Summer 2024

Exploring the Functional Significance of a YAP1 Missense Variant of Uncertain Significance in Caenorhabditis elegans

Nathan Jones
njones11@stu.jsu.edu

Exploring the Functional Significance of a *YAP1* Missense Variant of Uncertain Significance in *Caenorhabditis elegans*

A Thesis Submitted to the
Graduate Faculty
of Jacksonville State University
in Partial Fulfillment of the
Requirements for the Degree of
Master of Science
with a Major in
Biology

By
Nathan Lee Jones

Jacksonville, Alabama

August 2, 2024

Nathan Lee Jones August 2, 2024

Abstract

Polycystic ovary syndrome (PCOS) is a complex disorder with various implications, such as polycystic ovaries, visceral obesity, and increased risk of cancer. *YAP1* was recently identified as a gene of interest in the development of PCOS. Researchers have established that single nucleotide variants in *YAP1* are likely to play a role in PCOS development. This project aims to provide insight into the potential impact of a *YAP1* variant of uncertain significance (VUS). Studies in *C. elegans* have established *yap-1* as a nematode ortholog for human *YAP1*. A *YAP1* VUS was identified through ClinVar, *YAP1* c.1015A>G (p.Asn339Asp). Evolutionary conservation of the VUS loci was confirmed using multiple sequence alignments in Benchling. Additional analyses were carried out, examining the difference in amino acid chemical properties and the bioinformatics PolyPhen-2 prediction. PolyPhen-2 predicts this variant to be likely pathogenic (HumDiv score of 0.982). These findings supported further investigation, and we sought out to examine the VUS *in vivo* through the *C. elegans yap-1* ortholog and model. DNA primers were designed and optimized using PCR to amplify the *YAP1* VUS loci in the nematode *yap-1* region. A CRISPR RNA guide was designed and optimized to target *yap-1* using an *in vitro* nuclease assay. The VUS-*yap-1 C. elegans* strain was generated through microinjection of CRISPR-Cas9. The generated VUS strain will be assessed for possible phenotypes. This study has provided insight into the potential impact and significance of a patient VUS through *in silico* and *in vivo* studies. Future bioinformatic experimentation will include protein modeling to examine potential structural changes between encoded wildtype and mutant proteins.

Acknowledgements

Thank you to my friends and family that been a source of support throughout this journey. Thank you to Dr. Ashley Turner for her great mentorship and continued guidance. I would also like to thank my committee members, Dr. Jenna Ridlen and Dr. Chris Murdock, for their support and guidance throughout the project. I would also like to thank the JSU Biology Department for providing an environment that has provided me an environment that allowed me to enjoy and explore science.

This thesis is dedicated to an amazing grandmother. Thank you for always encouraging me to always dream big and to never stop chasing them.

Nathan Jones

Table of Contents

Page

List of Tables

Page

List of Figures

Page

INTRODUCTION

The functional consequences and clinical implications of single nucleotide variants (SNVs) identified within the human genome remain classified as a variant of uncertain significance (VUS). (Federici & Soddu, 2020). The abundance of VUS mutations brings to the forefront the importance of screening mutations through model organisms to gain insight into their potential impact. An estimated 6000-13,000 disease genes remain to be identified for mendelian diseases alone (Bamshad et al., 2019). Screening variants through model organisms, such as *Caenorhabditis elegans*, is the most effective way forward for such a large amount of mutations (Baldridge et al., 2021) Shining light on the potential impact of these genetic variants, including missense mutations, on protein function is pivotal in the development of treatment and diagnosis of human disease.

Polycystic ovary syndrome (PCOS) is a chronic hormonal condition that affects 8-13% of women of reproductive age worldwide, with 70% of PCOS cases remaining undiagnosed (*Polycystic Ovary Syndrome*, n.d.). PCOS is the leading cause of menstrual complications in women worldwide and often presents itself with a seemingly endless list of possible symptoms(S & A, 2017).

Symptoms of PCOS have historically included polycystic ovaries, weight gain, irregular periods, and excess body hair. However, PCOS can present symptoms such as acne, skin tags, or thick patches of skin (*Polycystic Ovary Syndrome (PCOS)*, 2022). PCOS is also characterized by elevated levels of androgens and insulin resistance. (Puurunen et al., 2009)(Diamanti-Kandarakis & Dunaif, 2012). The insulin resistance caused by PCOS puts PCOS patients at a higher risk for developing type 2 diabetes, and PCOS has even been recommended as a disease that requires screening for type 2 diabetes (Condorelli et al., 2017). The number of symptoms PCOS may or

may not have contributes greatly to its high undiagnosed rate. This leaves physicians with the only option of conducting a battery of tests which may or may not allow for diagnosis (Tremblay-Davis et al., 2021). A small questionnaire study conducted in Canada with 296 women from ages 18-60 concluded that out of the 296 women, 34% waited over 2 years for diagnoses and 41% had to see 3 or more doctors before diagnoses (Ismayilova & Yaya, 2022). Another study of 1385 women from North America, Europe, and other world regions had been diagnosed with PCOS, 33% reported having to wait over 2 years for diagnoses, and only 35% were satisfied with how they were diagnosed(Gibson-Helm et al., 2016). This consistent, multi-national delay in PCOS diagnoses is why any effort to expand the current genetic understanding of PCOS is critical to building a pathway to improved patient care.

Current research has implicated the *YAP1* (Yes – associated protein 1) human gene as a likely player in the pathophysiology of PCOS through genome-wide association studies (GWAS). A GWAS study conducted in the Han region of China with a sample size of 1115 PCOS and 1137 controls focused on three single nucleotide polymorphisms (SNPs) found in *YAP1*. For two of the three SNPs, PCOS patients who carried allele A tolerated glucose at a higher level than the control. PCOS patients who carried Allele G were found to be at a higher risk for high luteinizing hormone than the control (Li et al., 2012). These findings established that *YAP1* appears to be a gene of interest for PCOS. A meta-analysis of large scale GWAS studies was also conducted including 10,074 women diagnosed with PCOS and 103,164 controls from 7 groups who descend from European ancestry. Index variants were found in *YAP1* to be significantly associated with PCOS (Day et al., 2018). A major conclusion the authors drew was that SNPs found in *YAP1* appear to be possibly associated with the fertility and menstrual cycle

irregularity of PCOS. It was also concluded that when compared between diagnosed PCOS patients and controls, *YAP1* loci were significantly associated with PCOS (Day et al., 2018).

YAP1 is a transcriptional coactivator and a downstream effector of the Hippo pathway. YAP1 localizes in the nucleus where through phosphorylation, aids the Hippo pathway in promoting gene expression and stimulates cell proliferation and apoptosis. *YAP1* is one of the major effectors of the Hippo pathway. Due to the nature of YAP1's critical function in a cell's life cycle, *YAP1* is highly regulated. *YAP1*'s role in disease can vary based on how *YAP1* is being regulated. It was found that when YAP1 is in high levels in patients with pancreatic cancer, the high YAP1 concentration correlates to poor patient survival (Salcedo Allende et al., 2017). However, it was also found that YAP1 is in low concentration in breast cancer and can actually be considered a tumor suppressor (Yuan et al., 2008). It appears that the role of *YAP1* in disease can vary widely, depending on how *YAP1* is currently being utilized at that moment.

YAP1 plays its role within Hippo alongside another gene, *TAZ*. Researchers, up until recently, assumed that *YAP1* and *TAZ* were redundant. However, recent studies have shown that *YAP1* is much more vital than *TAZ*. It was found that even though *YAP1* and *TAZ* share a binding domain on *TEAD, YAP1* protein levels were significantly higher than *TAZ* in HEK293 cells (Plouffe et al., 2018). Knockouts in mice of Yap1 result in embryonic lethality, whereas knockouts of Taz are still viable for a mice embryo. This points toward *YAP1* being the primary actor within Hippo over *TAZ*. However, *YAP1* and *TAZ* do have overlapping functions, but to what extent they overlap is not well understood (Plouffe et al., 2018).

YAP1 is vital for healthy ovarian function. *YAP1* is essential for granulosa cell proliferation and follicle development in the ovaries. When *YAP1* is knocked-out in mice,

ovarian follicle development is disrupted *in vitro* and *in vivo (Lv et al., 2019) YAP1* contributes greatly to female fertility by interacting with several outside signaling pathways to aid in regulating granulosa cell proliferation and differentiation (Lv et al., 2019). *YAP1* was also confirmed to be required for proliferation of ovarian granulosa cells. (Jiang et al., 2017). When hypomethylation occurs on or near the *YAP1* promoter, YAP1 becomes overexpressed in ovary granulosa cells because of hypomethylation (Jiang et al., 2017). Precise regulation of YAP1 is vital for normal ovarian function as well as preventing the development or acceleration of PCOS (Jiang et al., 2017). *YAP1* is necessary for proper granulosa cell behavior in the ovaries. Prior to ovulation, oocytes will inhibit Hippo transcripts and will instead rely on *YAP1* to promote granulosa cell survival. If *YAP1* fails to take over, then premature differentiation will occur and result in an increate rate of progesterone production (Sun & Diaz, 2019).

It has become clear that *YAP1* is necessary for proper ovarian function, both physiologically and hormonally. This importance is a major reason why *YAP1* has become a gene of interest for PCOS. The ovarian organ system contributes greatly to the development of PCOS.

Outside of PCOS, *YAP1* plays a role in several human diseases. The most visible of these diseases, being cancer. YAP1 has become accepted as a major player in several cancers (Szulzewsky et al., 2021) *YAP1* concentration was found to be a possible prognostic marker for survival outcomes of ovarian cancer as well as contributing to the development of ovarian cancer through cell proliferation and chemoresistance (Hall et al., 2010). YAP1 immensely contributes to cell over-proliferation and cell survival through its interactions with *TEAD 1-4* (Zhao et al., 2007). Overall, *YAP1* is a gene that has a driving role in various diseases, not just PCOS. Better

understanding of *YAP1*'s role in the pathology of just one of these diseases could cause a cascading affect in the understanding of many diseases.

Caenorhabditis elegans is a cornerstone model organism in modern biological research and has contributed to major advances in molecular and cell biology (Nigon & Félix, 2018). *C. elegans* is one of the most abundant creatures on the planet today, playing a vital role in major world economies, agriculture, and overall human health (Murfin et al., 2012). Additionally, *C. elegans* was the first animal to have its genome fully sequenced, a precursor along the way toward sequencing the human genome (THE C. ELEGANS SEQUENCING CONSORTIUM, 1998). Over time *C. elegans* has acquired a meticulously annotated genome, providing a clear advantage for genetic studies over many model organisms (Stein et al., 2001). Several metabolic processes are evolutionary conserved between mammals and *C. elegans*. (Hashmi et al., 2013). Not only are vital metabolic pathways conserved in *C. elegans*, but they also require the same amino acids and vitamins as humans (Watson & Walhout, 2014).

C. elegans gene *F13E6.4*, now more commonly known as *yap-1*, has been determined to be the human *YAP1* ortholog (*Yes-Associated Protein Homolog, YAP-1, Is Involved in the Thermotolerance and Aging in the Nematode Caenorhabditis Elegans - ScienceDirect*, n.d.). Mutations in *yap-1* are known to lead to extended lifespan, stress induced lethality, and heat hypersensitivity in mutant nematodes compared to wildtype (*Yes-Associated Protein Homolog, YAP-1, Is Involved in the Thermotolerance and Aging in the Nematode Caenorhabditis Elegans - ScienceDirect*, n.d.).

Currently, little is known about the role of the SNVs in correlation to PCOS and there are no *YAP1* variants listed in ClinVar as an PCOS associated variant. As *YAP1* is an established

gene of interest for PCOS, then it is possible for a missense variant within *YAP1* to possibly play a role in PCOS physiology or susceptibility. If a human *YAP1* VUS can be identified, determined to be conserved in *C. elegans*, then that same mutation can be engineered into *C. elegans* and impact gene and protein function. If the gene variant can be successfully engineered in *C. elegans*, then a visible phenotype may be observed to shed light on its potential impact for other organisms.

MATERIALS AND METHODS

All experiments detailed in this study were conducted in accordance with established standard operating protocols and procedures as set forth by the research laboratory of Dr. Ashley Turner.

Identification of *YAP1* VUS in Conserved Loci

YAP1 missense VUS previously identified in human patients were obtained from ClinVar (Landrum et al., 2018). Multiple sequence alignments (MSAs) were conducted utilizing Benchling (Benchling [Biology Software]. (2022)) to examine the level of conservation for each VUS locus across five species. Coding DNA (cDNA) sequence files were imported into Benchling from Ensembl (Cunningham et al., 2022). The following gene IDs were used to import each species' cDNA file: human (ESNT00000282441), zebrafish (ENSDART00000098914), mouse (ENSMUST00000065353), cow (ENSBTAT0000006112), and nematode (WBGene00008748). Each MSA was constructed through a consensus alignment utilizing the Clustal Omega alignment program (Sievers et al., 2011). Evolutionary conservation was assessed at the DNA and amino acid levels across species.

PolyPhen-2 Analysis

The Polyphen-2 bioinformatics tool was used to predict the potential impact of an amino acid change on the human YAP1 protein. The YAP1 protein identifier (P46937) was obtained from Uniprot (UniProt Consortium, 2023) and input into PolyPhen-2 along with the amino acid substitution position (339) and the specific amino acid substitution (asparagine to aspartic acid) (I. Adzhubei et al., 2013). The human YAP1 protein sequence has 504 amino acids.

Primer Design and Resuspension

Two DNA primer sets were designed using the design tool within Benchling (Benchling [Biology Software] (2022)) to amplify the VUS locus in the *yap-1* gene. The full genomic *yap-1 C. elegans* gene file was used for design. Primers were obtained from IDT as dry, lyophilized oligos and resuspended initially to create a primer stock at 100 μM in TE buffer. Next a working primer dilution was created at 10 μM with molecular water and used for downstream PCR experiments.

Nematode Manipulation and DNA Extraction

Nematodes were manipulated with a custom platinum wire nematode tool for strain maintenance and experiment transfers. Genomic DNA was obtained by placing 10-20 N2 *C. elegans* hermaphrodite adults into a PCR tube with 30 μL of worm lysis solution containing Proteinase K (Cat. # V302B, Promega) and incubated overnight at -80 °C followed by 65°C for 2 hours and 95°C for 10 minutes.

Polymerase Chain Reactions and Gel Electrophoresis

DNA from lysis solution was used directly as a template for polymerase chain reactions (PCRs).A 12.5 μL PCR reaction was generated using the following reagents: 0.25 μL of 10 μM forward primer, 0.25 μL of 10 μM reverse primer, 6.25 μL of Taq 2X Master Mix (Cat. # M0270S, New England Biology), 1 μL of template DNA, nuclease-free water up to 12.5 μL. The PCR reaction tubes were then placed into a thermocycler at an annealing temperature of 55°C. For gradient PCR, the as stated procedure was replicated an additional eight times for each of the primer sets. PCR tubes for primer set 1 was annealed at ranges 55-65°C, while PCR tubes for primer set 2 were annealed at ranges 50-60°C.

Polyacrylamide Gel Electrophoresis

PCR products were visualized on a 6% polyacrylamide gel utilizing electrophoresis. Gels were stained using 0.5μL of GelRed (Cat. #41003, Biotium) and imaged using a UV gel imaging system. A 100-base pair (bp) ladder was used to estimate product size (608bp and 790bp). 2 μL of PCR product was mixed with loading dye was loaded into each gel well and allowed to resolve on the gel for 45 minutes at 150 V.

***In vitro* Nuclease Assay**

Reagents were assembled in an RNAse free station. The ribonucleoprotein (RNP) complexes were assembled using 10 μL of single guide RNA, 1.6 μL of Cas9 enzyme and 88.4 μL of PBS. The RNP complex solution was then allowed to incubate at room temperature for 10 minutes to allow the RNP complex to fully form. RNP complexes were formed for each sgRNA separately, including *yap-1* and *rol-6*. sgRNA's were obtained from IDT. Two sets of RNP complexes were formed utilizing Alt-R HiFi Cas9 (Cat. #1081060, IDT) and QB3Macrolabs HiFi2 Cas9 (Cat #HF2CAS9-200, MCLAB). After formation of the RNP complexes, the *in vitro* nuclease digestion was performed. The follow reagents were combined in an aliquot tube: 1 μL of 10x Cas9 nuclease reaction buffer (Cat#60617W, MCLAB), 1 μL of Cas9 RNP complex, 1 μL of targeted PCR product, and 7 μL of nuclease-free water. The reaction was then incubated at 37 °C for 60 minutes. Next, 2 μL of proteinase K was added and incubated at 56 °C for 10 minutes. The digestion was then visualized on a 6% polyacrylamide gel utilizing electrophoresis.

RESULTS

A multiple sequence alignment was conducted to screen for variant loci conservation primarily between *C. elegans* and human, but also to gauge evolutionary conservation of the VUS region across a wide range of species. As seen in Tables 1 and 2, the MSA revealed that not only was the VUS locus conserved between human and *C. elegans*, but also cow and mouse.

Table 1
Evolutionary conservation of YAP1 VUS locus across five species at DNA level

Multiple Sequence Alignment																			
Nematode	T	T	A	C	-	G	A	C	A	A	T	G	G	A	C	A	C	T	C
Human	G	C	A	C	A	G	C	A	A	A	T	T	C	T	C	C	A	A	A
Cow	G	C	A	C	A	G	C	C	A	A	T	T	C	T	C	C	A	A	A
Mouse	G	C	A	C	A	G	C	A	A	A	T	G	C	T	C	C	A	A	A
Zebrafish	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	A	G

Table 1
Evolutionary conservation of YAP1 VUS locus across five species at amino acid level

Multiple Sequence Alignment																			
Nematode	Q	N	R	Y	N	N	C	Y	D	N	G	H	S	S	R	S	L	P	S
Human	M	R	N	I	N	P	S	T	A	N	S	P	L	C	Q	E	L	A	L
Cow	M	R	N	I	N	P	S	T	A	N	S	P	K	C	Q	E	L	A	L
Mouse	I	R	N	I	N	P	S	T	A	N	S	P	K	C	Q	E	L	A	L
Zebrafish	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	R	N

The VUS region was not conserved within zebrafish. This high conservation between species points toward a possibility of evolutionary importance of this gene region. Two variants were screened using MSA, with variants listed in ClinVar as having unknown clinical significance. One VUS was identified as being conserved between *C. elegans* and human, *YAP1* c.1015A>G (p.Asn339Asp). Table 3 shows the location of the VUS in *C. elegans* and human genomes.

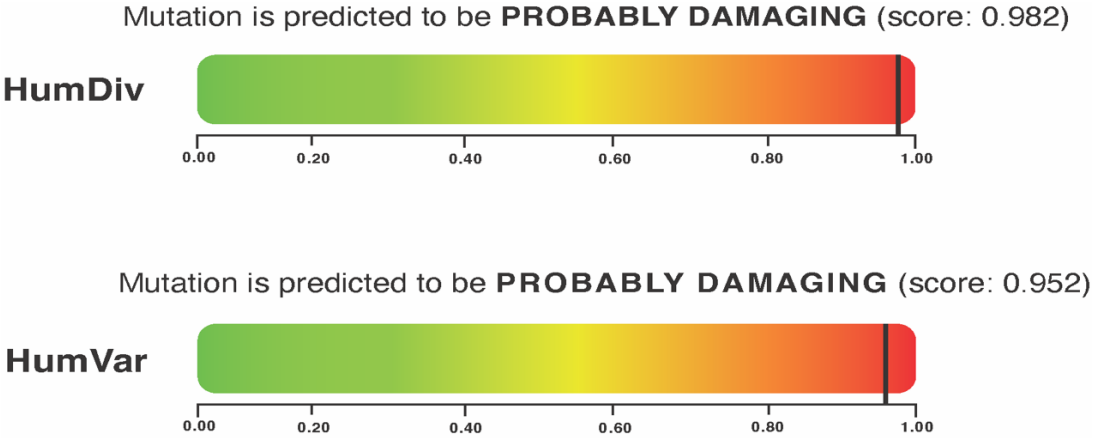
Table 3
Genetic Location of YAP1 VUS within H. sapiens and C. elegans orthologs

Species	VUS Location	Exon
C. elegans	*yap-1* c.1527A>G (p.Asn261Asp)	7
H. sapiens	*YAP1* c.1015A>G (p.Asn339Asp)	6

The potential impact of variant N339D on the structure and function of the encoded human YAP1 protein was estimated using the genetic algorithm within PolyPhen-2. The PolyPhen-2 HumDiv model predicted N339D to have a probably damaging score of 0.992, while the HumVar model predicted N339D to have a probably damaging score of 0.952. Both HumDiv and HumVar scores are shown in Figure 1.

Figure 1

Results of the PolyPhen-2 analysis predicting the potential impact of the YAP1 VUS c.1015A>G (p. Asn339Asp) substitution on the human YAP1 protein



PolyPhen-2 results displaying both HumDiv and HumVar score. HumDiv had a score of 0.982 and the HumVar had a score of 0.952.

Two DNA primer sets were then designed within Benchling to provide amplification of the target region within *yap-1* in *C. elegans*. Each primer set was created with a forward and reverse primer bordering the target VUS region. Primer set 1 had an expected amplification size of 608 base pairs and primer set 2 had an expected amplification size of 709 base pairs. Primer sequence information for both primer sets is provided in Table 4.

Table 4

DNA primer design for targeting the YAP1 VUS locus within yap-1

Primer set	VUS Locus	Forward Primer	Reverse Primer	Expected Product Size (bp)
1	*yap-1* c.1527A>G (p.Asn261Asp)	5' TGTTGAGCAGTTACCCATGCCCC 3'	5' TCCAGAATTGATCAGCTGCGCG 3'	608
2	*yap-1* c.1527A>G (p.Asn261Asp)	5' AGCAGTTACCCATGCCCCAAGG 3'	5' TGCGGTGTCTGTTGTTGTTGAGA 3'	790

Initial PCRs were carried out with the annealing temperature of 55°C to assess the specificity of each primer set. PCR products were then run out onto gels. PCR bands were present for both primer sets, showing successful amplification of the *C. elegans* DNA. Primer set 1 amplified a PCR product band around 900 bp with non-specific amplification present with additional bands. Primer set 2 amplified a prominent PCR product band around 800 bp with a single non-specific band. Both primer sets are shown in Figure 2. Lanes 1-2 and 2-2 were amplified using control *C. elegans* DN, ensuring that the DNA being amplified is from *C. elegans*. There was no amplification an either of the two negative control water lanes. This signifies that were was likely no contamination while preparing the PCR products as well as signifying that the amplification seen was coming from DNA.

Figure 2

Initial PCR results of the two designed DNA primer sets amplifying the YAP1 VUS locus in yap-1 through gel electrophoresis

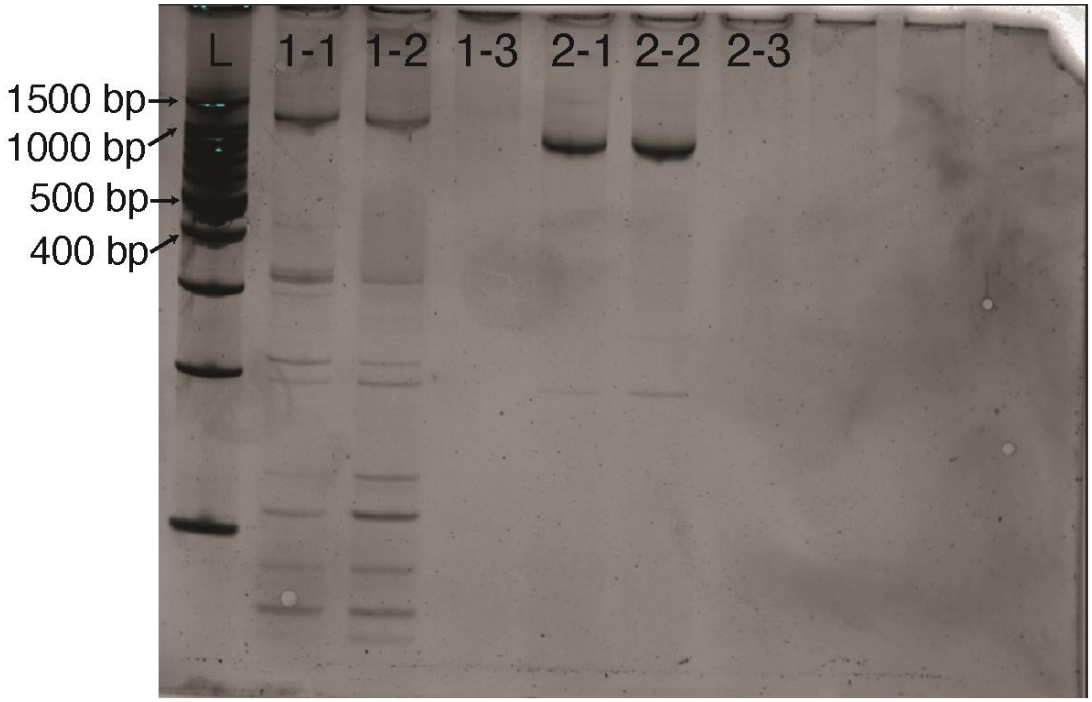
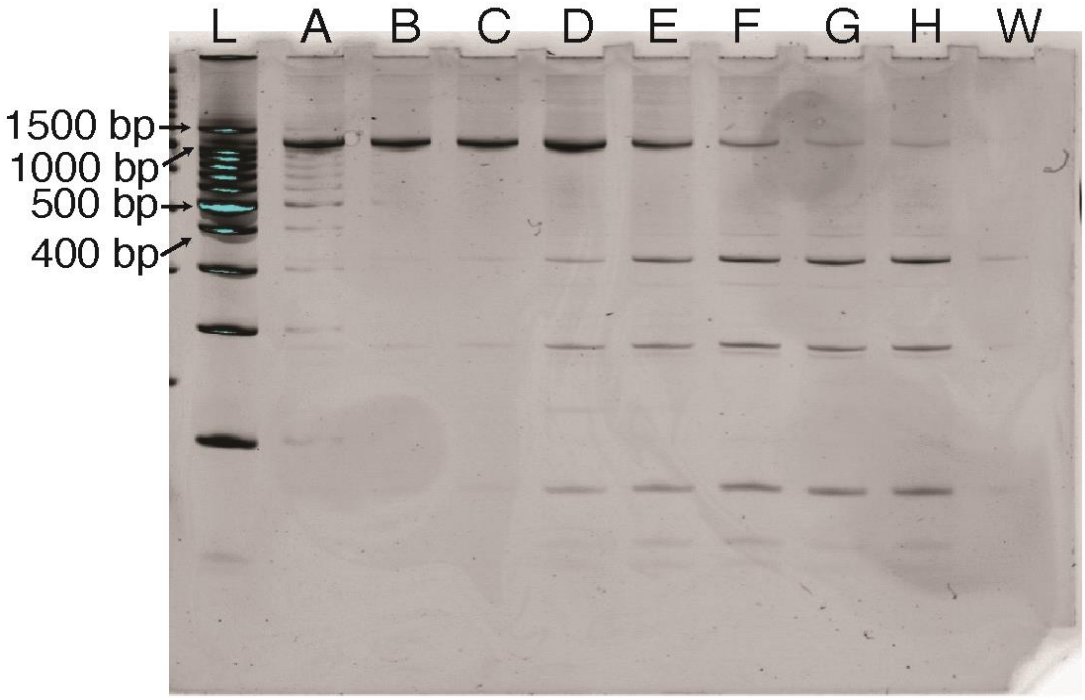


Image of polyacrylamide gel showing initial PCR amplification for each DNA primer set using 55 ºC annealing temperature. Gel is stained with GelRed and imaged with UV and gel documentation system. L = 100 bp ladder, primer set 1 PCR reactions (1-1 = digested N2 nematode DNA, 1-2 = positive control N2 nematode DNA, 1-3 = water, no template control), primer set 2 PCR reactions (2-1 = digested N2 nematode DNA, 2-2 = positive control N2 nematode DNA, 2-3 = water, no template control).

Gradient PCR was then used to optimize the annealing temperature to reduce non-specific amplification for both primer sets. Primer set 1 had a dark, single band at around 700 bp at 59.3°C. Primer set 2 had a dark single band at around 800 bp at 64.3°C. Primer set 1 is shown in Figure 3 and primer set 2 is shown in Figure 4.

Figure 3

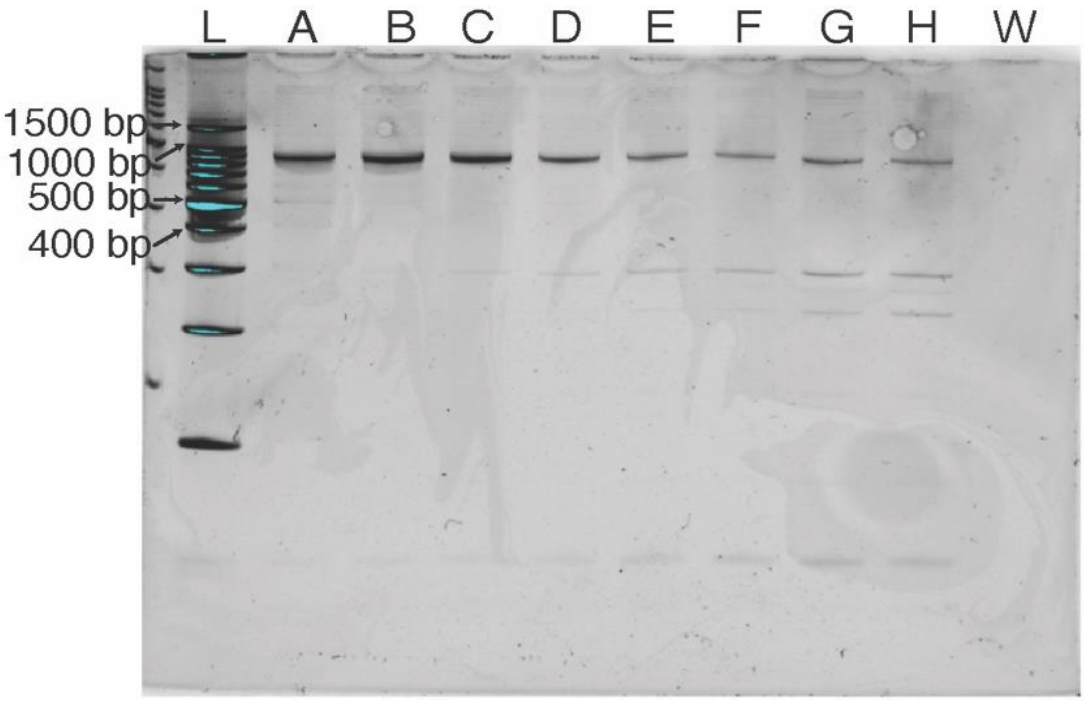
PCR results and annealing temperature optimization of primer set 1 amplification of the YAP1 VUS locus in yap-1 through gel electrophoresis



Images of polyacrylamide gels showing PCR amplification of primer sets one and two utilizing thermal gradients. N2 *C. elegans* DNA was used across all reactions. Annealing temperatures for primer set 1 ranging from 65-55°C. Thermal gradient includes: A = 65°C, B= 64.3°C, C = 63.1°C, D = 61.3°C, E = 59°C, F = 57.3°C, G = 56°C, H = 55°C, W = water.

Figure 4

PCR results and annealing temperature optimization of primer set 2 amplifications of the YAP1 VUS locus in yap-1 through gel electrophoresis



Images of polyacrylamide gels showing PCR amplification of primer sets one and two utilizing thermal gradients. N2 *C. elegans* DNA was used across all reactions. Annealing temperatures for primer set 2 ranging from 60-50°C. Thermal gradient includes: A = 60°C, B = 59.3°C, C = 58.1°C, D = 56.3°C, E = 54°C, F = 52.3 °C, G = 50.9°C, H = 50°C, W = water.

Following successful amplification, an RNA guide was then designed to target the *yap-1* N261D region in *C. elegans*. The guides were designed for use with Cas9 protein for CRISPR-Cas9 targeting and editing. The generated RNA guides targeted the VUS within 20 base pairs shown in Figure 5. RNA guide design is shown in Table 5 and Figure 6. The homology repair template for introducing the VUS into the nematode was generated within Benchling and is displayed in Table 6.

Table 5

CRISPR RNA guide design for yap-1 targeting

RNA Guide Target	Strand	RNA Guide Sequence	PAM
***yap-1* c.1527A>G (p.Asn261Asp)**	Top	TATAATAATTGTTACGACAA	TGG

Table 6

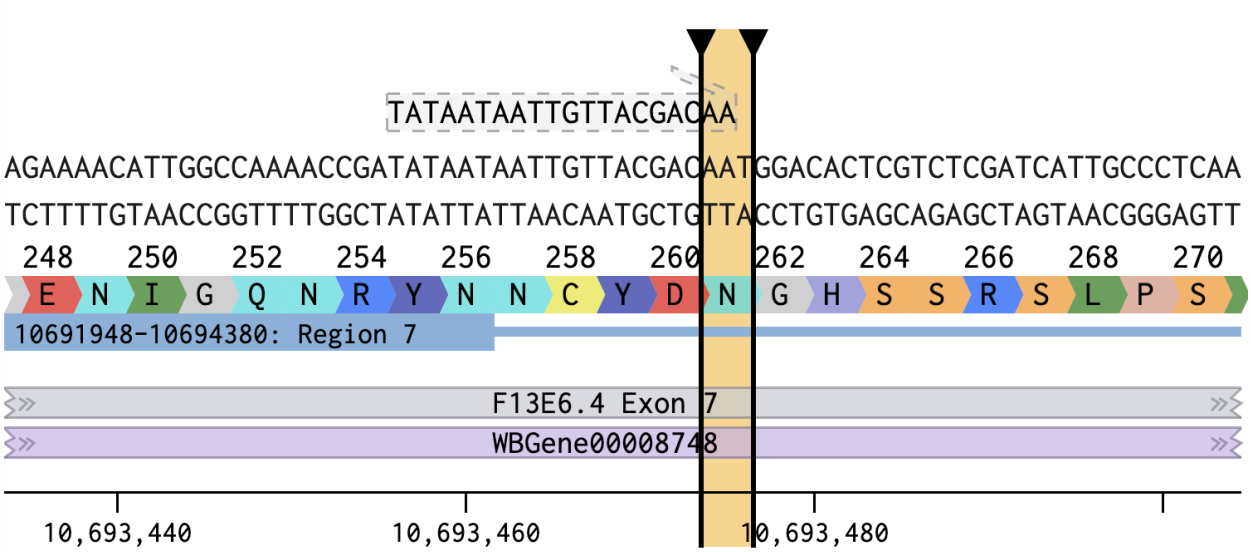
Homology repair template design for YAP1 VUS locus editing within yap-1

***yap-1* wildtype sequence**	CACATTAAATATTAATTTATTTCAACTTTCAGAACAAGAACAAACTG GTTTTGGATTAGGAGAAAACATTGGCCAAAACCGATATAATAATTGT TACGACAATGGACACTCGTCTCGATCATTGCCCTCAATTCATCAGCAC CAGCAGATGATTCCCAATCATCCACAACCACAATACTCTTCTCAGCA GCAAATGGATT
***yap-1* homology repair template sequence**	CACATTAAATATTAATTTATTTCAACTTTCAGAACAAGAACAAACTG GTTTTGGATTAGGAGAAAACATTGGCCAAAACCGATATAATAATTGC TATGACGATGGACACTCGTCTCGATCATTGCCCTCAATTCATCAGCAC CAGCAGATGATTCCCAATCATCCACAACCACAATACTCTTCTCAGCA GCAAATGGATT

Note. Sequence of the 200 bp homology repair template as well as 200 bp of the wildtype sequence. Red coloring signifies *yap-1* c.1527A>G (p.Asn261Asp) change. Introduction of the VUS into DNA region creates a Hyp99I restriction enzyme cut site that will be used for downstream screening and genotyping of mutant VUS animals. The restriction enzyme is signified by orange letters.

Figure 5

CRISPR-Cas9 targeting in the yap-1 gene for generation of the yap-1 VUS C. elegans model

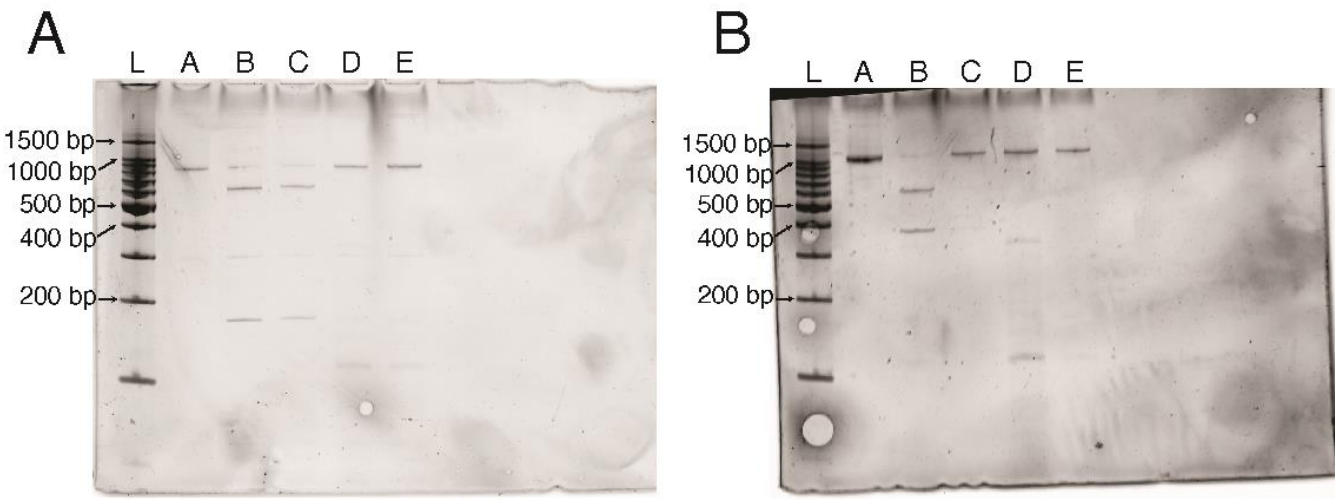


Single guide RNA design within Benchling. Gray-colored arrow signifies the sequence of the single guide RNA. Highlighted section displays VUS region in *C. elegans*.

Figures 6 and 7 displays the *in vitro* nuclease assay for *yap-1* DNA primer sets 1 and 2. The *in vitro* nuclease digests displayed cutting in the RNA guide targeted yap-1 region. The IVA was used to test guide efficiency as well as test cutting efficiency of Alt-R Cas9 enzyme versus wildtype Cas9 enzyme. The digests showed no Cas9 activity with *rol-6* RNP complexes or N2 PCR product. Cas9 was estimated to cut at 419bp and 144bp for primer set 1 and 418bp and 361bp for primer set 2.

Figure 6

In-vitro nuclease assay results displaying cutting of Cas9 in targeted regions



In vitro nuclease digestion reactions for *yap-1* primer set 1 and 2. Part A shows reactions with PCR products from *yap-1* primer set 1. Part B shows reactions with PCR products from *yap-1* primer set 2. L = ladder, A = uncut N2 *yap-1* PCR product, B = N2 *yap-1* PCR product + Alt-R Cas9 protein + *yap1* RNA guide (RNP complexes), C = N2 *yap-1* PCR product + QB3MacroLab Cas9 protein + *yap-1* RNA guide (RNP complexes), D = N2 *yap-1* PCR product + Alt-R Cas9 protein + *rol-6* RNA guide (RNP complexes), E = N2 *yap-1* PCR product + QB3MacroLab Cas9 protein + *rol-6* RNA guide (RNP complexes).

CONCLUSIONS

Multiple sequence alignment revealed that the variant of uncertain significance N339D locus to be conserved between *C. elegans*, human, cow, and mice genomes. The gene files were aligned using the Clustal Omega algorithm as it allows for highly accurate alignment and prevents the reversing of sequences (Sievers et al., 2011). The MSA revealed the VUS to be conserved across multiple species and the locus within ten base pairs on either side. Due to the high evolutionary conservation of this region, it can be assumed that the region of the VUS is possibly biological important. This high evolutionary conservation is expected, as the Hippo pathway itself highly conserved across species. However, *YAP1* along with other key Hippo genes were already present in unicellular organisms (Sebé-Pedrós et al., 2012). This reinforces the notion that *YAP1* is extremely important for biological function.

When looking at the protein region, the *YAP1* VUS c.1015A>G (p.Asn339Asp) lies in a structurally important position, a loop. Loops are vital regions of a protein, where they play a role protein function, stability, the protein's ability to properly fold and maintain that precise fold (Gavrilov et al., 2015). Changes in a loop region could possibly destabilize the loop, possibly resulting in altered protein function. N339D is a change from asparagine to aspartic acid. Asparagine and aspartic acid share many similarities, as asparagine is the amide derivative of aspartic acid. It is however important to note that asparagine holds a neutral charge, whereas aspartic acid carries a negative charge. This distinction is important in relation to protein stability. The change from a neutral to negative charge will result in a change in electrostatic interactions. Changes in electrostatics could have devastating results on the protein(Sinha & Smith-Gill, 2002).

Following confirmation of evolutionary conservation, PolyPhen-2 was utilized to predict the possible functional effect of the VUS on the human protein. Both HumVar and HumDiv scores were firmly within the range of probably damaging, HumDiv is the metric of primary interest. HumDiv is more important as it is the primary estimator of pathogenicity for non-Mendelian diseases (I. A. Adzhubei et al., 2010). This distinction is important as PCOS is a complex disease possible involving many genes and genotypes.

Following positive *in silico* experimentation, planning of *in vivo* experimentation with *YAP1* VUS c.1015A>G (p. Asn339Asp) began. As previously stated, *C. elegans* was the model organism chosen for *in vivo* experimentation. A side from previously stated advantages of *C. elegan*s. *C. elegans* is a power house for genetic research and already serves as a model for a wide variety of human disease, simulating the disease-related phenotypes almost exactly (Markaki & Tavernarakis, 2020). It is therefore safe to assume that if a phenotype is observed from the VUS in a mutant nematode compared to wildtype, then it is possible to predict the same VUS is impacting gene function in a human.

It was decided to utilize the CRISPR-Cas9 system, one of the most effective ways to engineer a VUS into a *C. elegans* organism to generate a mutant strain. This was decided based on the ease of utilizing CRISPR-Cas9 in *C. elegans* for genome editing as well as screening for the edited organisms. To accomplish this, a PCR-based screening assay was designed and utilized for downstream screening of potentially edited organisms. Two primer sets were designed and tested for reactivity and specificity at 55 °C. Each primer set amplified and targeted the appropriate base pair region for the VUS. Two lanes containing water instead of DNA were used to confirm no contamination during loading as well as to confirm DNA was being

amplified. There was no amplification in the water lanes. The initial temperature of 55 °C resulted in non-specific amplification for both primer sets. This was optimized by utilizing gradient PCR to find a more appropriate annealing temperature for specific amplification. Primer set 1 was tested from 55-65 °C, with primer set 2 being test from 50°- 60C. It was determined that primer set 1 has an optimum annealing temperature of 58.1 °C, while primer set 2 had an optimum annealing temperature of 63.1 °C. Determining optimum annealing temperatures is vital for the PCR-based screening assay due to the use of a wildtype N2 *C. elegans* DNA ran out on a gel with a single amplification band acting as an uncut control.

Following primer optimization, RNA guides were designed and tested for efficiency. An *in vitro* nuclease assay was used to test for RNA guide efficiency. As previously stated, the RNP complexes were assembled and incubated, PCR amplicons added, and heated for protein denaturation. The *in vitro* assay allows for confirmation of RNA guide accuracy as well as Cas9 cutting efficiency prior to use in *in vivo* experiments downstream. The experiment resulted in Cas9 cuts in lanes B and C for both primer sets. The gene *rol-6* was used as a negative control to show that the designed *yap-1* guide was targeting the guides designated gene sequence within *yap-1*. *rol-6* was chosen as it could be used in a downstream screening assay as a visual indicator that Cas9 has been active within the nematode. *rol-6* was the optimum gene to choose for this as *rol-6* causes the nematode to consistently roll in a circle. This experiment also examined the cutting efficiency of Alt-R Cas9 and QB3Macrolab Cas9. The digest reactions showed that Alt-R Cas9 is highly effective, nearly completely digesting the target band. Alt-R Cas9 also shows visibly stronger cutting bands for primer set 2. Based on the fading of the target band as well as

visibly stronger cut bands, it was concluded that the Alt-R Cas9 has a higher cutting efficiency for downstream *in vivo* targeting.

Polycystic ovary syndrome is an unbelievably complex disease that is understudied and underfunded (Brakta et al., 2017). As previously stated, PCOS has no known definitive cause, this makes any advancement in understanding of the genetic underlying of PCOS vital to advancing treatment toward treating PCOS itself instead of focusing on managing symptoms caused by PCOS. Personalized medicine is already a concept being applied to several prominent diseases (Goetz & Schork, 2018). It is likely that personalized medicine will become the prominent method of PCOS treatment. Understanding the genetic underlying of PCOS' pathology and which SNVs correlate to PCOS will greatly benefit the development of personalized medication for PCOS.

BIBLIOGRAPHY

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, *7*(4), 248–249. https://doi.org/10.1038/nmeth0410-248

Adzhubei, I., Jordan, D. M., & Sunyaev, S. R. (2013). Predicting functional effect of human missense mutations using PolyPhen-2. *Current Protocols in Human Genetics*, *Chapter 7*, Unit7.20. https://doi.org/10.1002/0471142905.hg0720s76

Baldridge, D., Wangler, M. F., Bowman, A. N., Yamamoto, S., Acosta, M. T., Adam, M., Adams, D. R., Agrawal, P. B., Alejandro, M. E., Alvey, J., Amendola, L., Andrews, A., Ashley, E. A., Azamian, M. S., Bacino, C. A., Bademci, G., Baker, E., Balasubramanyam, A., Baldridge, D., … Undiagnosed Diseases Network. (2021). Model organisms contribute to diagnosis and discovery in the undiagnosed diseases network: Current state and a future vision. *Orphanet Journal of Rare Diseases*, *16*(1), 206. https://doi.org/10.1186/s13023-021-01839-9

Bamshad, M. J., Nickerson, D. A., & Chong, J. X. (2019). Mendelian Gene Discovery: Fast and Furious with No End in Sight. *American Journal of Human Genetics*, *105*(3), 448–455. https://doi.org/10.1016/j.ajhg.2019.07.011

Brakta, S., Lizneva, D., Mykhalchenko, K., Imam, A., Walker, W., Diamond, M. P., & Azziz, R. (2017). Perspectives on Polycystic Ovary Syndrome: Is Polycystic Ovary Syndrome Research Underfunded? *The Journal of Clinical Endocrinology and Metabolism*, *102*(12), 4421–4427. https://doi.org/10.1210/jc.2017-01415

Condorelli, R. A., Calogero, A. E., Di Mauro, M., & La Vignera, S. (2017). PCOS and diabetes mellitus: From insulin resistance to altered beta pancreatic function, a link in evolution. *Gynecological Endocrinology*, *33*(9), 665–667. https://doi.org/10.1080/09513590.2017.1342240

Cunningham, F., Allen, J. E., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Austine-

Orimoloye, O., Azov, A. G., Barnes, I., Bennett, R., Berry, A., Bhai, J., Bignell, A., Billis, K., Boddu, S., Brooks, L., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., … Flicek, P. (2022). Ensembl 2022. *Nucleic Acids Research*, *50*(D1), D988–D995. https://doi.org/10.1093/nar/gkab1049

Day, F., Karaderi, T., Jones, M. R., Meun, C., He, C., Drong, A., Kraft, P., Lin, N., Huang, H., Broer, L., Magi, R., Saxena, R., Laisk, T., Urbanek, M., Hayes, M. G., Thorleifsson, G., Fernandez-Tajes, J., Mahajan, A., Mullin, B. H., … Welt, C. K. (2018). Large-scale genome-wide meta-analysis of polycystic ovary syndrome suggests shared genetic architecture for different diagnosis criteria. *PLOS Genetics*, *14*(12), e1007813. https://doi.org/10.1371/journal.pgen.1007813

Diamanti-Kandarakis, E., & Dunaif, A. (2012). Insulin Resistance and the Polycystic Ovary Syndrome Revisited: An Update on Mechanisms and Implications. *Endocrine Reviews*, *33*(6), 981–1030. https://doi.org/10.1210/er.2011-1034

Federici, G., & Soddu, S. (2020). Variants of uncertain significance in the era of high-throughput genome sequencing: A lesson from breast and ovary cancers. *Journal of Experimental & Clinical Cancer Research*, *39*(1), 46. https://doi.org/10.1186/s13046-020-01554-6

Gavrilov, Y., Dagan, S., & Levy, Y. (2015). Shortening a loop can increase protein native state entropy. *Proteins*, *83*(12), 2137–2146. https://doi.org/10.1002/prot.24926

Gibson-Helm, M., Teede, H., Dunaif, A., & Dokras, A. (2016). Delayed Diagnosis and a Lack of Information Associated With Dissatisfaction in Women With Polycystic Ovary Syndrome. *The Journal of Clinical Endocrinology and Metabolism*, *102*(2), 604–612. https://doi.org/10.1210/jc.2016-2963

Goetz, L. H., & Schork, N. J. (2018). Personalized medicine: Motivation, challenges, and progress. *Fertility and Sterility*, *109*(6), 952–963. https://doi.org/10.1016/j.fertnstert.2018.05.006

Hall, C. A., Wang, R., Miao, J., Oliva, E., Shen, X., Wheeler, T., Hilsenbeck, S. G., Orsulic, S., & Goode,

S. (2010). Hippo Pathway Effector Yap is an Ovarian Cancer Oncogene. *Cancer Research*, *70*(21), 8517–8525. https://doi.org/10.1158/0008-5472.CAN-10-1242

Hashmi, S., Wang, Y., Parhar, R. S., Collison, K. S., Conca, W., Al-Mohanna, F., & Gaugler, R. (2013). A C. elegans model to study human metabolic regulation. *Nutrition & Metabolism*, *10*(1), 31. https://doi.org/10.1186/1743-7075-10-31

Ismayilova, M., & Yaya, S. (2022). "I felt like she didn't take me seriously": A multi-methods study examining patient satisfaction and experiences with polycystic ovary syndrome (PCOS) in Canada. *BMC Women's Health*, *22*(1), 47. https://doi.org/10.1186/s12905-022-01630-3

Jiang, L.-L., Xie, J.-K., Cui, J.-Q., Wei, D., Yin, B.-L., Zhang, Y.-N., Chen, Y.-H., Han, X., Wang, Q., & Zhang, C.-L. (2017). Promoter methylation of yes-associated protein (YAP1) gene in polycystic ovary syndrome. *Medicine*, *96*(2), e5768. https://doi.org/10.1097/MD.0000000000005768

Landrum, M. J., Lee, J. M., Benson, M., Brown, G. R., Chao, C., Chitipiralla, S., Gu, B., Hart, J., Hoffman, D., Jang, W., Karapetyan, K., Katz, K., Liu, C., Maddipatla, Z., Malheiro, A., McDaniel, K., Ovetsky, M., Riley, G., Zhou, G., … Maglott, D. R. (2018). ClinVar: Improving access to variant interpretations and supporting evidence. *Nucleic Acids Research*, *46*(D1), D1062–D1067. https://doi.org/10.1093/nar/gkx1153

Li, T., Zhao, H., Zhao, X., Zhang, B., Cui, L., Shi, Y., Li, G., Wang, P., & Chen, Z.-J. (2012). Identification of YAP1 as a novel susceptibility gene for polycystic ovary syndrome. *Journal of Medical Genetics*, *49*(4), 254–257. https://doi.org/10.1136/jmedgenet-2011-100727

Lv, X., He, C., Huang, C., Wang, H., Hua, G., Wang, Z., Zhou, J., Chen, X., Ma, B., Timm, B. K., Maclin, V., Dong, J., Rueda, B. R., Davis, J. S., & Wang, C. (2019). Timely expression and activation of YAP1 in granulosa cells is essential for ovarian follicle development. *The FASEB Journal*, *33*(9), 10049–10064. https://doi.org/10.1096/fj.201900179RR

Markaki, M., & Tavernarakis, N. (2020). Caenorhabditis elegans as a model system for human diseases.

Current Opinion in Biotechnology, *63*, 118–125. https://doi.org/10.1016/j.copbio.2019.12.011

Murfin, K. E., Dillman, A. R., Foster, J. M., Bulgheresi, S., Slatko, B. E., Sternberg, P. W., & Goodrich-Blair, H. (2012). Nematode-Bacterium Symbioses—Cooperation and Conflict Revealed in the "Omics" Age. *The Biological Bulletin*, *223*(1), 85–102. https://doi.org/10.1086/BBLv223n1p85

Nigon, V. M., & Félix, M.-A. (2018). History of research on C. elegans and other free-living nematodes as model organisms. In *WormBook: The Online Review of C. elegans Biology [Internet]*. WormBook. https://www.ncbi.nlm.nih.gov/books/NBK453431/

Plouffe, S. W., Lin, K. C., Moore, J. L., Tan, F. E., Ma, S., Ye, Z., Qiu, Y., Ren, B., & Guan, K.-L. (2018). The Hippo pathway effector proteins YAP and TAZ have both distinct and overlapping functions in the cell. *The Journal of Biological Chemistry*, *293*(28), 11230–11240. https://doi.org/10.1074/jbc.RA118.002715

Polycystic ovary syndrome. (n.d.). Retrieved May 3, 2024, from https://www.who.int/news-room/fact-sheets/detail/polycystic-ovary-syndrome

Polycystic Ovary Syndrome (PCOS). (2022, February 28). https://www.hopkinsmedicine.org/health/conditions-and-diseases/polycystic-ovary-syndrome-pcos

Puurunen, J., Piltonen, T., Jaakkola, P., Ruokonen, A., Morin-Papunen, L., & Tapanainen, J. S. (2009). Adrenal Androgen Production Capacity Remains High up to Menopause in Women with Polycystic Ovary Syndrome. *The Journal of Clinical Endocrinology & Metabolism*, *94*(6), 1973–1978. https://doi.org/10.1210/jc.2008-2583

S, B., & A, S. (2017). The pathogenesis and treatment of polycystic ovary syndrome: What's new? *Advances in Clinical and Experimental Medicine : Official Organ Wroclaw Medical University*, *26*(2). https://doi.org/10.17219/acem/59380

Salcedo Allende, M. T., Zeron-Medina, J., Hernandez, J., Macarulla, T., Balsells, J., Merino, X., Allende,

H., Tabernero, J., & Ramon Y Cajal, S. (2017). Overexpression of Yes Associated Protein 1, an Independent Prognostic Marker in Patients With Pancreatic Ductal Adenocarcinoma, Correlated With Liver Metastasis and Poor Prognosis. *Pancreas*, *46*(7), 913–920. https://doi.org/10.1097/MPA.0000000000000867

Sebé-Pedrós, A., Zheng, Y., Ruiz-Trillo, I., & Pan, D. (2012). Premetazoan origin of the Hippo signaling pathway. *Cell Reports*, *1*(1), 13–20. https://doi.org/10.1016/j.celrep.2011.11.004

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, *7*, 539. https://doi.org/10.1038/msb.2011.75

Sinha, N., & Smith-Gill, S. J. (2002). Electrostatics in protein binding and function. *Current Protein & Peptide Science*, *3*(6), 601–614. https://doi.org/10.2174/1389203023380431

Stein, L., Sternberg, P., Durbin, R., Thierry-Mieg, J., & Spieth, J. (2001). WormBase: Network access to the genome and biology of Caenorhabditis elegans. *Nucleic Acids Research*, *29*(1), 82–86.

Sun, T., & Diaz, F. J. (2019). Ovulatory signals alter granulosa cell behavior through YAP1 signaling. *Reproductive Biology and Endocrinology : RB&E*, *17*, 113. https://doi.org/10.1186/s12958-019-0552-1

Szulzewsky, F., Holland, E. C., & Vasioukhin, V. (2021). YAP1 and its fusion proteins in cancer initiation, progression and therapeutic resistance. *Developmental Biology*, *475*, 205–221. https://doi.org/10.1016/j.ydbio.2020.12.018

THE C. ELEGANS SEQUENCING CONSORTIUM. (1998). Genome Sequence of the Nematode C. elegans: A Platform for Investigating Biology. *Science*, *282*(5396), 2012–2018. https://doi.org/10.1126/science.282.5396.2012

Tremblay-Davis, A. C., Holley, S. L., & Downes, L. A. (2021). Diagnosis and Treatment of Polycystic

Ovary Syndrome in Primary Care. *The Journal for Nurse Practitioners*, *17*(10), 1226–1229. https://doi.org/10.1016/j.nurpra.2021.08.008

UniProt Consortium. (2023). UniProt: The Universal Protein Knowledgebase in 2023. *Nucleic Acids Research*, *51*(D1), D523–D531. https://doi.org/10.1093/nar/gkac1052

Watson, E., & Walhout, A. J. M. (2014). C. elegans Metabolic Gene Regulatory Networks Govern the Cellular Economy. *Trends in Endocrinology and Metabolism: TEM*, *25*(10), 502–508. https://doi.org/10.1016/j.tem.2014.03.004

Yes-associated protein homolog, YAP-1, is involved in the thermotolerance and aging in the nematode Caenorhabditis elegans—ScienceDirect. (n.d.). Retrieved October 14, 2023, from https://www.sciencedirect.com/science/article/abs/pii/S0014482713000384?via%3Dihub

Yuan, M., Tomlinson, V., Lara, R., Holliday, D., Chelala, C., Harada, T., Gangeswaran, R., Manson-Bishop, C., Smith, P., Danovi, S. A., Pardo, O., Crook, T., Mein, C. A., Lemoine, N. R., Jones, L. J., & Basu, S. (2008). Yes-associated protein (YAP) functions as a tumor suppressor in breast. *Cell Death and Differentiation*, *15*(11), 1752–1759. https://doi.org/10.1038/cdd.2008.108

Zhao, B., Wei, X., Li, W., Udan, R. S., Yang, Q., Kim, J., Xie, J., Ikenoue, T., Yu, J., Li, L., Zheng, P., Ye, K., Chinnaiyan, A., Halder, G., Lai, Z.-C., & Guan, K.-L. (2007). Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact inhibition and tissue growth control. *Genes & Development*, *21*(21), 2747–2761. https://doi.org/10.1101/gad.1602907