Title
Harnessing expression data to identify novel candidate genes in polycystic ovary syndrome.

Permalink
https://escholarship.org/uc/item/08f4f9sv

Journal
PloS one, 6(5)

ISSN
1932-6203

Authors
Jones, Michelle R
Chua, Angela
Chen, Yii-Der I
et al.

Publication Date
2011

DOI
10.1371/journal.pone.0020120

Peer reviewed
Harnessing Expression Data to Identify Novel Candidate Genes in Polycystic Ovary Syndrome

Michelle R. Jones¹, Angela Chua¹, Yii-Der I. Chen²,³, Xiaohui Li³, Ronald M. Krauss⁶, Jerome I. Rotter³,⁴, Richard S. Legro⁷, Ricardo Azziz²,⁴,⁵, Mark O. Goodarzi¹,²,³,⁴*

¹Division of Endocrinology, Diabetes and Metabolism, Department of Medicine, Cedars-Sinai Medical Center, Los Angeles, California, United States of America, ²Department of Obstetrics and Gynecology, Cedars-Sinai Medical Center, Los Angeles, California, United States of America, ³Medical Genetics Institute, Cedars-Sinai Medical Center, Los Angeles, California, United States of America, ⁴Department of Medicine, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California, United States of America, ⁵Department of Obstetrics and Gynecology, David Geffen School of Medicine, University of California Los Angeles, Los Angeles, California, United States of America, ⁶Children’s Hospital Oakland Research Institute, Oakland, California, United States of America, ⁷Department of Obstetrics and Gynecology, Pennsylvania State College of Medicine, Hershey, Pennsylvania, United States of America.

Abstract

Novel pathways in polycystic ovary syndrome (PCOS) are being identified in gene expression studies in PCOS tissues; such pathways may contain key genes in disease etiology. Previous expression studies identified both dickkopf homolog 1 (Dkk1) and DnaJ (Hsp40) homolog, subfamily B, member 1 (DNAJB1) as differentially expressed in PCOS tissue, implicating them as candidates for PCOS susceptibility. To test this, we genotyped a discovery cohort of 335 PCOS cases and 198 healthy controls for three DKK1 single nucleotide polymorphisms (SNPs) and four DNAJB1 SNPs and a replication cohort of 396 PCOS cases and 306 healthy controls for 1 DKK1 SNP and 1 DNAJB1 SNP. SNPs and haplotypes were determined and tested for association with PCOS and component phenotypes. We found that no single nucleotide polymorphisms were associated with PCOS risk; however, the major allele of rs1569198 from DKK1 was associated with increased total testosterone (discovery cohort P = 0.0035) and dehydroepiandrosterone sulfate (replication cohort P = 0.05). Minor allele carriers at rs3962158 from DNAJB1 had increased fasting insulin (discovery cohort P = 0.003), increased HOMA-IR (discovery cohort P = 0.006; replication cohort P = 0.036), and increased HOMA-%B (discovery cohort P = 0.004). Carriers of haplotype 2 at DNAJB1 also had increased fasting insulin, HOMA-IR, and HOMA-%B. These findings suggest that genetic variation in DKK1 and DNAJB1 may have a role in the hyperandrogenic and metabolic dysfunction of PCOS, respectively. Our results also demonstrate the utility of gene expression data as a source of novel candidate genes in PCOS, a complex and still incompletely defined disease, for which alternative methods of gene identification are needed.

Citation: Jones MR, Chua A, Chen Y-DI, Li X, Krauss RM, et al. (2011) Harnessing Expression Data to Identify Novel Candidate Genes in Polycystic Ovary Syndrome. PLoS ONE 6(5): e20120. doi:10.1371/journal.pone.0020120

Editor: Debashish Bhattacharya, Rutgers University, United States of America

Received January 25, 2011; Accepted April 13, 2011; Published May 17, 2011

Copyright: © 2011 Jones et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the National Institutes of Health [R01-HL069757 to R.M.K.]; National Center for Research Resources [M01-RR00425 to the Cedars-Sinai General Clinical Research Center]; Winnick Clinical Scholars Award [to M.O.G.]; and an endowment from the Helping Hand of Los Angeles, Inc. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist. The Helping Hand of Los Angeles, Inc. is a charitable support group at Cedars-Sinai Medical Center, whose fundraising efforts partly supported this study. This does not alter the authors’ adherence to all the PLoS ONE policies on sharing data and materials.

* E-mail: mark.goodarzi@cshs.org

Introduction

Familial aggregation and twin studies have established a genetic etiology for polycystic ovary syndrome (PCOS) [1]. The hallmark of PCOS is hyperandrogenemia; however, insulin resistance, pancreatic beta cell dysfunction and chronic inflammation are frequently present [1,2,3]. Many previous candidate gene studies focused on genes from androgen synthesis and insulin signaling pathways. Few susceptibility genes are widely agreed upon, potentially because candidate gene selection has been based on incomplete understanding of the disorder.

Recently a number of expression studies (mRNA and protein) have been performed in PCOS tissues, including ovary [4], omental fat [5] and lymphocytes [6]. Remarkably few genes have been reported as differentially expressed in PCOS tissues in more than one study. One such gene is dickkopf homolog 1 (Dkk1), an inhibitor of the Wnt signaling pathway and cell growth repressor [4,5]. Dkk1 interacts with the Wnt coreceptor low-density lipoprotein receptor-related proteins 5 and 6 and influences several functions, including embryogenesis and cell cycle regulation in cancer pathways [7]. It has been reported as being under expressed in omental fat [5] and over expressed in cultured ovarian theca [4] from PCOS subjects.

The second gene we selected for analysis in this study is DnaJ (Hsp40) homolog, subfamily B, member 1 (DNAJB1), whose expression was decreased in ovaries from PCOS women [8]. DNAJB1 was also selected as a positional and functional candidate. It acts in concert with molecular chaperones to regulate protein folding, protein complex assembly and disassembly, and transport of proteins across cellular membranes, particularly in the androgen signaling pathway, and is under transcriptional regulation by insulin [9]. It is also located within the chromosome 19p13.2 linkage region that has been identified in PCOS susceptibility [10], implicating DNAJB1 as a potential positional candidate.

Because logical candidates have led to few successes in PCOS genetics, novel methods of candidate gene selection are needed. In
the present study we used published expression data to identify putative molecular targets; selecting two genes, DKK1 and DNAJB1, from such analysis for association study with PCOS, conducted in two independent case/control cohorts. A SNP in DNAJB1 was associated with a measure of insulin resistance in women with PCOS in both cohorts.

Materials and Methods

Ethics statement
All subjects gave written informed consent; each study was approved by the Institutional Review Boards of the recruiting centers and Cedars-Sinai Medical Center.

Subjects and phenotyping

Discovery Cohort. We studied 335 unrelated White PCOS patients and 198 White control women recruited at two centers, the University of Alabama at Birmingham (UAB; 287 PCOS and 117 controls) and Cedars-Sinai Medical Center (CSMC; 48 PCOS and 11 controls). Cases were premenopausal, non-pregnant, on no hormonal therapy, including oral contraceptives, for at least three months; all PCOS subjects met 1990 NIH criteria [11]. Parameters for defining hirsutism, hyperandrogenemia, ovulatory dysfunction, and exclusion of related disorders were previously reported [12]. Controls were healthy women, with regular menstrual cycles and no evidence of hirsutism, acne, alopecia, or endocrine dysfunction and had not taken hormonal therapy (including oral contraceptives) for at least three months. Controls were recruited by word of mouth and advertisements calling for “healthy women.”

Replication Cohort. We assembled a cohort of 396 unrelated White PCOS patients and 306 White control women. The replication cohort was constituted from three sources: 380 PCOS subjects (all meeting 1990 NIH criteria) and 71 healthy controls previously recruited by R. S. Legro [13], 16 PCOS subjects and 2 healthy controls recruited at Cedars-Sinai Medical Center using the same criteria as those used in the discovery cohort; and 233 white control women derived from the Cholesterol and Atherosclerosis Pharmacogenetics (CAP) study, a component of the Pharmacogenomics and Risk of Cardiovascular Disease (PARC) Study [14].

Subjects recruited at UAB and CSMC were evaluated per a previously described protocol [12]. Fasting glucose and insulin were also obtained in a subset (70%) of the cases (non-diabetic). The subset of subjects with fasting glucose and insulin did not differ demographically or hormonally from the study subjects overall. The computer-based homeostasis model assessment (HOMA, www.dtu.ox.ac.uk/homa) was used to calculate indices of insulin resistance (HOMA-IR) and insulin secretion (i.e. percent beta-cell function or HOMA-%B) utilizing the fasting glucose and insulin levels. This computer model was also applied to generate HOMA data in the replication cohort, on whom fasting insulin and glucose were measured in >90% of subjects. Table 1 presents clinical characteristics of both cohorts.

Genotyping and haplotype determination

Discovery genotyping was performed on three DKK1 single nucleotide polymorphisms (SNPs) (rs2241529, rs1569198, rs2288355) and four DNAJB1 SNPs (rs7003, rs1803768, rs4926222, rs3962138) selected using genotyping data of the CEU (Utah residents with ancestry from northern and western Europe) population of the HapMap database (release 24, http://hapmap.ncbi.nlm.nih.gov/). These SNPs were selected because they are predicted to tag SNPs across the entirety of each gene, plus 10 kb upstream and 10 kb downstream. These SNPs capture 21 of 26 (81%) of the CEU HapMap SNPs at r²>0.8 for the two genes. Replication genotyping was performed on two of the discovery SNPs; rs1569198 from DKK1 and rs3962138 from DNAJB1. All genotyping was performed using Applied Biosystems Taqman Assays-on-Demand (Applied Biosystems, Foster City, CA) according to manufacturer’s instructions.

Haploview (version 4.1) was used to calculate linkage disequilibrium (LD, the D’ statistic) between each pair of SNPs and determine haplotypes and their frequencies [15]. The solid spine of LD algorithm in Haploview was used to determine haplotype blocks. Only subjects whose haplotype assignment was >95% certain were analyzed.

Statistical analysis

Unpaired t-tests and chi-square tests were used to compare clinical characteristics between cases and controls; quantitative traits were log- or square-root-transformed as appropriate to reduce non-normality. Data are presented as median (interquartile range). In the replication cohort, androgen measurements were available only in the cases and controls recruited by R. S. Legro. Abbreviations: BMI: body mass index; mFG: modified Ferriman-Gallwey hirsutism score; NA: Data not Available; HOMA-IR, homeostasis model assessment of insulin resistance; HOMA-%B, homeostasis model assessment of beta-cell function (insulin secretion).

doi:10.1371/journal.pone.0020120.0001

Table 1. Characteristics of PCOS and control subjects.

|                      | Discovery | Replication |
|----------------------|-----------|-------------|
|                      | Control   | PCOS        | Control   | PCOS        |
|                      | (n = 198) | (n = 335)   | (n = 306) | (n = 396)   |
| Age (yr)             | 32.5 (17.0) | 27.0 (11.1)* | 51.0 (23.7) | 28.0 (9.0)* |
| BMI (kg/m²)          | 24.7 (5.9)  | 33.4 (14.7)* | 25.9 (8.3)  | 34.9 (12.3)* |
| Total testosterone   | 1.40 (0.94) | 2.63 (1.14)* | 1.0 (0.52)  | 2.43 (1.11)* |
| (nmol/l)             |           |             |           |             |
| DHEAS (μmol/l)       | 2.58 (2.04) | 5.68 (4.57)* | 3.70 (1.75) | 5.72 (3.94)* |
| Insulin (pmol/l)     | 41.4 (39.0) | 96.0 (108.0)* | 72.0 (45.0) | 126.0 (96.0)* |
| Glucose (mmol/l)     | 4.77 (0.56) | 4.77 (0.72)  | 5.01 (0.68) | 4.86 (0.64)  |
| HOMA-IR              | 0.95 (0.84) | 2.01 (2.13)* | 1.58 (0.96) | 2.69 (1.83)* |
| HOMA-%B              | 104.1 (57.0) | 166.9 (108.8)* | 129.5 (58.3) | 196.6 (81.0)* |

*P<0.001 compared to control group. Data are median (interquartile range). In the replication cohort, androgen measurements were available only in the cases and controls recruited by R. S. Legro.

DKK1 and DNAJB1 Genes in Polycystic Ovary Syndrome

Results

Discovery cohort
We genotyped three SNPs from DKK1 and four SNPs from DNAJB1 in the discovery cohort (Table 2), where the overall
This haplotype carries the minor allele of rs3962158, which was also associated with each of these traits in the single marker analysis. These haplotype associations were not observed in the control women.

Repetition cohort

We selected the two significant SNPs from the discovery phase of the study for replication. The genotyping success rate for rs1569198 was 95.7% and for rs3962158 was 97.3%, with 100% concordance observed in duplicate samples run both within the replication study samples, and across the replication and discovery study samples. The minor allele frequencies of rs1569198 (PCOS 0.46, control 0.51, overall 0.49) and rs3962158 (PCOS 0.31, control 0.30, overall 0.31) were the same as the frequencies observed in the discovery cohort. We observed association (additive model) between DKK1 SNP rs1569198 and dehydroepiandrosterone sulfate (DHEAS), with increasing copies of the A allele correlating with increasing DHEAS (AA: 2229.0 (1551.3), AG: 2099.0 (1373.8), GG: 1978.0 (1520.0) nmol/l; P = 0.05). We did not replicate the association between total testosterone and rs1569198. We did, however, replicate the significant association between carriers of the minor T allele at rs3962158 of DNAJB1 with increased HOMA-IR under the same dominant model (CC: 2.57 (1.65) vs. CT/TT: 2.77 (1.98); P = 0.036).

Discussion

In an attempt to circumnavigate arbitrary bias introduced in the selection of candidate genes for PCOS, we evaluated mRNA and protein expression data reported from several PCOS tissues in order to identify novel susceptibility genes, DKK1 and DNAJB1. These results suggest that variation in gene expression may be an important factor in identifying relevant PCOS pathways. This strategy was employed to select the 17β-hydroxysteroid dehydrogenase type 6 (HSD17B6) gene, which exhibited increased activity in PCOS ovaries, as a candidate gene for PCOS [16]; we and others ultimately replicated association of variants in this gene with metabolic phenotypes of PCOS [17,18]. Others have performed network analysis on PCOS microarray data to identify pathways that may be most relevant to the development of the disorder [19].

A DKK1 variant, rs1569198, was associated with testosterone levels in PCOS subjects in our discovery cohort, with carriers of the major allele having an elevated testosterone level. This result is consistent with the observation that a murine homolog of DKK1 plays a role in testicular testosterone production [20]. DKK1 is expressed in a number of tissues including ovary, testis and adipose tissue, with increased expression in PCOS ovarian theca [4]. DKK1 has been identified as one of the most upregulated genes in testosterone responsive tissues, such as the dermal papilla [21], suggesting its expression may be regulated by androgens. In a recent report designed to identify proteins that act in androgen-dependent hair loss, DKK1 was reported as a dihydrotestosterone-inducible transcript, and also caused apoptosis in follicular keratinocytes [21]. These data suggest a role for DKK1 in the regulation of cell cycle in androgen responsive tissues, and further studies in ovary, particularly ovarian theca may confirm this. In a study of genetic determinants of bone phenotypes in men, rs1569198 was associated with hip axis length [22], which indicate potential roles for DKK1 and this SNP in particular in a number of pathways, including those under hormonal regulation.

That we did not replicate the association of rs1569198 with testosterone deserves comment. The replication of associations such as these presents several challenges; the most important is the accumulation of an appropriate replication cohort, which must consist of an equal or large sample size of ethnically matched individuals with PCOS.
subjects, with similar disease phenotype and quantitative trait measures. The appropriateness of our replication cohort is substantiated by our replication of the association between rs3962158 and HOMA-IR. The replication of genetic association with testosterone poses a particular challenge, given the difficulty in precisely measuring this trait, particularly in women [23,24]. In the replication cohort, we identified a nominal association of the major allele of rs1569198 with increased DHEAS, representing either a coincidence or supporting a role in androgen production or clearance. If the SNP does affect both testosterone and DHEAS levels, the lack of observation of association with both traits in both cohorts may be related to statistical chance or imprecision of androgen measurement in women.

Of particular interest as a candidate identified via differential expression was DNAJB1 as it lies in the previously identified PCOS linkage region (microsatellite marker D19S884) on chromosome 19p13.2 [10] and acts as a molecular chaperone. Androgen receptor function (including ligand binding and nuclear translocation) depends on its interaction with heat shock proteins and their co-chaperones [25]. DNAJB1 acts as a co-chaperone in a number of pathways including androgen receptor and glucocorticoid receptor signaling. An androgenic protein co-chaperone, DNAJB1 is under transcriptional regulation by insulin, with increased hepatic expression demonstrated under conditions of reduced insulin [9]. DNAJB1 may thus represent a common factor at the nexus of both the androgenic and insulin pathways that are frequently dysfunctional in PCOS.

The functional role of the associated DKK1 and DNAJB1 SNPs is unknown. The DKK1 SNP rs1569198 is an intronic SNP, and the DNAJB1 SNP rs3962158 does not change the amino acid at position 4. These SNPs, purposefully selected as tagging SNPs, may not be causal, but may be in LD with functional variants elsewhere in their genes or in the promoters. In particular, HapMap data indicate that the DKK1 SNP is in LD with several SNPs in the promoter region of the gene. Tissue specific gene expression studies in genotyped subjects should be undertaken to further elucidate the role of variation in these genes on gene expression.

Despite many candidate gene studies in PCOS, few biologically selected candidates have been replicated. The use of expression data from PCOS subjects may result in the identification of novel susceptibility loci. In the present study, we used published expression data to select two putative candidate genes for analysis, DKK1 and DNAJB1. By using a discovery and replication approach we have identified and replicated DNAJB1 as a potential gene important in the insulin resistance of PCOS. While not associated with PCOS itself, genetic variation in DNAJB1 and perhaps DKK1 may act as modifiers, affecting the metabolic and androgenic pathways of PCOS, respectively. In conclusion, gene expression data appears to be a useful source of candidate genes in complex and poorly understood diseases such as PCOS.

**Author Contributions**

Conceived and designed the experiments: Y-DIC RMK JIR RSL RA MOG. Performed the experiments: MRJ AC. Analyzed the data: MRJ AC XL MOG. Wrote the paper: MRJ MOG.
References

1. Goodarzi MO, Aziz R (2006) Diagnosis, epidemiology, and genetics of the polycystic ovary syndrome. Best Pract Res Clin Endocrinol Metab 20: 193–205.
2. Goodarzi MO, Erickson S, Port SC, Jennrich RI, Korresman SG (2005) Beta-cell function: a key pathological determinant in polycystic ovary syndrome. J Clin Endocrinol Metab 90: 310–315.
3. Diamanti-Kandarakis E, Paterakis T, Kandarakis HA (2006) Indices of low-grade inflammation in polycystic ovary syndrome. Ann N Y Acad Sci 1092: 175–186.
4. Wood JR, Nelson VL, Ho C, Jansen E, Wang CY, et al. (2003) The molecular phenotype of polycystic ovary syndrome (PCOS) theca cells and new candidate PCOS genes defined by microarray analysis. J Biol Chem 278: 26380–26390.
5. Corton M, Botella-Carretero JI, Benguria A, Villuendas G, Zaballos A, et al. (2007) Differential gene expression profile in omental adipose tissue in women with polycystic ovary syndrome. J Clin Endocrinol Metab 92: 328–337.
6. Borro M, Gentile G, Sigligiano A, Muzzi S, Toscano V, et al. (2007) Proteomic analysis of peripheral T lymphocytes, suitable circulating biosensors of strictly related diseases. Clin Exp Immunol 150: 494–501.
7. Aguilera O, Fraga MF, Ballestar E, Paz MF, Herranz M, et al. (2006) Epigenetic inactivation of the Wnt antagonist DICKKOPF-1 (DKK-1) gene in human colorectal cancer. Oncogene 25: 4116–4121.
8. Jansen E, Hwang D, Seo S, Kim Y, Kim C, et al. (2007) Microarray analysis of insulin-regulated gene expression in the liver: the use of transgenic mice co-expressing insulin-siRNA and human IDE as an animal model. Int J Mol Med 20: 829–835.
9. Urbaneck M, Legro RS, Driscoll DA, Aziz R, Ehrmann D, et al. (1999) Thirty-seven candidate genes for polycystic ovary syndrome: strongest evidence for linkage is with follistatin. Proc Natl Acad Sci USA 96: 8573–8578.
10. Zawadzki J, Dunaif A (1992) Diagnostic criteria for polycystic ovary syndrome: towards a rational approach. In: Dunaif A, Givens JR, Haseltine FP, Merriam GR, eds. Polycystic Ovary Syndrome. Boston: Blackwell Scientific. pp 377–384.
11. Aziz R, Woods KS, Reyna R, Key TJ, Knochenhauer ES, et al. (2004) The prevalence and features of the polycystic ovary syndrome in an unselected population. J Clin Endocrinol Metab 89: 2745–2749.
12. Legro RS, Driscoll D, Strauss JF, Fox J, Dunaif A (1998) Evidence for a genetic basis for hyperandrogenemia in polycystic ovary syndrome. Proc Natl Acad Sci U S A 95: 14956–14960.
13. Legro RS, Driscoll D, Port SC, Jennrich RI, Kooreman SG (2005) Beta-cell function: a key pathological determinant in polycystic ovary syndrome. J Clin Endocrinol Metab 90: 310–315.
14. Simon JA, Lin F, Huyle SB, Blanche PJ, Waters D, et al. (2006) Phenotypic predictors of response to simvastatin therapy among African-Americans and Caucasians: the Cholesterol and Pharmacogenetics (CAP) Study. Am J Cardiol 97: 843–850.
15. Barrett JC, Fry B, Maller J, Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21: 263–265.
16. Jones MR, Italiano L, Wilson SG, Mullin BH, Mead R, et al. (2006) Polymorphism in HSD17B6 is associated with key features of polycystic ovary syndrome. Fertil Steril 86: 1438–1446.
17. Ke L, Che YN, Cao YX, Xu HX, Yu YL, et al. (2010) Polymorphisms of the HSD17B6 and HSD17B5 genes in Chinese women with polycystic ovary syndrome. J Womens Health (Larchmt) 19: 2227–2232.
18. Jones MR, Mathur R, Cui J, Guo X, Aziz R, et al. (2009) Independent confirmation of association between metabolic phenotypes of polycystic ovary syndrome and variation in the type 6 17beta-hydroxysteroid dehydrogenase gene. J Clin Endocrinol Metab 94: 5058–5063.
19. Mohamed-Hussein ZA, Harun S (2009) Construction of a polycystic ovarian syndrome (PCOS) pathway based on the interactions of PCOS-related proteins retrieved from bibliomic data. Theor Biol Med Model 6: 18.
20. Dakhova O, O’Day D, Kinet N, Vucer N, Wiese M, et al. (2009) Dickkopf-like1 regulates postpubertal spermatocyte apoptosis and testosterone production. Endocrinology 150: 404–412.
21. Kweck MH, Sung YK, Chung EJ, Im SU, Ahn JS, et al. (2008) Dihydrotestosterone-inducible dickkopf 1 from balding dermal papilla cells causes apoptosis in follicular keratinocytes. J Invest Dermatol 128: 262–269.
22. Petris E, Balemans W, Nielsen TL, Andersen M, Boudin E, et al. (2010) Common genetic variation in the DKK1 gene is associated with hip axis length but not with bone mineral density and bone turnover markers in young adult men: results from the Odense Androgen Study. Calcif Tissue Int 86: 271–281.
23. Rosner W, Auchus RJ, Azziz R, Sluss PM, Raff H (2007) Position statement: Utility, limitations, and pitfalls in measuring testosterone: an Endocrine Society position statement. J Clin Endocrinol Metab 92: 405–413.
24. Legro RS, Schlaff WD, Diamond MP, Coutifaris C, Casson PR, et al. (2010) Total testosterone assays in women with polycystic ovary syndrome: precision and correlation with hirsutism. J Clin Endocrinol Metab 95: 5305–5313.
25. Prescott J, Coetzee GA (2006) Molecular chaperones throughout the life cycle of the androgen receptor. Cancer Lett 231: 12–19.