92. https://doi.org/10.4161/fly.19695

Degalez, F., Jehl, F., Muret, K., Bernard, M., Lecerf, F., Lagoutte, L., Desert, C., Pitel, F., Klopp, C., & Lagarrigue, S. (2021). Watch Out for a Second SNP: Focus on Multi-Nucleotide Variants in Coding Regions and Rescued Stop-Gained. *Frontiers in Genetics*, *12*, 659287–659287. https://doi.org/10.3389/fgene.2021.659287

Dreyer, C. A., VanderVorst, K., & Carraway, rd. (2022). Vangl as a Master Scaffold for Wnt/Planar Cell Polarity Signaling in Development and Disease. *Frontiers in Cell and Developmental Biology*, *10*, 887100–887100. https://doi.org/10.3389/fcell.2022.887100

Exons - Danio_rerio - Ensembl genome browser 110. (n.d.). http://useast.ensembl.org/Danio_rerio/Transcript/Exons?db=core%3Bg

Garrison, E., & Marth, G. (2012). *Haplotype-based variant detection from short-read sequencing*. https://doi.org/10.48550/arxiv.1207.3907

Gehring, N. H., & Roignant, J.-Y. (2021). Anything but Ordinary – Emerging Splicing Mechanisms in Eukaryotic Gene Regulation. *Trends in Genetics*, *37*(4), 355–372. https://doi.org/10.1016/j.tig.2020.10.008

Horn, M. (2023). *Elucidating the Role of Vangl2 and Its Interactions with Wnt/Beta-Catenin Signaling* [Master's thesis, University of Guelph]. https://atrium.lib.uoguelph.ca/xmlui/bitstream/handle/10214/27555/Horn_Matthew_202305_Msc.pdf?sequence=1&isAllowed=y

Howe, K., Torroja, C. ., Torrance, J., Collins, J. ., Humphray, S., McLaren, K., McLaren, S., Sealy, I., Caccamo, M., Churcher, C., Scott, C., Barrett, J. ., Rauch, G. ., Chow, W., Kilian, B., Quintais, L. ., Guerra-Assuncao, J. ., Zhou, Y., Eyre, T., … Fu, B. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature (London)*, *496*(7446), 498–503. https://doi.org/10.1038/nature12111

Jessen, J. R., Topczewski, J., Bingham, S., Sepich, D. S., Marlow, F., & Chandrasekhar, A. (2002). Zebrafish trilobite identifies new roles for Strabismus in gastrulation and neuronal movements. *Nature Cell Biology*, *4*(8), 610–615. https://doi.org/10.1038/ncb828

Komiya, Y., & Habas, R. (2008). Wnt signal transduction pathways. *Organogenesis*, *4*(2), 68–75. https://doi.org/10.4161/org.4.2.5851

Korteling, J. E. H., van de Boer-Visschedijk, G. C., Blankendaal, R. A. M., Boonekamp, R. C., & Eikelboom, A. R. (2021). Human- versus Artificial Intelligence. *Frontiers in Artificial Intelligence*, *4*, 622364–622364. https://doi.org/10.3389/frai.2021.622364

Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, *25*(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324

Liu, J., Shen, Q., & Bao, H. (2022). Comparison of seven SNP calling pipelines for the next-generation sequencing data of chickens. *PLoS ONE*, *17*(1), e0262574. https://doi.org/10.1371/journal.pone.0262574

Liu, J., Xiao, Q., Xiao, J., Niu, C., Li, Y., Zhang, X., Zhou, Z., Shu, G., & Yin, G. (2022). Wnt/β-catenin signalling: function, biological mechanisms, and therapeutic opportunities. *Signal Transduction and Targeted Therapy*, *7*(1), 3–3. https://doi.org/10.1038/s41392-021-00762-6

Miles, L. B., Mizoguchi, T., Kikuchi, Y., & Verkade, H. (2017). A role for planar cell polarity during early endoderm morphogenesis. *Biology Open*, *6*(5), 531–539. https://doi.org/10.1242/bio.021899

Niwa, M., & Berget, S. M. (1991). Mutation of the AAUAAA polyadenylation signal depresses in vitro splicing of proximal but not distal introns. *Genes & Development*, *5*(11), 2086–2095. https://doi.org/10.1101/gad.5.11.2086

O'Rawe, J., Jiang, T., Sun, G., Wu, Y., Wang, W., Hu, J., Bodily, P., Tian, L., Hakonarson, H., Johnson, W. E., Wei, Z., Wang, K., & Lyon, G. J. (2013). Low concordance of multiple variant-calling pipelines: practical implications for exome and genome sequencing. *Genome Medicine*, *5*(3), 28–28. https://doi.org/10.1186/gm432

Qual, QD, and GQ formulation. Illumina. (n.d.). https://support.illumina.com/content/dam/illumina-

support/help/Illumina_DRAGEN_Bio_IT_Platform_v3_7_1000000141465/Content/SW/Informatics/Dragen/QUAL_QD_GQ_Formulation_fDG.htm

Shindo, A., Inoue, Y., Kinoshita, M., & Wallingford, J. B. (2019). PCP-dependent transcellular regulation of actomyosin oscillation facilitates convergent extension of vertebrate tissue. *Developmental Biology*, *446*(2), 159–167. https://doi.org/10.1016/j.ydbio.2018.12.017

Srinivasan, S., Kalinava, N., Aldana, R., Li, Z., van Hagen, S., Rodenburg, S. Y. A., Wind-Rotolo, M., Qian, X., Sasson, A. S., Tang, H., & Kirov, S. (2021). Misannotated Multi-Nucleotide Variants in Public Cancer Genomics Datasets Lead to Inaccurate Mutation Calls with Significant Implications. *Cancer Research (Chicago, Ill.)*, *81*(2), 282–288. https://doi.org/10.1158/0008-5472.CAN-20-2151

Torres, M. A., Yang-Snyder, J. A., Purcell, S. M., DeMarais, A. A., McGrew, L. L., & Moon, R. T. (1996). Activities of the Wnt-1 Class of Secreted Signaling Factors Are Antagonized by the Wnt-5A Class and by a Dominant Negative Cadherin in Early Xenopus Development. *The Journal of Cell Biology*, *133*(5), 1123–1137. https://doi.org/10.1083/jcb.133.5.1123

Varani, G., & Pérez Cañadillas, J. M. (2003). Recognition of GU-rich polyadenylation regulatory elements by human CstF-64 protein. *The EMBO Journal*, *22*(11), 2821–2830. https://doi.org/10.1093/emboj/cdg259

Xie, J., Libri, D., & Porrua, O. (2023). Mechanisms of eukaryotic transcription termination at a glance. *Journal of Cell Science*, *136*(1). https://doi.org/10.1242/jcs.259873

Zhang, Y., Liu, L., Qiu, Q., Zhou, Q., Ding, J., Lu, Y., & Liu, P. (2021). Alternative polyadenylation: methods, mechanism, function, and role in cancer. *Journal of Experimental & Clinical Cancer Research*, *40*(1), 51–51. https://doi.org/10.1186/s13046-021-01852-7