Quantitative Trait Meta-Analysis Identifies Rare Noncoding Variants Associated with Altered Hormone Levels in Polycystic Ovary Syndrome

Matthew Dapas[1], Ryan Sisk[1], Richard S. Legro[2], Margrit Urbanek[1,3,4], Andrea Dunaif[5,*], M. Geoffrey Hayes[1,3,6,*]

[1] Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL; [2] Department of Obstetrics and Gynecology, Penn State College of Medicine, Hershey, PA; [3] Center for Genetic Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL; [4] Center for Reproductive Science, Northwestern University Feinberg School of Medicine, Chicago, IL; [5] Division of Endocrinology, Diabetes and Bone Disease, Icahn School of Medicine at Mount Sinai, New York, NY; [6] Department of Anthropology, Northwestern University, Evanston, IL; * These authors jointly supervised this work.

Background

- Polycystic ovary syndrome (PCOS) is a complex genetic disorder characterized by hyperandrogenism, chronic anovulation, and polycystic ovarian morphology.
- PCOS affects up to 15% of premenopausal women worldwide [1].
- Common risk alleles identified to date confer modest increases in disease risk and account for a small proportion of the estimated genetic heritability of PCOS.

Hypothesis

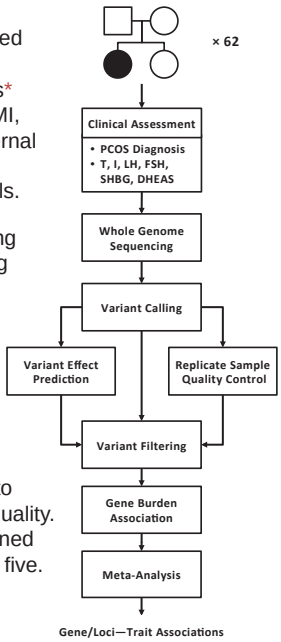
Rare variants contribute to the pathogenesis of PCOS.

Methods

Subjects: 62 two-generation families (avg. size = 4.2) with one or more affected daughters. All probands fulfilled NIH criteria for PCOS [2]. The following traits* were recorded for each subject: age, BMI, T, I, DHEAS, SHBG, LH, and FSH. Paternal hormone levels were excluded from all association tests, except for insulin levels.

Sequencing: Whole genome sequencing and variant calling were performed using the Complete Genomics pipeline on GRCh37. The Scripps Wellderly Genome Resource served as the control population [3].

Variant Selection: Rare variants (MAF ≤ 2%) were filtered for predicted, gene-specific deleterious effects [4-5]. Variants were filtered further according to Mendelian inheritance and variant-call quality. Optimal quality thresholds were determined using replicate samples from a family of five.



Association Testing: Variants in gene regions (including 3' UTR, introns, and 7.5kb upstream of 5' TSS) and sliding windows were tested for association with PCOS and its hormonal traits using a family-based variance component association test [6]. Age and BMI were included as covariates. Residuals were corrected for normality using an inverse-normal-transformation. Variants were weighted according to their predicted levels of deleteriousness [4]. Coding and noncoding variants were tested independently. Coding variant associations were further adjusted by relative variation intolerance [7].

Meta-Analysis: Associations between gene regions and distinct quantitative traits were combined into a single meta-statistic using a modified Fisher's combined probability test for correlated traits [8]. Results were corrected for multiple testing and genomic inflation, including all groupings with at least one variant. Only genes with rare, deleterious variants in at least 10% of cases were considered.

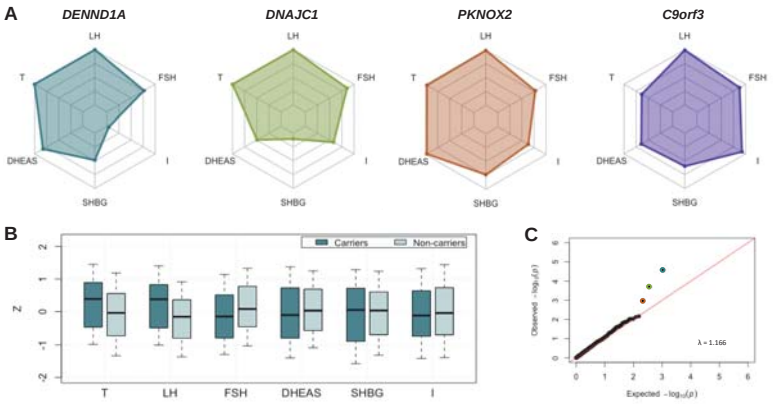
* T: testosterone; I: Insulin; LH: luteinizing hormone; FSH: follicle-stimulating hormone; SHBG: sex hormone-binding globulin; DHEAS: dehydroepiandrosterone

Results

Noncoding Quantitative Trait Association Meta-Analysis

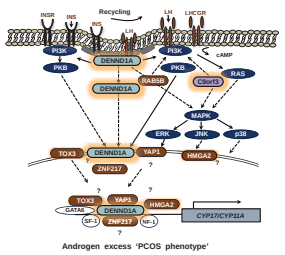
Chr	Gene / Locus	Length	Variants	Families[1]	Aff : Unaff OR[2]	*P*	P_{adj}
9	***DENND1A*** ***(q33.3)***	**550kb**	**30**	**47%**	**1.29** **[0.78 – 2.22]**	**2.66×10^{-5}**	**0.017**
10	***DNAJC1*** ***(p12.31)***	247kb	6	11%	1.40 [0.89 – 2.40]	1.95×10^{-4}	0.12
11	***PKNOX2*** ***(q24.2)***	269kb	18	39%	1.54 [0.91 – 2.81]	1.07×10^{-3}	0.67
9	***C9orf3***[3] ***(q22.32)***	420kb	9	19%	1.79 [0.74 - 5.04]	1.45×10^{-2}	1.00

1. Number of families with ≥1 gene variant present in multiple generations
2. Odds ratio estimates based on number of variants in affected vs. unaffected subjects
3. C9orf3 included for reference as next highest-ranking PCOS GWAS gene in meta-analysis results (#12 / 542)



A) Relative ranks of individual trait associations for top genes and *C9orf3*
B) Residual trait differences in subjects with ≥1 rare variant in *DENND1A*
C) P-values for meta-analysis of family-based SKAT trait associations

Hypothesized Androgen Signaling Cascade [9]



- Multiple association studies have found associations between PCOS and common, noncoding variants in both *DENND1A* and *C9orf3* [10-13].
- PCOS theca cells have higher levels of a *DENND1A* isoform, DENND1A.V2 [14].
- Overexpression of DENND1A.V2 in theca cells leads to increased androgen biosynthesis [14].

Conclusions

- Our findings suggest that rare, noncoding variants in *DENND1A* contribute to elevated androgen and LH levels in PCOS.
- The *DENND1A* and *C9orf3* variant associations further strengthen the evidence for their integral involvement in PCOS pathogenesis.
- A family-based, quantitative trait meta-analysis can be a powerful approach to rare variant association testing.

References

1. Fauser, 2012, *Fertil Steril* 97:28
2. Legro, 1998, *PNAS USA* 87:2134
3. Erikson, 2016, *Cell* 165:1002
4. Kircher, 2014, *Nat Genet* 46:310
5. Huang, 2017, *Nat Genet* 49:618
6. Schaid, 2013, *Genet Epi* 37:409
7. Petrovski, 2013, *PLOS Genet* 9
8. Yang, 2016, *BMC Bioinformatics* 17:19
9. McAllister, 2015, *TEM* 26
10. Chen, 2011, *Nat Genet* 43:55
11. Goodarzi, 2012, *J Med Genet* 49:90
12. Welt, 2012, *J Clin Endocrinol Metab* 97
13. Hayes, 2015, *Nat Commun* 6:7502
14. McAllister, 2014 *PNAS USA* 111:15

Northwestern Medicine®
Feinberg School of Medicine

ACKNOWLEDGEMENTS: This study was supported by US National Institutes of Health (NIH) grants R01 HD057223 (A.D.), P50 HD044405 (A.D.), U54 HD34449 (A.D.), R01 HD057450 (M.U.), and R01 HD056510 (R.S.L.), M.D. was primarily supported by NRSA fellowship T32 DK007169, and partially supported by NIH award R01LM011297 during the study. This study uses data from the Scripps Wellderly Genome Resource, which is funded under NIH grant 5 UL1 TR001114 Scripps Translational Science Institute CTSA Award. We thank Rick Tearle, PhD, (Complete Genomics, Inc.), who generated the variant call reports and provided informatics and data support. We thank Terry Farrah and Gustavo Glusman, PhD, from the Institute for Systems Biology for facilitating variant queries via Kaviar (http://db.systemsbiology.net/kaviar). We are very grateful to all of the families who took part in the study.