Calpain10 (CAPN10): Unraveling its Significance in Polycystic Ovary Syndrome (PCOS) and Insulin Resistance

Meena Easwaran

McWilliams School of Biomedical Informatics

The University of Texas Health Science Center at Houston

BMI 5330 Introduction to Bioinformatics

Dr. Wenjin. J. Zheng

December 08, 2023

Introduction and Background

Polycystic Ovary Syndrome (PCOS) is a hormonal and reproductive endocrine disorder, that is estimated to affect every one in ten women worldwide [1]. Anovulation (irregular or no ovulation), menstrual irregularities, acne, obesity, polycystic ovaries, hirsutism, or excessive hair growth due to excess androgen production are all common PCOS symptoms [2]. There is a wide range of phenotypic and symptomatic manifestations of this disorder, and these vary from woman to woman [2]. In general, PCOS phenotypes can be classified into four categories as shown in **Table 1** [3–7]. Clinically, these phenotypes can be used to diagnose PCOS symptoms and manage them. A common feature of three out of four PCOS phenotypes is insulin resistance, as illustrated in **Table 1** [5].

Table 1. Classification of PCOS Phenotypes

PCOS Phenotype	**Characteristics**			**Other nomenclature**	**Incidence of Insulin resistance**
	Anovulation	**Hyperandrogenism**	**Polycystic Ovary**		
1/A	Yes	Yes	Yes	Complete or Full-blown PCOS	Yes
2/B	Yes	Yes	No	Non polycystic ovary PCOS	Yes
3/C	No	Yes	Yes	Ovulatory PCOS	No
4/D	Yes	No	Yes	Non-hyperandrogenic PCOS	Yes

Insulin resistance is a metabolic condition that occurs when the body cells become less sensitive to insulin effects [8]. To compensate for reduced insulin effectiveness, the pancreas produces more insulin, a condition called hyperinsulinemia. This can result in hyperglycemia or high blood sugar levels, increasing the risk of Type 2 Diabetes mellitus [6, 7]. Besides insulin

resistance, women with PCOS are also at a higher risk of developing other metabolic conditions, such as Non-alcoholic fatty liver disease (NAFLD), hypertension, high cholesterol, cardiovascular diseases, and ovarian and endometrial cancers (**Fig. 1**) [4, 9]. These metabolic complications associated with PCOS are closely related and often interact with one another. A holistic approach to managing these PCOS-related health complications may require changes in lifestyle, changes to diet, weight management, and specific medication administration [7]. Unfortunately, there is currently no cure for PCOS and its associated phenotypic symptoms. Managing and mitigating PCOS complications requires close collaboration with health care professionals. The exact etiology of this multifaceted and multiphenotypic disorder is still not fully understood till date. PCOS has traditionally been regarded as an inherited disorder [10]. There is increasing evidence to suggest that this disorder may have a genetic component and that its symptoms may be associated with it [11]. Environmental, occupational, and lifestyle factors are also believed to be contributing factors [12].

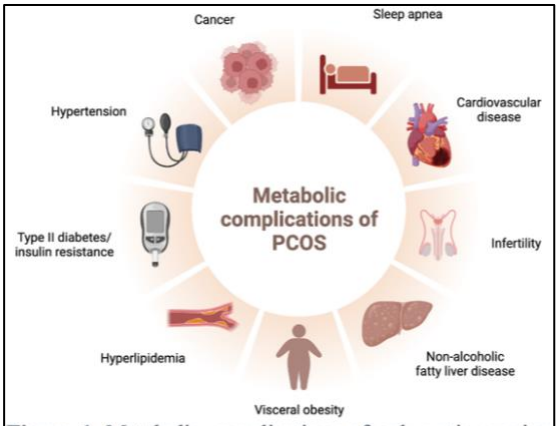

Figure 1. Metabolic complications of polycystic ovarian syndrome (PCOS). PCOS increases the likelihood that women will develop metabolic diseases, including Type II diabetes, insulin resistance, infertility, high blood pressure, hyperlipidemia, cardiovascular disease, obesity, non-alcoholic fatty liver disease (NAFLD), sleep apnea, endometrium, and ovarian cancers. Several of these conditions are interrelated, and women with PCOS may not show all these symptoms, and severity varies according to the individual. Created with BioRender.com.

My motivation to focus on PCOS for this project stems from a personal connection. I have been diagnosed with PCOS since my adolescence. In earlier generations, no women in my immediate family have been officially diagnosed with this condition, which is believed to be inherited. Considering this revelation, I am left to wonder whether PCOS remained undetected in these women due to the lack of advanced medical technologies at the time. Considering my situation, I also wonder if my potential PCOS diagnosis can be attributed to confounding factors

related to my lifestyle and environment. A bioinformatic approach may be able to help answer these questions. Embarking on this class project not only fulfills an academic requirement but also provides me with a personal journey of discovery and understanding. Taking a deep dive into the complexities of PCOS through bioinformatics will not only shed light on my own situation but will potentially contribute valuable insights for others with similar concerns. I am motivated to unravel the mysteries of PCOS through this project, which serves as a bridge between my academic pursuit and personal curiosity.

An important aspect of this motivational approach involves selecting a suitable gene for downstream analysis. Considering the insulin resistance aspect of PCOS, several abnormalities in genes like Calpain 10 (CAPN10), Peroxisome proliferator-activated receptor gamma (PPARγ) and genes associated with Insulin associated pathways such as Insulin (INS), Insulin-like growth factors (IGFs-IGF1 and 2), Insulin receptor (INSR), and Insulin receptor substrate protein (IRS1 and IRS2) have been previously identified [11, 13–15]. For this project, I have chosen to focus on CAPN10, as a candidate gene. There is a rationale behind this since it is one of the first genes to be associated with type 2 diabetes [16], which is closely associated to insulin resistance. It is also strikingly implicated as a genetic cause as of PCOS as a result of mutations [11, 13, 14]. Moreover, I learned about PCOS Knowledge Base (PCOSKB) [17, 18], a repository of curated information pertaining to PCOS through a literature search. A notable feature in this database is the Venn analysis tool, which allowed me to conduct a pilot comparative analysis of gene sets present between PCOS, insulin resistance syndrome, and Diabetes mellitus. The preliminary outcomes showed three genes that included INS and INSR, common to all three health conditions (**Fig. 2)**. CAPN10 gene was not among them. In the case of PCOS and insulin resistance, only one gene

was common, but this was not CAPN10 interestingly. CAPN10 turned out to be one in 96 common genes found between PCOS and Diabetes along with other genes like IGF1, IGF2, IRS1, and PPARγ (**Fig. 2)**. There is a considerable significance in this observation, particularly in light of established research emphasizing CAPN10's role in insulin resistance [19], a correlation which was later corroborated in pertinent PCOS-related literature [20, 21]. Furthermore, it is also possible that this database [17, 18] is not up-to-date as it has data curated from 2017 to 2019. To bridge these gaps in such PCOS databases and to possibly find novel insights, this project focused on CAPN10, as a candidate gene.

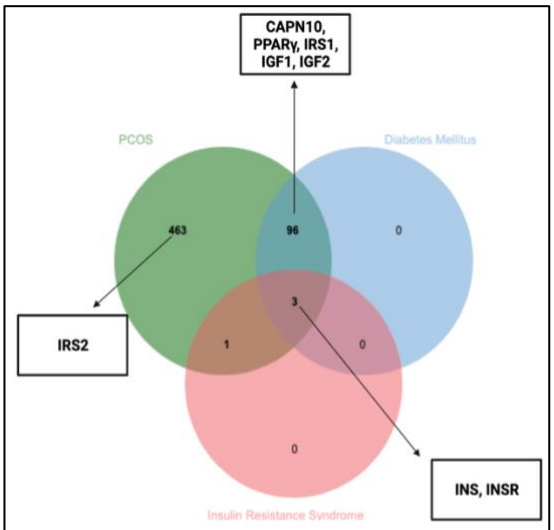

Figure 2. Genetic commonalties between PCOS, insulin resistance and diabetes mellitus. CAPN10, PPARγ, IRS1, IGF1, and IGF2 were common to PCOS and diabetes mellitus. INS and INSR were common to all three conditions. IRS1 was specific to only PCOS group. Created with PCOSKB and BioRender.com.

In this regard, the purpose of this project is to evaluate the role of CAPN10 gene in predisposing women with PCOS to insulin resistance and possibly Type 2 Diabetes. Specifically, this study will comprise of five distinct aims. First, it aims to investigate the expression levels of the CAPN10 gene throughout various human tissues. Second, the project seeks to assess the inter-species conservation status of CAPN10. Identifying genetic variations in CAPN10 and examining their associations with insulin resistance and Type 2 Diabetes will be the third and fourth aims of this project. Finally, a comprehensive analysis will be conducted to identify specific patterns in CAPN10 gene expression through differential expression and functional gene enrichment, using PCOS datasets from National Center for Biotechnology Information’s (NCBI) Gene Expression Omnibus (GEO).

Results

CAPN10 gene expression in human tissues

Gene expression patterns of CAPN10 in various human tissues were examined using three different tools such as Human Protein Atlas (HPA) [22], GTex portal [23], and EMBL-EBI Expression Atlas [24]. In terms of HPA outcomes, the number of transcripts for CAPN10 gene were identified as 6. CAPN10 expression pertained to being intracellular and to the cytosolic compartment as shown in **Fig. 3**. In terms of the tissue expression at gene and protein levels, CAPN10 was consistently expressed in organs across all functioning human systems (**Fig. 4**). In general, spleen seemed to have higher expression scores at gene and protein levels of measure.

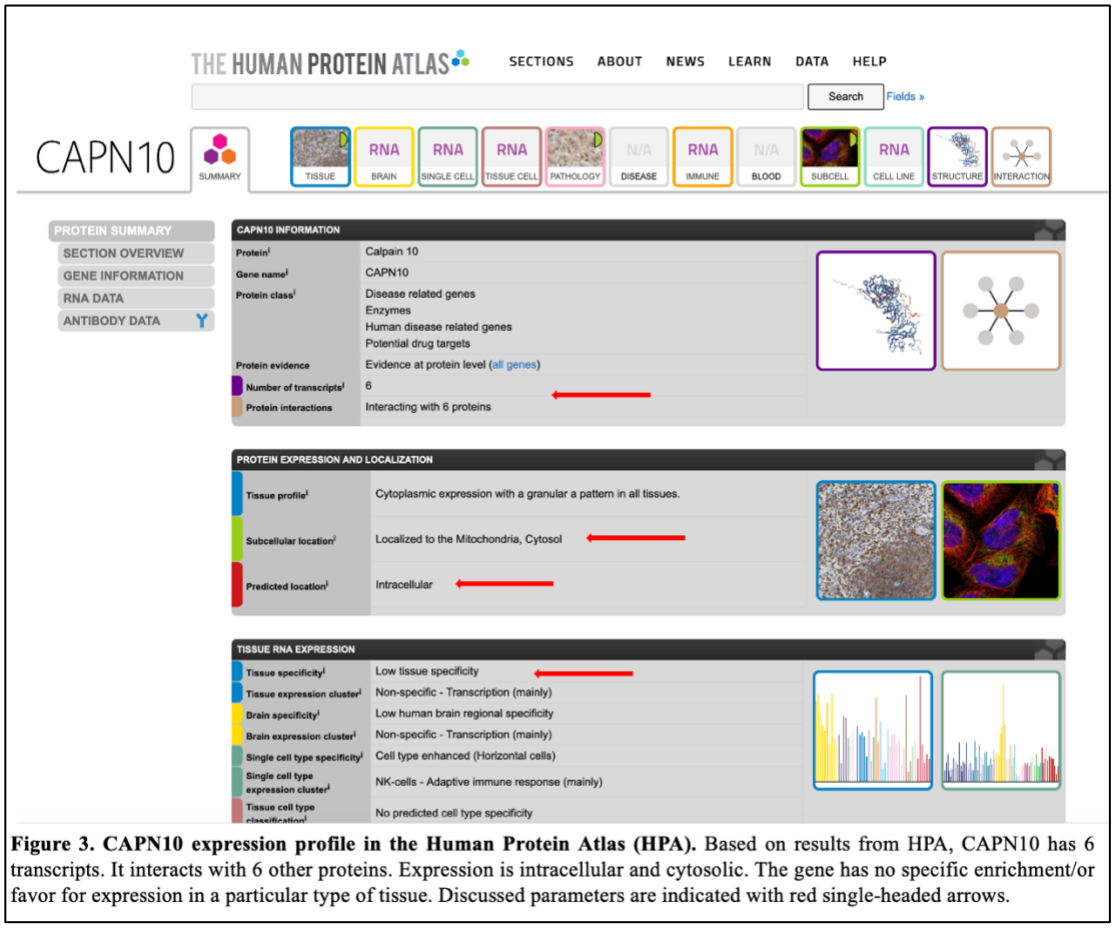

Figure 3. CAPN10 expression profile in the Human Protein Atlas (HPA). Based on results from HPA, CAPN10 has 6 transcripts. It interacts with 6 other proteins. Expression is intracellular and cytosolic. The gene has no specific enrichment/or favor for expression in a particular type of tissue. Discussed parameters are indicated with red single-headed arrows.

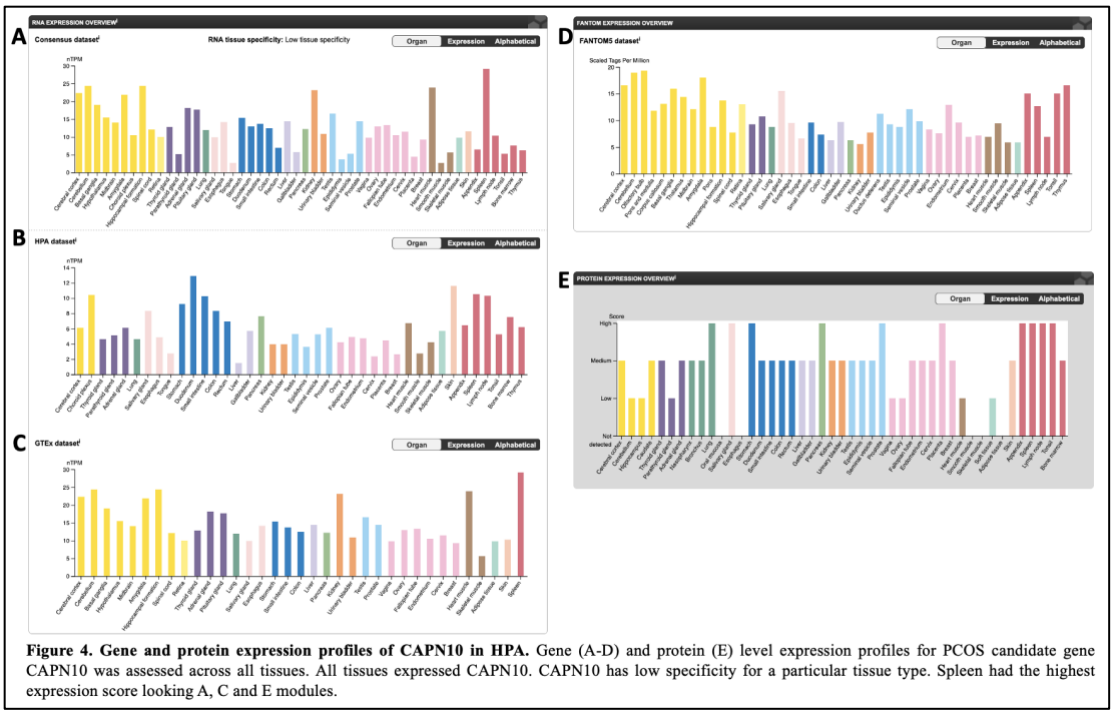

Figure 4. Gene and protein expression profiles of CAPN10 in HPA. Gene (A-D) and protein (E) level expression profiles for PCOS candidate gene CAPN10 was assessed across all tissues. All tissues expressed CAPN10. CAPN10 has low specificity for a particular tissue type. Spleen had the highest expression score looking A, C and E modules.

Although HPA has GTex portal expression levels (**Fig. 4C**), the CAPN10 expression levels were also directly assessed using the GTex portal [23]. The data (**Fig. 5**) show similar expression profiles of CAPN10 as previously observed via HPA. An interesting observation was that certain regions in the brain seem to have higher CAPN10 expression scores, in addition to spleen. Furthermore, outcomes from EMBL-EBI Expression Atlas [24] (**Fig. 6**) were quite similar to the earlier results. A wide variety of tissues seem to have low-high expression scores for CAPN10. Specifically, brain tissues have higher levels of expression in Expression Atlas outcomes.

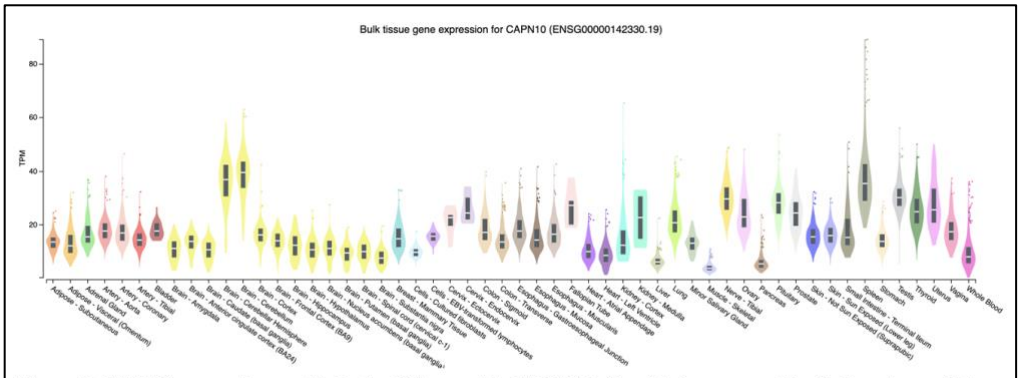

Figure 5. CAPN10 expression profile in the GTex portal. CAPN10 is found to be expressed in all tissue types. Spleen and specific regions of brain seem to have high expression.

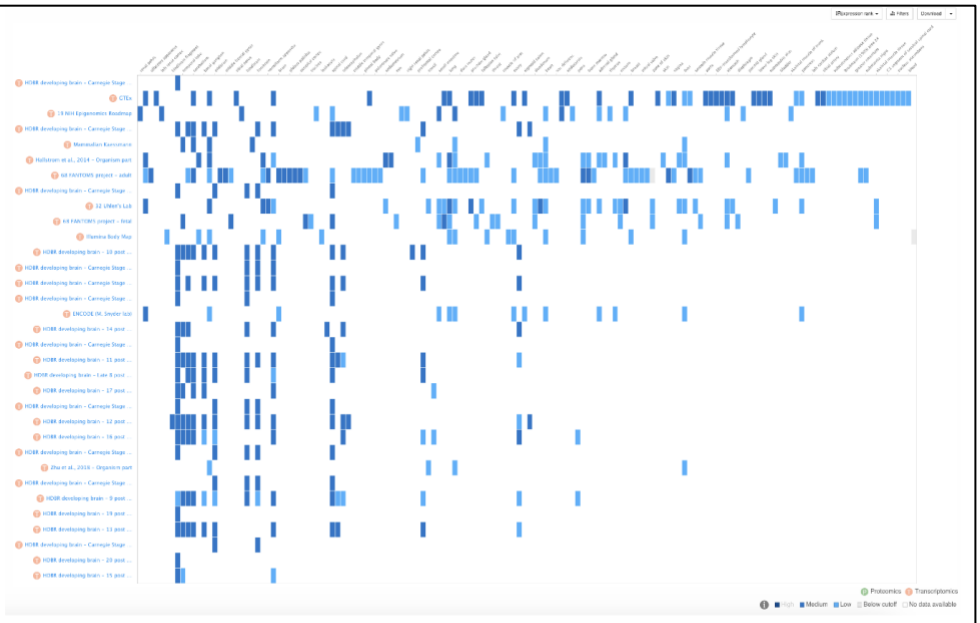

Figure 6. CAPN10 expression profile in EMBL-EBI expression atlas. CAPN10 is found to be expressed in all tissue types across different studies. The expression profiles vary from high to low, with brain regions having higher levels across studies.

Inter-species conservation of CAPN10

CAPN10 gene conservation across different species was analyzed using the University of California, Santa Cruz (UCSC) Genome browser [25] in reference to human reference assembly GRCh38/hg38. The results showed well conserved exon or coding regions for most of the laboratory animal models used in PCOS research such as rhesus monkey, mice, rats, cows, and dogs, except for pigs (**Fig. 7**). CAPN10 gene was found to contain 12 exonic and 11 intronic regions. 5' and 3'Untranslated regions (UTR) were found beside exons 1 and 12, respectively (**Fig. 7**).

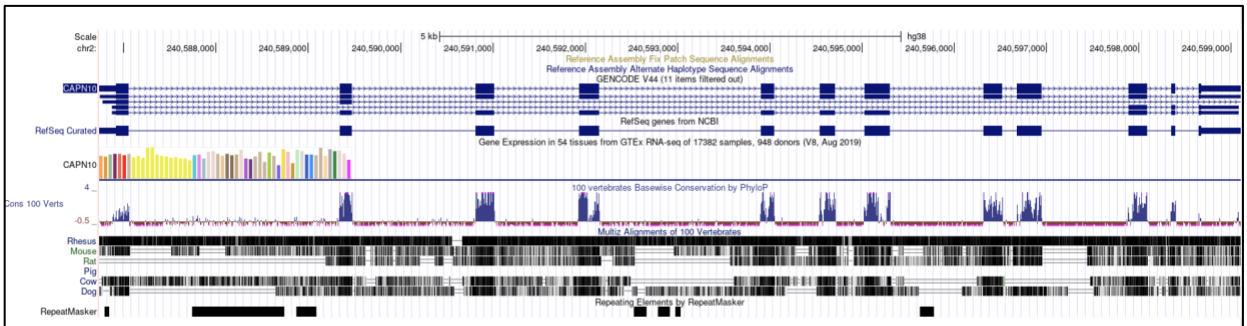

Figure 7. CAPN10 gene conservation across PCOS animal models in UCSC Genome browser. CAPN10 is found with 12 exon regions and 11 intron regions. Exonic regions are well conserved across most animal species typically used as PCOS research models.

CAPN10 gene conservation was also compared across entire species that are available on this browser. The results show that this gene has well conserved exonic regions when compared across all available species (**Fig. 8**). Moreover, it was interesting to notice well conserved intronic regions specifically in non-human primates.

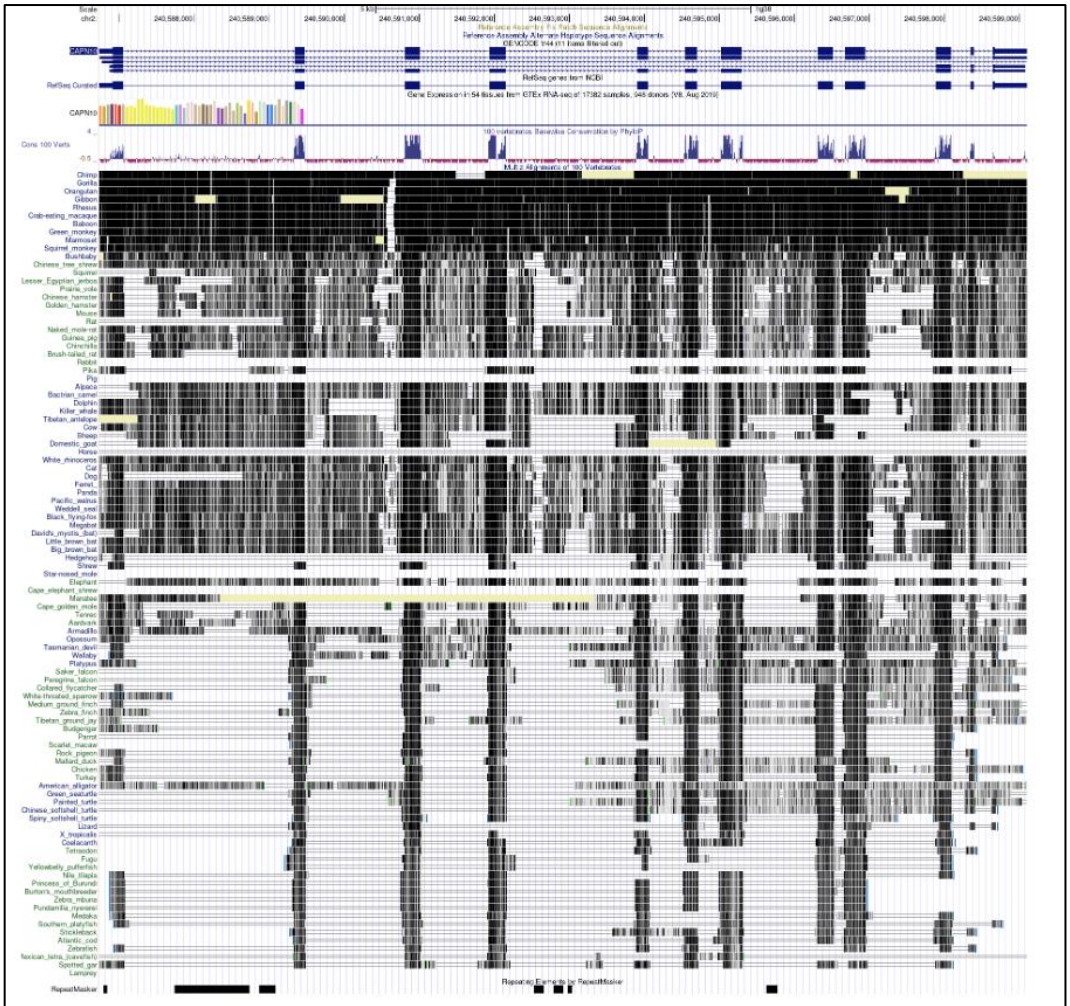

Figure 8. CAPN10 gene conservation across all species available in UCSC Genome browser. Exonic regions are well conserved across all species. Non-human primates also seem to have well conserved intronic regions.

CAPN10 genetic variations and its clinical relevance to PCOS and insulin resistance

Information about single nucleotide polymorphisms (SNPs) and structural variants, were obtained from NCBI resources such as dBSNP [26] and dbVAR [27] respectively, for the PCOS candidate gene, CAPN10 in humans. In terms of the SNPs, filtered results for clinical significance

in the “pathogenic and “risk-factor” categories from dBSNP [26], showed 2 single nucleotide variations (SNVs) and 3 deletion insertions (delins) for human CAPN10 gene as shown in **Table 2**. After screening for clinical condition relevance with ClinVaR[28], one delin (rs3842570) and two SNVs (rs3792267 and rs2975760) were found to be linked to Type 2 Diabetes mellitus. Moreover, snv rs2975760 had a clinical relevance to PCOS in addition to Type 2 Diabetes mellitus as well.

Table 2. Single nucleotide variations and deletion insertions on human CAPN10 gene

Variant type	**Snp_id**	**Clinical significance**	**ClinVar Accession ID**	**Clinvar Conditions**
delins	1559425580	pathogenic	VCV000984989.2	Inborn genetic conditions
delins	1182803771	pathogenic	data not found	data not found
delins	3842570	risk-factor	VCV000005094.4	Type 2 diabetes mellitus 1, susceptibility to
snv	3792267	risk-factor	VCV000005093.5	Type 2 diabetes mellitus 1, susceptibility to
snv	2975760	risk-factor	VCV000005096.5	Polycystic ovary syndrome, susceptibility to; Type 2 diabetes mellitus 1, susceptibility to

The SNP’s rs3842570, rs3792267 and rs2975760 were then cross verified using the SNP data found in PCOSKB [17, 18]. These three terms matched with the database, however a new fourth term rs5030952 was identified as an SNP associated with CAPN10 (**Fig. 9A**). Data found on ClinVaR[28] for this SNP ID corroborates with the data found in PCOSKB [17, 18] (**Fig. 9B**). Moreover, data indicates its possible relevance in Type 2 Diabetes as well. For all four identified

SNP IDs, there were no explicit clinical terminologies related to the terms “insulin resistance or insulin resistance syndrome”.

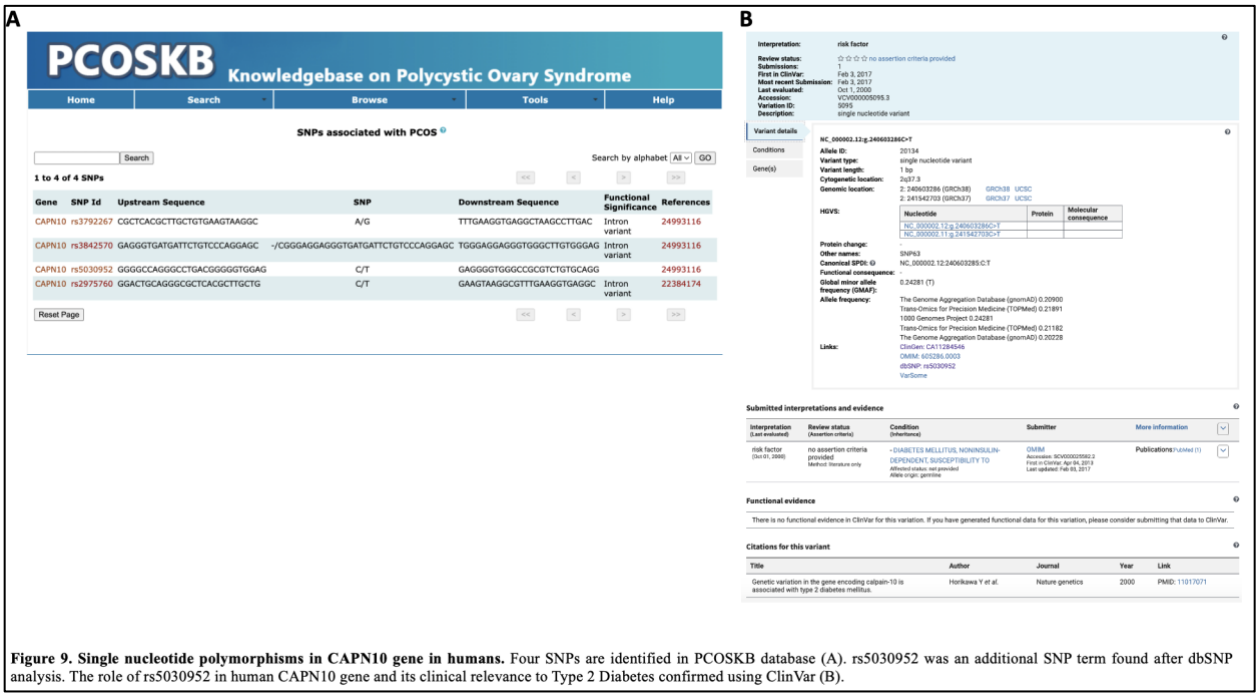

Figure 9. Single nucleotide polymorphisms in CAPN10 gene in humans. Four SNPs are identified in PCOSKB database (A). rs5030952 was an additional SNP term found after dbSNP analysis. The role of rs5030952 in human CAPN10 gene and its clinical relevance to Type 2 Diabetes confirmed using ClinVar (B).

In terms of the structural variants, filtered results for clinical significance in the “pathogenic” and “likely pathogenic” categories from dbVAR [27], showed a total of 91 copy number variations (CNVs) present in the human CAPN10 gene as shown in **Table 3**. 84 CNVs were in the “pathogenic” category and the remaining 7 were “likely pathogenic”. After screening for clinical condition relevance with ClinVaR[28], none of the terms were found to be linked to PCOS, insulin resistance or Type 2 Diabetes mellitus.

Table 3. Copy number variations in humanCAPN10 gene

Variant	ID	Variant Type	Clinical Significance
1. nsv3910630	48473985	Copy Number Variation (CNV)	Pathogenic
2. nsv3898929	48462284	Copy Number Variation (CNV)	Pathogenic
3. nsv4768328	50453133	Copy Number Variation (CNV)	Pathogenic
4. nsv6315436	53680383	Copy Number Variation (CNV)	Pathogenic
5. nsv3871027	48434382	Copy Number Variation (CNV)	Pathogenic
6. nsv3910501	48473856	Copy Number Variation (CNV)	Pathogenic

7. nsv3894767	48458122	Copy Number Variation (CNV)	Pathogenic
8. nsv3892808	48456163	Copy Number Variation (CNV)	Pathogenic
9. nsv3877071	48440426	Copy Number Variation (CNV)	Pathogenic
10. nsv4674624	50271449	Copy Number Variation (CNV)	Pathogenic
11. nsv3899818	48463173	Copy Number Variation (CNV)	Pathogenic
12. nsv3907279	48470634	Copy Number Variation (CNV)	Pathogenic
13. nsv3892556	48455911	Copy Number Variation (CNV)	Pathogenic
14. nsv3908885	48472240	Copy Number Variation (CNV)	Pathogenic
15. nsv3893680	48457035	Copy Number Variation (CNV)	Pathogenic
16. nsv3881079	48444434	Copy Number Variation (CNV)	Pathogenic
17. nsv3904002	48467357	Copy Number Variation (CNV)	Pathogenic
18. nsv3888182	48451537	Copy Number Variation (CNV)	Pathogenic
19. nsv6636673	54355502	Copy Number Variation (CNV)	Pathogenic
20. nsv3894098	48457453	Copy Number Variation (CNV)	Pathogenic
21. nsv3904272	48467627	Copy Number Variation (CNV)	Pathogenic
22. nsv6315457	53680404	Copy Number Variation (CNV)	Pathogenic
23. nsv3903735	48467090	Copy Number Variation (CNV)	Pathogenic
24. nsv3904021	48467376	Copy Number Variation (CNV)	Pathogenic
25. nsv6636841	54355670	Copy Number Variation (CNV)	Pathogenic
26. nsv5381323	51636580	Copy Number Variation (CNV)	Pathogenic
27. nsv3892921	48456276	Copy Number Variation (CNV)	Pathogenic
28. nsv4347295	49342208	Copy Number Variation (CNV)	Pathogenic
29. nsv3895143	48458498	Copy Number Variation (CNV)	Pathogenic
30. nsv4674036	50270861	Copy Number Variation (CNV)	Pathogenic
31. nsv6314846	53679120	Copy Number Variation (CNV)	Pathogenic
32. nsv3901109	48464464	Copy Number Variation (CNV)	Pathogenic
33. nsv6291434	53636829	Copy Number Variation (CNV)	Pathogenic
34. nsv3885702	48449057	Copy Number Variation (CNV)	Pathogenic
35. nsv3908432	48471787	Copy Number Variation (CNV)	Pathogenic
36. nsv3907251	48470606	Copy Number Variation (CNV)	Pathogenic
37. nsv3880305	48443660	Copy Number Variation (CNV)	Pathogenic
38. nsv3903264	48466619	Copy Number Variation (CNV)	Pathogenic
39. nsv6311660	53675531	Copy Number Variation (CNV)	Pathogenic
40. nsv3893525	48456880	Copy Number Variation (CNV)	Pathogenic
41. nsv3902809	48466164	Copy Number Variation (CNV)	Pathogenic
42. nsv3899161	48462516	Copy Number Variation (CNV)	Pathogenic
43. nsv3889388	48452743	Copy Number Variation (CNV)	Pathogenic
44. nsv4347297	49342210	Copy Number Variation (CNV)	Pathogenic
45. nsv6634398	54348701	Copy Number Variation (CNV)	Pathogenic

46. nsv6291102	53636497	Copy Number Variation (CNV)	Pathogenic
47. nsv3902426	48465781	Copy Number Variation (CNV)	Pathogenic
48. nsv4436652	49580256	Copy Number Variation (CNV)	Pathogenic
49. nsv3888880	48452235	Copy Number Variation (CNV)	Pathogenic
50. nsv3882185	48445540	Copy Number Variation (CNV)	Pathogenic
51. nsv4674409	50271234	Copy Number Variation (CNV)	Pathogenic
52. nsv3878365	48441720	Copy Number Variation (CNV)	Pathogenic
53. nsv4673906	50270731	Copy Number Variation (CNV)	Pathogenic
54. nsv3882086	48445441	Copy Number Variation (CNV)	Pathogenic
55. nsv3904723	48468078	Copy Number Variation (CNV)	Pathogenic
56. nsv3883138	48446493	Copy Number Variation (CNV)	Pathogenic
57. nsv3908556	48471911	Copy Number Variation (CNV)	Pathogenic
58. nsv3908282	48471637	Copy Number Variation (CNV)	Pathogenic
59. nsv3909019	48472374	Copy Number Variation (CNV)	Pathogenic
60. nsv1398390	30348053	Copy Number Variation (CNV)	Pathogenic
61. nsv3904190	48467545	Copy Number Variation (CNV)	Pathogenic
62. nsv491548	1198794	Copy Number Variation (CNV)	Pathogenic
63. nsv6315390	53680337	Copy Number Variation (CNV)	Pathogenic
64. nsv3874648	48438003	Copy Number Variation (CNV)	Pathogenic
65. nsv3885544	48448899	Copy Number Variation (CNV)	Pathogenic
66. nsv3882615	48445970	Copy Number Variation (CNV)	Pathogenic
67. nsv4728725	50372362	Copy Number Variation (CNV)	Pathogenic
68. nsv3890898	48454253	Copy Number Variation (CNV)	Pathogenic
69. nsv3896499	48459854	Copy Number Variation (CNV)	Pathogenic
70. nsv3908112	48471467	Copy Number Variation (CNV)	Pathogenic
71. nsv3892823	48456178	Copy Number Variation (CNV)	Pathogenic
72. nsv3871121	48434476	Copy Number Variation (CNV)	Pathogenic
73. nsv3894025	48457380	Copy Number Variation (CNV)	Pathogenic
74. nsv3884002	48447357	Copy Number Variation (CNV)	Pathogenic
75. nsv3897593	48460948	Copy Number Variation (CNV)	Pathogenic
76. nsv3905997	48469352	Copy Number Variation (CNV)	Pathogenic
77. nsv4673913	50270738	Copy Number Variation (CNV)	Pathogenic
78. nsv3885422	48448777	Copy Number Variation (CNV)	Pathogenic
79. nsv3902416	48465771	Copy Number Variation (CNV)	Pathogenic
80. nsv3890832	48454187	Copy Number Variation (CNV)	Pathogenic
81. nsv6311563	53675434	Copy Number Variation (CNV)	Pathogenic
82. nsv4674620	50271445	Copy Number Variation (CNV)	Pathogenic
83. nsv1397950	30347613	Copy Number Variation (CNV)	Pathogenic
84. nsv3904659	48468014	Copy Number Variation (CNV)	Pathogenic

85. nsv4347296	49342209	Copy Number Variation (CNV)	Likely Pathogenic
86. nsv3877136	48440491	Copy Number Variation (CNV)	Likely Pathogenic
87. nsv3898945	48462300	Copy Number Variation (CNV)	Likely Pathogenic
88. nsv3887397	48450752	Copy Number Variation (CNV)	Likely Pathogenic
89. nsv3906254	48469609	Copy Number Variation (CNV)	Likely Pathogenic
90. nsv3892918	48456273	Copy Number Variation (CNV)	Likely Pathogenic
91. nsv3881218	48444573	Copy Number Variation (CNV)	Likely Pathogenic

CAPN10 gene expression patterns in PCOS GEO datasets

To investigate the differential expression of the CAPN10 gene in the context of PCOS, differential expression of microarray data obtained from three distinct datasets (GDS3104 [29], GDS4133 [30], and GDS4399 [31]) from NCBI's GEO were performed. CAPN10 gene had no significant differential expression when comparisons were performed between PCOS with insulin resistance and healthy controls as illustrated in **Fig. 10.** Therefore, no further analysis like functional enrichments were performed.

GEO ID	Gene	Nucleotide.Title	logFC	AveExpr	P.Value	adj.P.Val
GDS3104	CAPN10	Homo sapiens calpain 10 (CAPN10), transcript variant 1, mRNA	-0.079428623	7.404371443	0.447657316	0.817949449
GDS4133	CAPN10	Homo sapiens calpain 10 (CAPN10), transcript variant 1, mRNA	0.091858604	7.742038328	0.395067046	0.606698613
GDS4399	CAPN10	Homo sapiens calpain 10 (CAPN10), transcript variant 1, mRNA	1.940137935	7.121850684	0.010174355	0.479233975

Figure 10. CAPN10 Gene expression in GEO PCOS datasets. Significant differential expression was not found in human CAPN10 gene across three different PCOS GEO datasets.

PCOSKB Functional Enrichment Analysis

In order to get a generic idea about possible pathway and ontology enrichments between PCOS, Insulin resistance and diabetes, a Venn analysis was performed using the tool on PCOSKB [17, 18]. As seen in **Fig. 11,** several terms related to insulin signaling were observed among the 45 pathway and 196 ontology terms that were common between all three conditions. Interestingly, no

commonalties were found between insulin resistance and diabetes conditions. In terms of PCOS and insulin resistance, pathway and ontology terms related to thyroid hormone synthesis overlapped.

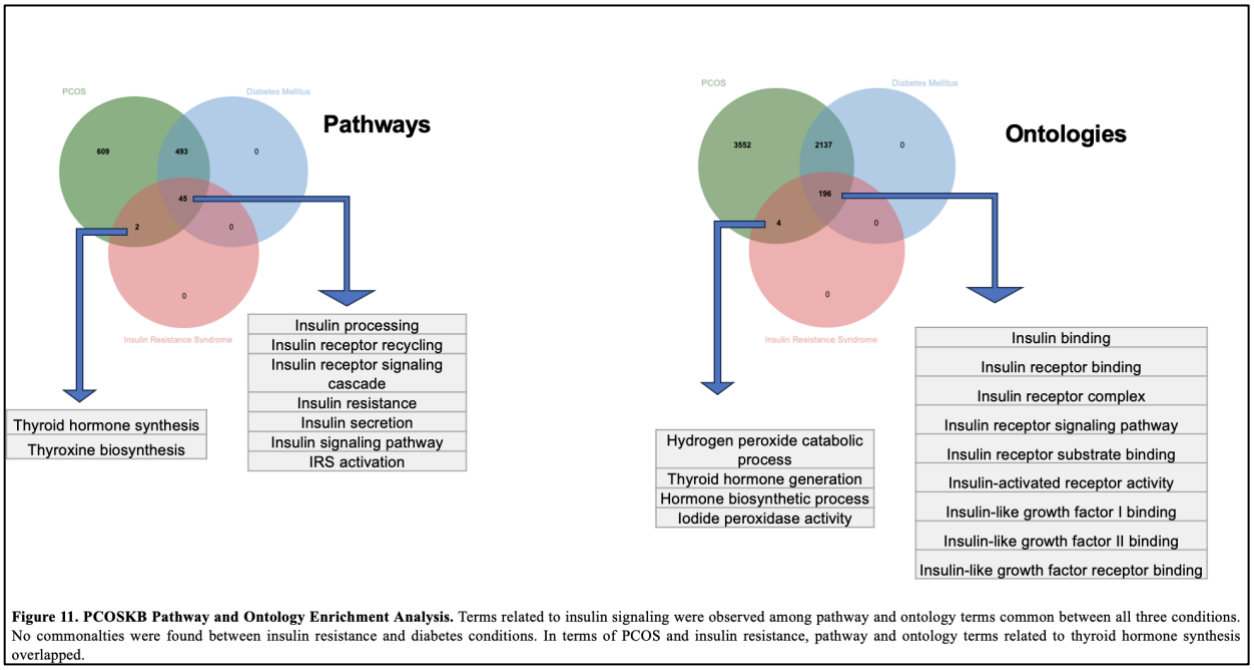

Figure 11. PCOSKB Pathway and Ontology Enrichment Analysis. Terms related to insulin signaling were observed among pathway and ontology terms common between all three conditions. No commonalties were found between insulin resistance and diabetes conditions. In terms of PCOS and insulin resistance, pathway and ontology terms related to thyroid hormone synthesis overlapped.

Discussion

The findings from this project reveal that the CAPN10 gene is expressed across various human issues, indicating its ubiquitous presence. Additionally, the gene is well conserved among different species, highlighting its evolutionary significance. PCOS is a human condition. However, animal models, specifically rodents, have been extensively used to study PCOS [32, 33]. These models do not naturally exhibit PCOS-like symptoms. Excess insulin in various tissues of PCOS women promotes insulin resistance, leading to excess androgen production or hyperandrogenism [34], which is one of the main criteria for PCOS diagnosis (**Table 1**). Therefore, to develop an animal model that mimics PCOS phenotype, androgen induction is necessary. It is interesting to note that some animals show polycystic ovaries but not hyperandrogenism [33]. From an

evolutionary perspective, insulin resistance is conserved across all species, including humans [35], similar to what our results demonstrated with CAPN10 gene. The question as to why humans and women develop PCOS and its symptoms, such as insulin resistance and hyperandrogenism, while other species do not despite having insulin resistance, remains unanswered. This raises intriguing research possibilities, highlighting a potential avenue for further exploration into other distinct factors genes, and mechanisms that may play a role.

The identified SNPs on CAPN10 gene in this project were found to be implicated in Diabetes and PCOS, although no terminology associated with insulin resistance was explicitly observed. Notably, 91 structural variants on CAPN10 were observed, but these were found to not have any clinical relevance. Future investigations are needed to examine these CAPN10 CNVs in the context of its relevance to PCOS, insulin resistance or Diabetes.

In terms of the current PCOS databases, our analysis revealed insulin signaling related enriched terms between PCOS, Diabetes, and insulin resistance. However, an interesting outcome was the enrichment of terms related to thyroid hormone-related processes linking PCOS and insulin resistance. Previous research has shown close associations between PCOS, insulin resistance and thyroid functions [36–38]. However, whether CAPN10 plays a role in the interplay between PCOS, insulin resistance and thyroid functions remains to be investigated. Another interesting aspect of the current PCOS database is that it needs to be updated. Evidence suggests that PCOS is only linked to Type 2 Diabetes mellitus [6, 7]. With the Venn analysis that was performed on PCOSKB in this project, it is unclear if the PCOSKB datasets are linked to only one type of Diabetes mellitus (type 1/2) or encompass both, which warrants major database upgrades. Irrespective of this missing aspect in this database, CAPN10 is strongly associated with PCOS, Type 2 Diabetes and insulin resistance, corroborating with literature evidence. Moreover, it is to

be noted that the existing information stored in this database give us an idea of possible functional overlaps between PCOS and associated conditions for future therapeutic interventions and research.

In this project, CAPN10 did not exhibit significant differential expression in three distinct PCOS GEO datasets. Upcoming research should also focus on incorporating transcriptomics to gain a comprehensive knowledge on CAPN10 and other related genes in the context of PCOS, Diabetes, and insulin resistance.

The initial intent of this project was to assess four PCOS candidate genes CAPN10, PPARγ, INS, and INSR. First, each gene was searched on dBSNP and dbVAR for human species "Homo Sapiens" only. Without adding any additional filter criteria, this search resulted in preliminary hits as demonstrated in **Fig. 12**. Considering the volume and complexity of the data resulting from this analysis of all four genes, the project plan had to be changed and limited to one gene. As my goal was to understand, the role of genetic and environmental factors in PCOS, insulin resistance, and possibly type 2 Diabetes Mellitus, this project was primarily focused on CAPN10 only. The other three genes PPARγ, INS, and INSR, were kept as secondary or backup if limited data was obtained upon genetic analysis for CAPN10. The data was limited for the project in terms of CAPN10. Another deviation from the initial intent of the project was to not use tools such Weighted gene correlated network analysis (WGCNA), Qiagen's Ingenuity Pathway Analysis tool (IPA) and StringDB on

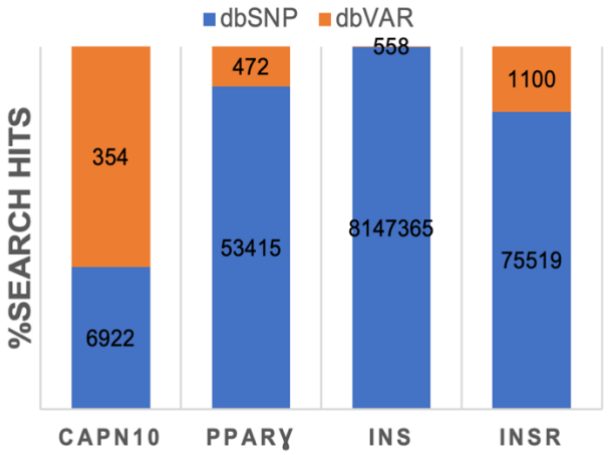

Figure 12. Number of search hits for four PCOS candidate genes. CAPN10, PPARγ, INS, and INSR are the four candidate genes searched on dbSNP and dbVAR for variation incidences. CAPN10 had 6922 and 354 hits in dbSNP and dbVAR, respectively. PPAR γ, INS and INSR had 53415, 8147365 and 75519 hits in dbSNP, respectively. PPARγ, INS and INSR had 472, 558 and 1100 hits in dbVAR, respectively.

possible datasets to look at smoking or vaping impact on PCOS. This was not accomplished due to lack of smoking or vaping related datasets available, complex tools and licensing issues such as for IPA. However, other ways were found to address this and to successfully complete the project.

Overall, the findings from this project provided a comprehensive overview of CAPN10 and its association with Diabetes, and insulin resistance.

Methods and Datasets

CAPN10 gene expression in human tissues

Tissue-specific gene and protein expression levels of PCOS candidate gene, CAPN10 was compiled using the Human Protein Atlas (HPA; Version: 23.0; Atlas updated: 2023-06-19) [22], GTex portal (GTEx dbGaP release V8) [23], and EMBL-EBI Expression Atlas (Expression Atlas release 39 – July 2023) [24]. The tissue expression profiles obtained from these resources in humans were compared with each other to look for unique CAPN10 expression patterns at RNA and or protein levels.

Inter-species conservation of CAPN10

UCSC Genome browser [25] (Last updated 08 Dec 2023) was utilized to observed CAPN10 gene conservation across different species in reference to human reference assembly GRCh38/hg38 (Human Dec. 2013 (GRCh38/hg38) (hg38, GRCh38 Genome Reference Consortium Human Reference 38 (GCA_000001405.15). The number of intron or non-coding regions, number of exons or coding regions, untranslated regions or UTR were examined on the CAPN10 gene. The human CAPN10 gene was first compared against animal models used for PCOS research such as rhesus monkey, mice, rats, cows, and dogs, and pigs. A follow-up comparison with all species available on UCSC Genome browser was also performed.

CAPN10 genetic variations and its clinical relevance to PCOS and insulin resistance

NCBI resources such as dBSNP [26] and dbVAR [27] was used to collect information on all gene mutations like single nucleotide polymorphisms and structural variants in human CAPN10 gene. Specifically, the search term used was "CAPN10 AND Homo sapiens" for conducting both searches. Subsequently, the outcomes were filtered based on their clinical significance categories such as "Risk-factor", "Pathogenic" and "Likely pathogenic". These filtered terms were then utilized for determining clinical relevance to the terms "PCOS, insulin resistance or Type 2 Diabetes mellitus" with ClinVaR[28] and PCOSKB [17, 18] (last updated in 2019).

CAPN10 gene expression patterns in PCOS GEO datasets

Three microarray datasets, GDS3104 [29], GDS4133 [30], and GDS4399 [31], each comprising of PCOS patients with insulin resistance and healthy controls, were retrieved from the NCBI GEO database using the GEOquery package [39] in R. Differential expression analysis comparing gene expression profiles between PCOS patients and healthy controls within each dataset was performed using the limma package in R [40]. The resulting dataset was filtered based on p-values less than 0.05 to isolate genes possibly including CAPN10, exhibiting statistically significant differential expression.

PCOSKB Functional Enrichment Analysis

Venn analysis tool on PCOSKB [17, 18] (last updated in 2019). was used to perform pathway and ontology functional enrichments between PCOS, insulin resistance and Diabetes. Possible pathway and ontology term overlaps between these conditions were determined.

References

1. Deswal R, Narwal V, Dang A, Pundir CS. The Prevalence of Polycystic Ovary Syndrome: A

Brief Systematic Review. J Hum Reprod Sci. 2020;13:261.

2. Goodarzi MO, Dumesic DA, Chazenbalk G, Azziz R. Polycystic ovary syndrome: etiology, pathogenesis and diagnosis. Nat Rev Endocrinol 2011 74. 2011;7:219–31.

3. Mumusoglu S, Yildiz BO. Polycystic ovary syndrome phenotypes and prevalence: Differential impact of diagnostic criteria and clinical versus unselected population. Curr Opin Endocr Metab Res. 2020;12:66–71.

4. Azziz R. Reproductive endocrinology and infertility: Clinical expert series polycystic ovary syndrome. Obstet Gynecol. 2018;132:321–36.

5. Panidis D, Tziomalos K, Misichronis G, Papadakis E, Betsas G, Katsikis I, et al. Insulin resistance and endocrine characteristics of the different phenotypes of polycystic ovary syndrome: a prospective study. Hum Reprod. 2012;27:541–9.

6. Sachdeva G, Gainder S, Suri V, Sachdeva N, Chopra S. Comparison of the Different PCOS Phenotypes Based on Clinical Metabolic, and Hormonal Profile, and their Response to Clomiphene. Indian J Endocrinol Metab. 2019;23:326.

7. Singh S, Pal N, Shubham S, Sarma DK, Verma V, Marotta F, et al. Polycystic Ovary Syndrome: Etiology, Current Management, and Future Therapeutics. J Clin Med. 2023;12:12.

8. Straznicky NE, Nestel PJ, Esler MD. Autonomic Nervous System: Metabolic Function. Encycl Neurosci. 2009;:951–9.

9. Vassilatou E. Nonalcoholic fatty liver disease and polycystic ovary syndrome. World J Gastroenterol. 2014;20:8351.

10. Govind A, Obhrai MS, Clayton RN. Polycystic Ovaries Are Inherited as an Autosomal Dominant Trait: Analysis of 29 Polycystic Ovary Syndrome and 10 Control Families. J Clin Endocrinol Metab. 1999;84:38–43.

11. Khan MJ, Ullah A, Basit S. Genetic Basis of Polycystic Ovary Syndrome (PCOS): Current Perspectives. Appl Clin Genet. 2019;12:249.

12. Kshetrimayum C, Sharma A, Mishra VV, Kumar S. Polycystic ovarian syndrome: Environmental/occupational, lifestyle factors; an overview. J Turkish Ger Gynecol Assoc. 2019;20:255.

13. Lakkakula BVKS, Thangavelu M, Godla UR. Genetic variants associated with insulin signaling and glucose homeostasis in the pathogenesis of insulin resistance in polycystic ovary syndrome: a systematic review. J Assist Reprod Genet. 2013;30:883.

14. Urbanek M. The genetics of the polycystic ovary syndrome. Nat Clin Pract Endocrinol Metab 2007 32. 2007;3:103–11.

15. Babatunde Sikiru A, Adeniran A, Akinola K, Behera H, Kalaignazhal G, Sunday S, et al. Unraveling the complexity of the molecular pathways associated with polycystic ovary syndrome (PCOS) and identifying molecular targets for therapeutic development: a review of literature. East Fertil Soc J. 2023;28:16.

16. Ridderstråle M, Nilsson E. Type 2 diabetes candidate gene CAPN10: First, but not last. Curr Hypertens Rep. 2008;10:19–24.

17. Joseph S, Barai RS, Bhujbalrao R, Idicula-Thomas S. PCOSKB: A KnowledgeBase on genes, diseases, ontology terms and biochemical pathways associated with PolyCystic Ovary Syndrome. Nucleic Acids Res. 2016;44:D1032–5.

18. Sharma M, Barai RS, Kundu I, Bhaye S, Pokar K, Idicula-Thomas S. PCOSKBR2: a database of genes, diseases, pathways, and networks associated with polycystic ovary syndrome. Sci Reports 2020 101. 2020;10:1–11.

19. Sáez ME, González-Sánchez JL, Ramírez-Lorca R, Martínez-Larrad MT, Zabena C, González

A, et al. The CAPN10 Gene Is Associated with Insulin Resistance Phenotypes in the Spanish Population. PLoS One. 2008;3:e2953.

20. Dasgupta S, Sirisha PVS, Neelaveni K, Anuradha K, Reddy BM. Association of CAPN10 SNPs and Haplotypes with Polycystic Ovary Syndrome among South Indian Women. PLoS One. 2012;7:e32192.

21. Li Y, Han T, Wang Y, Gao J, Zhang J, Wu Y, et al. Association of Calpain10 polymorphisms with polycystic ovarian syndrome susceptibility: a systematic review and meta-analysis with trial sequential analysis. Front Genet. 2023;14:1153960.

22. Pontén F, Jirström K, Uhlen M. The Human Protein Atlas--a tool for pathology. J Pathol. 2008;216:387–93.

23. Lonsdale J, Thomas J, Salvatore M, Phillips R, Lo E, Shad S, et al. The Genotype-Tissue Expression (GTEx) project. Nat Genet 2013 456. 2013;45:580–5.

24. Moreno P, Fexova S, George N, Manning JR, Miao Z, Mohammed S, et al. Expression Atlas update: gene and protein expression in multiple species. Nucleic Acids Res. 2022;50:D129–40.

25. Kent WJ, Sugnet CW, Furey TS, Roskin KM, Pringle TH, Zahler AM, et al. The Human Genome Browser at UCSC. Genome Res. 2002;12:996–1006.

26. Sherry ST, Ward MH, Kholodov M, Baker J, Phan L, Smigielski EM, et al. dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. 2001;29:308.

27. Lappalainen I, Lopez J, Skipper L, Hefferon T, Spalding JD, Garner J, et al. DbVar and DGVa: public archives for genomic structural variation. Nucleic Acids Res. 2013;41 Database issue.

28. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res. 2018;46:D1062–7.

29. Skov V, Glintborg D, Knudsen S, Jensen T, Kruse TA, Tan Q, et al. Reduced expression of

nuclear-encoded genes involved in mitochondrial oxidative metabolism in skeletal muscle of insulin-resistant women with polycystic ovary syndrome. Diabetes. 2007;56:2349–55.

30. Skov V, Glintborg D, Knudsen S, Tan Q, Jensen T, Kruse TA, et al. Pioglitazone Enhances Mitochondrial Biogenesis and Ribosomal Protein Biosynthesis in Skeletal Muscle in Polycystic Ovary Syndrome. PLoS One. 2008;3:e2466.

31. Kaur S, Archer KJ, Devi MG, Kriplani A, Strauss JF, Singh R. Differential Gene Expression in Granulosa Cells from Polycystic Ovary Syndrome Patients with and without Insulin Resistance: Identification of Susceptibility Gene Sets through Network Analysis. J Clin Endocrinol Metab. 2012;97:E2016–21.

32. Stener-Victorin E. Update on Animal Models of Polycystic Ovary Syndrome. Endocrinology. 2022;163.

33. Ryu Y, Kim SW, Kim YY, Ku SY. Animal Models for Human Polycystic Ovary Syndrome (PCOS) Focused on the Use of Indirect Hormonal Perturbations: A Review of the Literature. Int J Mol Sci. 2019;20.

34. Marshall JC, Dunaif A. All Women With PCOS Should Be Treated For Insulin Resistance. Fertil Steril. 2012;97:18.

35. Erol A. Insulin resistance is an evolutionarily conserved physiological mechanism at the cellular level for protection against increased oxidative stress. Bioessays. 2007;29:811–8.

36. Mueller A, Schöfl C, Dittrich R, Cupisti S, Oppelt PG, Schild RL, et al. Thyroid-stimulating hormone is associated with insulin resistance independently of body mass index and age in women with polycystic ovary syndrome. Hum Reprod. 2009;24:2924–30.

37. Dittrich R, Kajaia N, Cupisti S, Hoffmann I, Beckmann MW, Mueller A. Association of thyroid-stimulating hormone with insulin resistance and androgen parameters in women with

PCOS. Reprod Biomed Online. 2009;19:319–25.

38. Palomba S, Colombo C, Busnelli A, Caserta D, Vitale G. Polycystic ovary syndrome and thyroid disorder: a comprehensive narrative review of the literature. Front Endocrinol (Lausanne). 2023;14:1251866.

39. Sean D, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics. 2007;23:1846–7.

40. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015;43:e47–e47.