Final Progress Report for Research Projects Funded by Health Research Grants

Instructions: Please complete all of the items as instructed. Do not delete instructions. Do not leave any items blank; responses must be provided for all items. If your response to an item is “None”, please specify “None” as your response. “Not applicable” is not an acceptable response for any of the items. There is no limit to the length of your response to any question. Responses should be single-spaced, no smaller than 12-point type. The report must be completed using MS Word. Submitted reports must be Word documents; they should not be converted to pdf format. Questions? Contact Health Research Program staff at 717-783-2548.

1. **Grantee Institution:** Magee-Womens Research Institute and Foundation

2. **Reporting Period (start and end date of grant award period):** 1/1/13-12/31/13

3. **Grant Contact Person (First Name, M.I., Last Name, Degrees):** Cheryl A Richards, MBA

4. **Grant Contact Person’s Telephone Number:** 412-4-641-8932

5. **Grant SAP Number:** 4100062212

6. **Project Number and Title of Research Project:** Project 6 - Genetic Biomarkers of Inherited and Acquired Forms of Male and Female F Germ-cell Infertility

7. **Start and End Date of Research Project:** 1/1/13-12/31/13

8. **Name of Principal Investigator for the Research Project:** Alexander Yatsenko, MD

9. **Research Project Expenses.**

   9(A) Please provide the total amount of health research grant funds spent on this project for the entire duration of the grant, including indirect costs and any interest earned that was spent:

   \[ \$ 83,686.88 \]

   9(B) Provide the last names (include first initial if multiple individuals with the same last name are listed) of all persons who worked on this research project and were supported with health research funds. Include position titles (Principal Investigator, Graduate Assistant, Post-doctoral Fellow, etc.), percent of effort on project and total health research funds expended for the position. For multiple year projects, if percent of effort varied from year to year, report in the % of Effort column the effort by year 1, 2, 3, etc. of the project (x% Yr 1; z% Yr 2-3).
<table>
<thead>
<tr>
<th>Last Name, First Name</th>
<th>Position Title</th>
<th>% of Effort on Project</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yatsenko, Alexander</td>
<td>Principal Investigator</td>
<td>5%</td>
<td>$3,862</td>
</tr>
<tr>
<td>Zorilla, Michelle</td>
<td>Research Assistant</td>
<td>25%</td>
<td>$11,804</td>
</tr>
<tr>
<td>Kishore, Archana</td>
<td>Temporary Research Assistant</td>
<td>60%</td>
<td>$2,274</td>
</tr>
</tbody>
</table>

9(C) Provide the names of all persons who worked on this research project, but who were not supported with health research funds. Include position titles (Research Assistant, Administrative Assistant, etc.) and percent of effort on project. For multiple year projects, if percent of effort varied from year to year, report in the % of Effort column the effort by year 1, 2, 3, etc. of the project (x% Yr 1; z% Yr 2-3).

<table>
<thead>
<tr>
<th>Last Name, First Name</th>
<th>Position Title</th>
<th>% of Effort on Project</th>
</tr>
</thead>
<tbody>
<tr>
<td>Georgiadis, Andrew</td>
<td>Research Assistant</td>
<td>20%</td>
</tr>
</tbody>
</table>

9(D) Provide a list of all scientific equipment purchased as part of this research grant, a short description of the value (benefit) derived by the institution from this equipment, and the cost of the equipment.

<table>
<thead>
<tr>
<th>Type of Scientific Equipment</th>
<th>Value Derived</th>
<th>Cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

10. Co-funding of Research Project during Health Research Grant Award Period. Did this research project receive funding from any other source during the project period when it was supported by the health research grant?

Yes____X______  No__________

If yes, please indicate the source and amount of other funds:

Magee-Womens Research Institute start-up fund, $10,000

11. Leveraging of Additional Funds

11(A) As a result of the health research funds provided for this research project, were you able to apply for and/or obtain funding from other sources to continue or expand the research?
Yes____X______ No__________

If yes, please list the applications submitted (column A), the funding agency (National Institutes of Health—NIH, or other source in column B), the month and year when the application was submitted (column C), and the amount of funds requested (column D). If you have received a notice that the grant will be funded, please indicate the amount of funds to be awarded (column E). If the grant was not funded, insert “not funded” in column E.

Do not include funding from your own institution or from CURE (tobacco settlement funds). Do not include grants submitted prior to the start date of the grant as shown in Question 2. If you list grants submitted within 1-6 months of the start date of this grant, add a statement below the table indicating how the data/results from this project were used to secure that grant.

<table>
<thead>
<tr>
<th>A. Title of research project on grant application</th>
<th>B. Funding agency (check those that apply)</th>
<th>C. Month and Year Submitted</th>
<th>D. Amount of funds requested:</th>
<th>E. Amount of funds to be awarded:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genomic signatures of X-linked and autosomal candidate-genes in azoospermia</td>
<td>☒ NIH ☐ Other federal (specify:__________) ☐ Nonfederal source (specify: ____________)</td>
<td>10/2013</td>
<td>$274,632</td>
<td>pending</td>
</tr>
<tr>
<td>☐ NIH ☐ Other federal (specify:__________) ☐ Nonfederal source (specify: ____________)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

11(B) Are you planning to apply for additional funding in the future to continue or expand the research?

Yes____X______ No__________

If yes, please describe your plans:

We plan to submit an R21 NIH grant proposal in 2014. It will be concentrated on genomic and functional aspects of azoospermia and male infertility.

12. Future of Research Project. What are the future plans for this research project?

Expand preliminary results and apply for NIH R01 grant during next 1-2 years.
13. **New Investigator Training and Development.** Did students participate in project supported internships or graduate or post-graduate training for at least one semester or one summer?

Yes____X_____ No__________

If yes, how many students? Please specify in the tables below:

<table>
<thead>
<tr>
<th></th>
<th>Undergraduate</th>
<th>Masters</th>
<th>Pre-doc</th>
<th>Post-doc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Undergraduate</th>
<th>Masters</th>
<th>Pre-doc</th>
<th>Post-doc</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hispanic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Undergraduate</th>
<th>Masters</th>
<th>Pre-doc</th>
<th>Post-doc</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

14. **Recruitment of Out-of-State Researchers.** Did you bring researchers into Pennsylvania to carry out this research project?

Yes______ No______X____

If yes, please list the name and degree of each researcher and his/her previous affiliation:

15. **Impact on Research Capacity and Quality.** Did the health research project enhance the quality and/or capacity of research at your institution?

Yes______X____ No______

If yes, describe how improvements in infrastructure, the addition of new investigators, and other resources have led to more and better research.
We have set up a complete process for genomic Comparative Genome Hybridization (CGH) analysis and validated whole exome sequencing protocol in local settings at MWRI. PA state funding enabled us to validate experimental procedures at MWRI and collect initial preliminary results for genomic investigation.

16. **Collaboration, business and community involvement.**

16(A) Did the health research funds lead to collaboration with research partners outside of your institution (e.g., entire university, entire hospital system)?

Yes____X____ No________

If yes, please describe the collaborations:

We established collaboration with Dr. Sanfilippo, director of Reproductive Endocrinology and Infertility center at Magee Womens Hospital, UPMC.

16(B) Did the research project result in commercial development of any research products?

Yes____X____ No________

If yes, please describe commercial development activities that resulted from the research project:

16(C) Did the research lead to new involvement with the community?

Yes____X____ No________

If yes, please describe involvement with community groups that resulted from the research project:

17. **Progress in Achieving Research Goals, Objectives and Aims.**

List the project goals, objectives and specific aims (as contained in the grant agreement). Summarize the progress made in achieving these goals, objectives and aims for the period that the project was funded (i.e., from project start date through end date). Indicate whether or not each goal/objective/aim was achieved; if something was not achieved, note the reasons why. Describe the methods used. If changes were made to the research goals/objectives/aims, methods, design or timeline since the original grant application was submitted, please describe the changes. Provide detailed results of the project. Include evidence of the data that was generated and analyzed, and provide tables, graphs, and figures of the data. List published abstracts, poster presentations and scientific meeting presentations at the end of the summary of progress; peer-reviewed publications should be listed under item 20.
This response should be a DETAILED report of the methods and findings. It is not sufficient to state that the work was completed. Insufficient information may result in an unfavorable performance review, which may jeopardize future funding. If research findings are pending publication you must still include enough detail for the expert peer reviewers to evaluate the progress during the course of the project.

Health research grants funded under the Tobacco Settlement Act will be evaluated via a performance review by an expert panel of researchers and clinicians who will assess project work using this Final Progress Report, all project Annual Reports and the project’s strategic plan. After the final performance review of each project is complete, approximately 12-16 months after the end of the grant, this Final Progress Report, as well as the Final Performance Review Report containing the comments of the expert review panel, and the grantee’s written response to the Final Performance Review Report, will be posted on the CURE Web site.

There is no limit to the length of your response. Responses must be single-spaced below, no smaller than 12-point type. If you cut and paste text from a publication, be sure symbols print properly, e.g., the Greek symbol for alpha (α) and beta (β) should not print as boxes (□) and include the appropriate citation(s). DO NOT DELETE THESE INSTRUCTIONS.
Research Project 6: Project Title and Purpose

Genetic Biomarkers of Inherited and Acquired Forms of Male and Female Germ-cell Infertility – We propose a novel pilot study to investigate genome-wide DNA coding mutations related to male and female infertility using novel oligonucleotide microarray Comparative Genome Hybridization (CGH) and Next Generation Sequencing. We expect to identify a significant number of genomic, genetic, inherited and de-novo germ-cell-specific defects characteristic for infertile male and female germ cells (sperm and oocyte). We hypothesize that these findings will highlight disruption(s) of specific molecular mechanisms of spermatogenesis and oogenesis and would serve as future infertility-associated diagnostic markers of pathological reproductive tissues.

Project Overview
Infertility is defined as the reproductive failure or inability of a couple to conceive in one year. Commonly, male factor etiology is associated with semen deficiencies and female infertility relates to various oocyte defects. Among the most common semen abnormalities observed in infertile males are azoospermia (AZ, absence of male germ cells in semen) and oligozoospermia (OS, low sperm concentration) accounting for ~50% of all cases. The most common ovarian pathologies are polycystic ovarian syndrome, premature ovarian insufficiency/failure, amenorrhea, and ovulatory disorders (reduced or absent ovulation). At least 30% of infertile couples are diagnosed with unknown etiology (aka idiopathic male or female infertility). This reflects our poor understanding of the underlying pathological molecular mechanisms in spermatogenesis, oogenesis, and fertilization. To date, only a small proportion of infertile men and women are diagnosed with definitive genetic defects, while the vast majority of patients have descriptive clinical diagnosis without a known cause and have no medical cure. We propose a pilot genome-wide study to investigate genomic mutations using novel oligonucleotide micro array Comparative Genome Hybridization (aCGH) and Next Generation Sequencing. We anticipate identifying genomic and genetic germ-cell defects characteristic for infertile male and female germ cells (sperm and oocyte). We hypothesize that we will detect a significant number of genomic aberrations, DNA mutations that disrupt vital molecular mechanisms in spermatogenesis and oogenesis that would serve as future biomarkers of infertility. Identification of specific gene defects associated with infertility will help with the future development of novel diagnostic procedures in infertile couples. This research will serve as proof of principle of our abilities to generate and analyze enormous genomic data sets of a CGH and Whole Exome Sequencing technologies.

Principal Investigator

Alexander Yatsenko, MD, PhD
Visiting Assistant Professor
Magee-Womens Research Institute and Foundation and University of Pittsburgh
204 Craft Avenue
Pittsburgh, PA 15213

Other Participating Researchers
Michelle Zorilla, MS – employed by Magee-Womens Research Institute and Foundation
Expected Research Outcomes and Benefits

Infertility is a global health problem that affects 10-15% of American couples. Yet a majority of infertile males receive descriptive diagnosis. Widely used to help infertile couples, assisted reproductive technologies while helping many couples, could introduce underlying genetic causes of infertility to assisted offspring. Thus, it is essential to better understand the underlying genetic etiologies of male and female infertility both for the diagnosis and treatment of the infertile males and females as well as for appropriate genetic counseling. Despite a wealth of evidence from animal models that many autosomal genes could contribute to male and female reproductive defects, current genetic testing is limited. There are karyotype and Y microdeletion and fragile X chromosome analyses used for men and women with clinical descriptive and idiopathic infertility. We propose to apply highly efficient genomic approaches, such as genome-wide microarray CGH analysis of genomic imbalances, and high throughput exome sequencing of all known genes, to comprehensively examine genetic contribution to germ-line infertility and detect mutations with highest outcome available. We have partnered with an internationally renowned Reproductive Endocrinology and Andrology Clinic at Magee Womens Hospital UPMC, to enroll and design a compelling approach to accomplish our aims. This research aims to generate new information that will advance clinical genetic testing and counseling for the infertile couples. We hypothesize that our efforts will generate preliminary information that can be utilized in the future diagnostic genetic testing in couples with infertility. Moreover, our studies are likely to discover novel biologic markers of human male and female reproduction.

Summary

Specific Aim 1. Identify infertility-associated genomic aberrations in AZ and POI patients using oligonucleotide micro array Comparative Genome Hybridization (aCGH) technology.

Specific Aim 1 was achieved.

We focused our efforts on the identification of genomic aberrations in azoospermia and male infertility. We recruited and enrolled a total of 31 patients with male infertility: 19 males with various semen defects, and 12 individuals with normozoospermia. DNA was isolated from whole blood by the Gentra Puregene blood core kit according to the manufacturer’s protocol (Qiagen). DNA sample amount and quality were assessed by NanoDrop 2000 spectrophotometer (Thermo Scientific), Qubit 2.0 fluorometer (Invitrogen), and gel electrophoresis. Genomic DNA was required to be fully intact (~30-50 Kb) and of high purity (260/280 and 260/230 >1.8). High-quality DNA specimens were selected from 10 azoospermic males and 10 normozoospermic males.

With these 20 samples, we performed a high-resolution genome-wide 400K comparative genome hybridization (CGH; Agilent) microarray study to identify genomic abnormalities associated with azoospermia, not present in normozoospermic males. Experimental samples and an equal amount of male control DNA (Promega) were prepared for hybridization by the SureTag DNA labeling kit according to the manufacturer’s protocol (Agilent). DNA was restriction digested and labeled by cyanine 5 (Cy5) for experimental samples and Cy3 for control samples. The hybridization procedure was completed according to the Agilent protocol. Microarray results
were obtained via Agilent SureScan microarray D scanner. Scan quality control was assessed via Agilent’s Feature Extraction for CytoGenomics. Evaluation metrics, including background noise, signal intensity, signal-to-noise ratio, and derivative log ratio, were required to be in the “excellent” range to proceed with analysis. Analysis of array data was performed with Agilent CytoGenomics 2.5.7.0.

CGH data were screened for aberrations by CytoGenomics default Aberration Detection Method 2. The algorithm identifies statistically significant copy number variations (CNVs) by determining the consistency of the log ratio across all probes in the aberration. The quality of probe log ratio, based on its signal-to-noise ratio, is also factored into the significance score. Aberrations were required to have at least three consecutive probes in the CNV region and a minimum absolute log ratio of plus or minus 0.3 in the region of interest. We identified between 18 and 51 CNVs per patient, with an average of 43 CNVs and aberration sizes ranging from 40 kilobases (Kb) to 3 megabases (Mb).

In our preliminary study of 10 azoospermic males we identified a total of 461 CNVs (on average, 46/patient). Manual filtering was performed to identify biologically plausible CNVs. After removing polymorphic CNVs found in our controls and in the Database of Genomic Variants (DGV), a database of polymorphic CNVs and all intergenic variants, we had 370 CNVs in gene regions. We hypothesized that azoospermia is an arrest of spermatogenesis probably caused by conditions stemming from loss-of-function defects. We focused on 211 deletions and removed all duplications. We then removed all variants also found in the 10 normozoospermic samples, leaving 49 CNVs. The remaining genes, with aberrations, were annotated for tissue expression and protein function, reducing our list to 4 CNVs with biological plausibility for male infertility.

The first gene deletion was of ~160 Kb in ADAM3A, located on chromosome 8 (Fig. 2). A piece of neighboring gene ADAM5P was also deleted. Another patient had a small, nearly 90-Kb deletion, on chromosome X that removed a genomic region inside the gene TEX11, including three exons of the gene, exons 8, 9 and 10 (Fig. 3). A third patient had a 12-Kb deletion inside the gene SBF1 on chromosome 22. The last deletion was of 11 Kb and removed a portion of the gene RIF1 on chromosome 2. ADAM3A encodes an important testis-specific metalloproteinase, cyritestin; TEX11 is a protein involved in meiotic division; SFB1 encodes a testis-specific psuedophosphatase, and RIF1 is a homeobox transcription factor. Previously, knock-out models of these genes in male mice were shown to cause spermatogenesis defects, azoospermia, and infertility.

Of special interest is the 90-Kb deletion found in TEX11 that covers 3 exons and removes 79 amino acids from a critical protein domain (SPO11) and probably has a severe effect (Fig. 3). Importantly, Tex11 knockout causes meiotic arrest and azoospermia in male mice. We confirmed the deletion with quantitative polymerase chain reaction (qPCR). Further study of the X-linked TEX11 deletion breakpoint, using an X-chromosome CGH array (Fig. 2), long-range PCR (Fig. 4), and qPCR (Fig. 5), confirmed our finding. Other CNVs were confirmed by qPCR.

To corroborate our findings, we will perform replication mutation screening studies on our growing collection of azoospermic samples and, in collaboration with Dr. Tuettelmann at the
University of Munster, where there is a vast collection of azoospermic samples. **Such replication studies are beyond the financial scope of the current project.** Effects of novel, potentially pathologic CNVs can be tested by generating transgenic mouse models of identified genes. Mice are excellent models for azoospermia, and such models will be of great utility in coming to know the mechanisms of testicular failure and in identifying targets for future therapeutic interventions to reverse sperm loss in human azoospermia. **Animal transgenic studies are beyond the financial scope of the current project.**

**Materials and Methods**

**Array Comparative Genomic Hybridization**

We performed an array CGH with DNAs from 20 patients and with Promega male DNA as a control (Promega, Fitchburg, WI). The study was carried out according to the oligonucleotide array CGH protocol (Agilent, Santa Clara, CA). Genomic DNA was digested by Alu I and Rsa I, and samples were fluorescently labeled with Cy5-deoxyuridine triphosphate (dUTP) (experimental) or Cy3-dUTP (control) by the SureTag DNA labeling kit (Agilent). Labeled samples were purified by SureTag labeling kit columns, and matched amounts of experimental and control samples were hybridized to SurePrint G3 human CGH 2x400K oligo microarrays and in-house-designed 4x180K X chromosome-specific oligo microarrays (Agilent). Hybridized slides were read by a SureScan microarray scanner (Agilent). Microarray CGH data were analyzed using Agilent CytoGenomics 2.5.7.0. CNVs previously described in the DGV were removed from further consideration. Data was filtered for significant CNVs using the National Center for Biotechnology Information databases Online Mendelian Inheritance in Mammals (OMIM), UniGene, and the Conserved Domain Database, as well as BioGPS (www.biogps.org).

**Quantitative PCR, PCR, Long-Range PCR and Sanger DNA Sequencing**

qPCR was performed with 20 ng of DNA and iQ SYBR Green Supermix (Bio-Rad Laboratories, Hercules, CA). To confirm the deletion of exons 8-10 in patients P67 and P89, we created primer sets exclusively within each TEX11 exons 8-10 and in control Beta-actin (ACTB). We calculated fold change by the ΔΔCt method. \( \Delta Ct \) for each sample was calculated as: \( \Delta Ct = Ct_{\text{experimental \, TEX11 \, exon}} - Ct_{\text{ACTB}} \). \( \Delta \Delta Ct \) for each experimental sample was calculated as: \( \Delta \Delta Ct = \Delta Ct_{\text{experimental \, sample}} - \Delta Ct_{\text{control \, sample}} \). Fold change was calculated as: fold change = \( 2^{(-\Delta \Delta Ct)} \). Long range PCR was performed with 200 ng of DNA and TaKaRa LA PCR kit v2.1 (Clontech, Mountain View, CA). We used the following forward and reverse primers, TCTGTCCGAAAAGTCACATATCTCTGTTTCTG: TATACAGTTTGCTATGGACCGAATGTTTGTGTC, to identify the length and intronic breakpoint locations of the exon 8-10 deletion in patients P67 and P89. PCR was performed with 15 ng of DNA and KAPA HiFi HotStart DNA polymerase (KAPA Biosystems, Woburn, MA). A total of 30 primer sets were used to amplify TEX11 exons 1a, 1b, and 2-29, including at least 50 base pairs (bp) of flanking intronic regions (between 274-632 bp, average 409 bp). PCR products were sequenced by an ABI Prism 3130XL sequencer, using BigDye V3.1 reagent (Applied Biosystems, Foster City, CA). Sanger sequencing data was aligned and screened for mutations using Sequencher DNA sequencing software (Gene Codes, Ann Arbor, MI).

From our preliminary data, we expect pathogenic azoospermic variants to be rare in the control population (incidence of genetic azoospermic in men under age 40 is less than 0.5%). It is therefore unlikely that we will observe most azoospermia-associated CNVs in the controls.
Entries in the DGV from 10,000 “healthy” individuals will also help to eliminate benign CNVs, as will records in the International Standards for Cytogenomic Arrays (ISCA) database. Small CNVs found only in azoospermia will be validated using quantitative real-time PCR or fluorescence in situ hybridization (FISH). We have access to the complete library of FISH probes in the Pittsburgh Cytogenetics Lab. In cases of sporadic findings (i.e., de novo events), if family members are available, we will perform segregation and non-paternity analyses.

Identified variants are sorted into five types: definitely pathogenic, likely pathogenic, variants of unknown significance, likely benign, and benign variants (see the American College of Medical Genetics guidelines). To further prioritize CNVs, we use additional information, such as the effect of whole-gene or large deletion versus a small deletion of an exon or regulatory region; the gene’s expression in testis; or the variant’s presence in the control population. We will focus on the first two pathogenic categories as being the most significant, while benign CNVs and those of unknown significance will be of lesser priority. Among CNVs, we give priority to deletions, as the functional effect of duplications is more difficult to interpret. We will not discard duplications given that they can cause well-known human syndromes. CNVs that are present in more than one patient and segregate with the azoospermic phenotype (present in other patients, absent in unaffected siblings and controls, de novo, or present in fathers) will be ranked highest. We will identify polymorphic CNVs using the in-house database of genomic variants that we will build using 200 controls, integrated also with benign CNVs from the ISCA database and the DGV, derived from high-resolution whole-genome tiling CGH array (42 million probes) on 450 HapMap individuals. Currently, we have incorporated 461 CNVs (500 bp or greater) into our analysis and have had no difficulty in the preliminary interpretation of array CGH results in 10 azoospermic men.

**Specific Aim 2.** Identify infertility-associated mutations and candidate genes by a genome-wide screen with Next Generation DNA Sequencing (Whole Exome).

Specific Aim 2 was achieved.

We concentrated our efforts on the identification of specific nucleotide variants in azoospermia and male infertility. From collected specimens, we selected eight DNA samples from two families with azoospermia. DNA was isolated and the quality tested by the same techniques as in Specific Aim 1. Family 1 had two infertile males and their parents; Family 2 consisted of one infertile male, his fertile brother, and their parents (Fig. 1).

We performed a standard human exome capture procedure using Agilent exome capture library V4.2 + UTRs, which covers 71 Mb of coding and flanking sequences and enriches coding regions. Whole exome sequencing (WES) was run on a HiSeq 2000 machine (Illumina). Sanger sequencing was used to confirm positive WES findings. Bioinformatic analysis was performed with open access and commercial NextGENe software (SoftGenetics). To rule out genomic aberrations and test for homozygosity, we ran the 400K CGH (Agilent) and 700K SNP (Illumina) arrays, respectively.

For each family member in each of the two families, we obtained on average ~6.0 Gb of sequencing results. The average exome coverage was 62X. Nearly 94% of the targeted exome sequence had 10X coverage or higher, with Phred-like consensus quality of 30. Phred quality >9 (chance of error = 0.001) was used to remove sequencing run errors.
Family AZ-67. We obtained approximately 6.0 Gb of whole exome sequence for a father, mother, and affected sons. Approximately 96% of the targeted exome regions were covered more than 10X, overall, with an average coverage of 52X. Good quality variant calls, indicated via Phred-like read quality, are ≥9. After read alignment to the genome, we detected a total of 69,329 and 69,095 variants (single-nucleotide polymorphisms [SNPs], deletions, duplications) in the two affected Sertoli cell-only syndrome (SCOS) brothers. Considering autosomal recessive or X-linked inheritance modes as most plausible in the family pedigree, we reduced the number of shared variants between the two affected brothers to 10,152 homozygous mutations, and 11,777 compound heterozygotes.

Out of these variants, 1,666 homozygous mutations had non-synonymous substitutions in the coding regions, and 1,340 were compound heterozygous for non-synonymous substitutions. These variants were filtered against known SNPs from the dbSNP135 database, and variants with a frequency >1% were excluded. Phred quality >9 (chance of error is 0.001) was used to remove sequencing run errors. We did not test the deleterious effect of variants and conservation due to strong elimination of genes. The next filter used was testis expression, which led to the identification of 21 deleterious novel sequence variants in a total of 10 genes.

Using bioinformatic filtering for autosomal recessive and X-linked inheritance models, we retained the 1,666 homozygous and 1,340 compound heterozygous non-synonymous variants in exons or flanking intronic regions. These variants were filtered against known SNPs in the dbSNP135 database, and SNPs with a population frequency >1% were excluded, leaving 569 novel variants. Further filtering, using male reproductive tissue-specific expression and protein function, led to a total of 13 novel homozygous, compound heterozygous, and hemizygous sequence variants (Table 1). Preliminary analysis indicates that these 13 selected azoospermia candidate genes are likely to have deleterious effects.

Family AZ-45. In the infertile male in Family 2, we identified ~57,000 SNPs and indels (Fig. 6). In this family we considered autosomal recessive or X-linked inheritance. In this family, we had an unaffected brother, which allowed us to remove all shared variants between brothers, leaving 8,437 homozygous mutations, and 8,823 compound heterozygotes.

Using bioinformatic filtering for autosomal recessive and X-linked inheritance models, we retained 1,259 homozygous and 2,140 compound heterozygous non-synonymous variants in exons or flanking intronic regions. These variants were filtered against known SNPs in the dbSNP135 database, and SNPs with a population frequency >1% were excluded, leaving 411 novel variants. Further filtering, using male reproductive tissue-specific expression and protein function, led to a total of nine novel homozygous, compound heterozygous, and hemizygous sequence variants. Preliminary analysis indicates that seven azoospermia candidate genes are likely to have deleterious effects (Table 2).

To corroborate our findings, we will perform replication mutation screening studies on our collection of azoospermic samples and, in collaboration with Dr. Tuettelmann at the University of Munster, where there is a vast collection of azoospermic samples. Such replication studies are beyond the financial scope of the current project. Loss of function in certain genes will be tested by generating transgenic mouse models of novel gene-candidates. Mice are excellent models for azoospermia, and such models will be of great utility in coming to know the mechanisms of testicular failure and in identifying targets for future therapeutic interventions to reverse sperm loss in human azoospermia. Animal transgenic studies are beyond the financial scope of the current project.
Materials and Methods
Whole exome capture and sequencing
We enriched genomic DNA with the whole exome capture and comprehensively sequenced all exons and exon-intron junctions via a next-generation sequencing (NGS) protocol that successfully captures the exome exonic mutations. Capture was carried out using exome capture “baits” in the SureSelect 71Mb (Agilent) solution (120-nucleotide-long oligos that complement targeted regions in the genome) that bind to the coding exons of the vast majority of the human exome. Only 3 μg of genomic DNA was needed for exome capture with the SureSelect exome capture system (Agilent). We routinely obtain 100 μg of genomic DNA from 5 mL of peripheral blood. We carried out the NGS run at the MWRI Genomics Laboratory (PI, Dr. David Peters). The captured exome was sequenced using the Illumina HiSeq HD 2000 System, a platform that enables massive parallel sequencing of DNA fragments linked to beads. One sequencing flow cell is divided into eight flow channels and yields up to 100 Gb of sequence per run. Exomes captured with SureSelect baits will be transferred directly into emulsion PCR for sequencing. Samples were portioned on the Illumina flow cell to achieve an average of 50X sequencing coverage, which increases accurate detection.

Recovery of Human Exome Variants
Quality control of sequencing includes removing short reads (30 bp or less) and unaligned reads which do not match the reference genomic sequence and corresponding human exons. Our analytic platform is NextGENe (Softgenetics), which aligns, assembles, and annotates sequence and calls SNPs and small deletion/insertion polymorphisms (DIPs), the latter ranging from 1 to 10 base pairs. The program also integrates data with online databases. We will utilize the NGS quality control pipeline as recently described. Sequencing data was processed, aligned, and mapped to the human genome with the help of Dr. Michael Barmada, the Director of the Center for Computational Genetics at the University of Pittsburgh. SNP and DIP variants were filtered through public sequence databases, such as dbSNP, the HapMap project, and the 1000 Genomes project. We systematically investigated candidate genes that have sequence variants with minor allele frequencies less than 1%. Variants with an allele frequency greater than 1% were eliminated from consideration as not likely to be disease causing, given the relatively low prevalence of SCOS. Based on recent publications, we compared an estimated number of coding SNPs to the number retrieved from each patient. Given that our chip covered 71 Mb of coding and flanking sequences, we would recover ~1,980 coding SNPs per patient, including novel, non-coding SNPs and cryptic splicing defects. We would filter ~90-94% of these against dbSNP, getting ~120-200 “novel” SNPs per patient, and ~50% of them (60-100) would be non-synonymous SNPs. Nearly 25-30% of these SNPs will be deleterious to protein function, resulting in up to 15-30 deleterious sequence variants or mutations per patient.

Figure 1. Two families that have male subjects with azoospermia (AZ) and male infertility. Family AZ-67 is already enrolled, screened via array CGH and being sequenced by NGS. Preliminary results of NGS sequencing for this family are discussed. Family AZ-45 is being enrolled.
Figure 2. CGH analysis for two selected azoospermic patients with male infertility. 
A. Homozygous ~50-Kb deletion that removes more than half of ADAM3A gene. 
B. Nearly 90-Kb deletion of Tex11 X-linked gene affecting three coding exons of the gene.
Figure 3. Hemizygous deletion of *TEX11* exons 8-10 and flanking intronic regions in two azoospermic males. **A.** Comparative genomic hybridization with ~400K array, showing a deletion on the X13.2 region of the X chromosome, highlighted in blue. **B.** The deletion includes 12 probes in the gene *TEX11*, spanning about 90 kb. **C, D.** A custom X chromosome array with ~180K probes further defined the deletion to 91,042 (C, P67) and 92,064 bp (D, P89), encompassing exons 8-10. In sample P67, the deletion covers chrX: 69,954,488-70,045,530 (hg19), and in P89 the deletion covers chrX:69,954,488-70,046,552 (hg19).

Figure 4. Genomic structure of the *TEX11* deleted region defined by long-range PCR and sequence analysis. The deletion covering exons 8-10 is shown in red. Alu repeats are shown in dark blue. Green boxes, ex8, ex9, and ex10, abbreviate *TEX11* exons 8-10. 5’ and 3’ breakpoints and genomic coordinates are displayed above and below the diagram.
**Figure 5.** qPCR analysis of deletion exons 8-10 of *TEX11*. Fold change of four azoospermic samples and Promega male control, calculated by delta Ct, with Actin beta serving as an internal control. P67 and positive control showed a fold change of ~0, with no amplification of *TEX11* exons 8-10. Azoospermic samples P1 and P46 and normal Promega male DNA serving as controls display a fold change of 1.19 +/- 0.12. Exon 8 is represented by a black bar, exon 9 by a grey bar, and exon 10 by a white bar.

**Figure 6.** Bioinformatic filtering of exome sequencing results for family AZ-45, an azoospermic brother, normozoospermic brother, and their parents. Two modes of inheritance were considered, autosomal recessive and X-linked.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Chr</th>
<th>Exon</th>
<th>DNA Mutation</th>
<th>Protein Change, aa</th>
<th>M Expr</th>
<th>H Expr</th>
<th>Mouse Study</th>
<th>Protein Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACRC</td>
<td>X</td>
<td>7</td>
<td>G&gt;T</td>
<td>R441S</td>
<td>brain</td>
<td>testis, nerve</td>
<td>-</td>
<td>may play a role in chromatin structure</td>
</tr>
<tr>
<td>APLF</td>
<td>2</td>
<td>7</td>
<td>delGAA</td>
<td>R314del1aa</td>
<td>bone</td>
<td>thyroid</td>
<td>abnormal DNA repair</td>
<td>DNA break repair</td>
</tr>
<tr>
<td>CYP4F8</td>
<td>19</td>
<td>9</td>
<td>insC</td>
<td>V151X, PTC</td>
<td>prostate, seminal vesicles</td>
<td>-</td>
<td>-</td>
<td>catalyzes reactions involved in lipid/drug metabolism synthesis</td>
</tr>
<tr>
<td>DNAJC28</td>
<td>21</td>
<td>1</td>
<td>delTTAAA</td>
<td>F314fs, PTC</td>
<td>testis high</td>
<td>testis high</td>
<td>null male mice, fertile</td>
<td>protein folding, chaperone</td>
</tr>
<tr>
<td>JRK</td>
<td>8</td>
<td>1</td>
<td>delCA</td>
<td>V461fs, PTC</td>
<td>testis highest</td>
<td>testis average blood, muscle</td>
<td>DNA binding protein</td>
<td></td>
</tr>
<tr>
<td>KIR3DL1</td>
<td>19</td>
<td>3</td>
<td>C&gt;T</td>
<td>S107L</td>
<td>-</td>
<td>blood, muscle</td>
<td>-</td>
<td>transmembrane glycoproteins</td>
</tr>
<tr>
<td>NINL</td>
<td>20</td>
<td>16</td>
<td>delCCCAC</td>
<td>E959fs, PTC</td>
<td>testis low</td>
<td>testis low</td>
<td>overexpressed male mice, fertile</td>
<td>microtubule organization interphase cells</td>
</tr>
<tr>
<td>PABPC1</td>
<td>8</td>
<td>-</td>
<td>comp hetero</td>
<td>testis high</td>
<td>testis high</td>
<td>testis high</td>
<td>-</td>
<td>poly(A) binding protein</td>
</tr>
<tr>
<td>PTCHD3</td>
<td>10</td>
<td>1</td>
<td>insC</td>
<td>L306fs, PTC</td>
<td>testis</td>
<td>testis</td>
<td>pachytene stage spermatogenesis</td>
<td>regulates sperm motility</td>
</tr>
<tr>
<td>SPANXD</td>
<td>X</td>
<td>2</td>
<td>G&gt;T, nonsense</td>
<td>S48X</td>
<td>testis only</td>
<td>testis only</td>
<td>-</td>
<td>encodes a sperm protein associated with the nucleus, role in spermatogenesis</td>
</tr>
<tr>
<td>SRD5A2</td>
<td>2</td>
<td>1</td>
<td>insG</td>
<td>K29fs, PTC</td>
<td>male reprod tissues</td>
<td>male reprod tissues</td>
<td>null mutant males fertile, have small prostates and seminal vesicles</td>
<td>converts testosterone into dihydrotestosterone</td>
</tr>
<tr>
<td>TAF7L</td>
<td>X</td>
<td>10</td>
<td>delCTCATC</td>
<td>D340del2aa</td>
<td>placenta, germ cells</td>
<td>testis, germ cells</td>
<td>decreased sperm production, decreased fertility</td>
<td>pre-meiotic spermatogenesis, RNA polymerase</td>
</tr>
<tr>
<td>TPTE2</td>
<td>13</td>
<td>-</td>
<td>comp hetero</td>
<td>testis average</td>
<td>testis highest</td>
<td>testis highest</td>
<td>-</td>
<td>membrane-associated phosphatase</td>
</tr>
</tbody>
</table>
Table 2. Final summary of variants/mutations identified in two affected brothers with azoospermia and male infertility from family AZ-45. These variants have the highest likelihood of having a damaging effect on spermatogenesis and male reproductive function. Chr abbreviates chromosome; aa is amino acid; ESP6500 is NIH based database of SNPs in normal population; Mouse model shows knock-out mouse model phenotype related to male infertility; Protein Function is the known role of encoded protein.

<table>
<thead>
<tr>
<th>#</th>
<th>Gene</th>
<th>Chr</th>
<th>Location</th>
<th>DNA change</th>
<th>Protein change</th>
<th>Population% (ESP6500)</th>
<th>Mouse models</th>
<th>Protein Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>ATM</td>
<td>11</td>
<td>108199929</td>
<td>c.T7271G</td>
<td>p.V2424G</td>
<td>0.0077</td>
<td>abnormal male reproductive system</td>
<td>PI3/P4- cell cycle checkpoint kinase</td>
</tr>
<tr>
<td>3</td>
<td>C19orf21</td>
<td>19</td>
<td>758613</td>
<td>c.G1667A</td>
<td>p.R556H</td>
<td>0.5075</td>
<td>NA</td>
<td>cell division mitosis</td>
</tr>
<tr>
<td>5</td>
<td>DHX33</td>
<td>17</td>
<td>5347788</td>
<td>c.C1342T</td>
<td>p.R448C</td>
<td>1.7453</td>
<td>NA</td>
<td>putative DEAD box RNA helicase involved in cell division, growth embryogenesis, spermatogenesis</td>
</tr>
<tr>
<td>6</td>
<td>ETV4</td>
<td>17</td>
<td>41606033</td>
<td>c.C478T</td>
<td>p.R160C</td>
<td>0.8842</td>
<td>abnormal ejaculation male infertility</td>
<td>-</td>
</tr>
<tr>
<td>7</td>
<td>FOXN1</td>
<td>17</td>
<td>26861409</td>
<td>G988A</td>
<td>p.E330K</td>
<td>0</td>
<td>reduced male fertility</td>
<td>correlated with T-cell immunodeficiency</td>
</tr>
<tr>
<td>8</td>
<td>IFT81</td>
<td>12</td>
<td>110566907</td>
<td>c.A401G</td>
<td>p.D134G</td>
<td>1.4455</td>
<td>NA</td>
<td>intraflagellar transport 81 homolog (Chlamydomonas) highly expressed in testis</td>
</tr>
<tr>
<td>9</td>
<td>KIAA0408</td>
<td>6</td>
<td>127767954</td>
<td>c.C1510T</td>
<td>p.P504S</td>
<td>0.7381</td>
<td>NA</td>
<td>interacts with TEX11</td>
</tr>
<tr>
<td>10</td>
<td>KIAA1217</td>
<td>10</td>
<td>24762833</td>
<td>c.G1523A</td>
<td>p.R508H</td>
<td>0.4229</td>
<td>male infertility</td>
<td>interacts with TEX11</td>
</tr>
<tr>
<td>11</td>
<td>SIK3</td>
<td>11</td>
<td>116729161</td>
<td>c.A2702G</td>
<td>p.Y901C</td>
<td>0.2924</td>
<td>NA</td>
<td>ovarian cancer</td>
</tr>
<tr>
<td>12</td>
<td>SZT2</td>
<td>1</td>
<td>43897496</td>
<td>c.G5027A</td>
<td>p.R1676H</td>
<td>1.0995</td>
<td>NA</td>
<td>resistance to oxidative stress</td>
</tr>
</tbody>
</table>
18. Extent of Clinical Activities Initiated and Completed. Items 18(A) and 18(B) should be completed for all research projects. If the project was restricted to secondary analysis of clinical data or data analysis of clinical research, then responses to 18(A) and 18(B) should be “No.”

18(A) Did you initiate a study that involved the testing of treatment, prevention or diagnostic procedures on human subjects?

___X___ Yes
___ ___ No

18(B) Did you complete a study that involved the testing of treatment, prevention or diagnostic procedures on human subjects?

______Yes
___X___ No

If “Yes” to either 18(A) or 18(B), items 18(C) – (F) must also be completed. (Do NOT complete 18(C-F) if 18(A) and 18(B) are both “No.”)

18(C) How many hospital and health care professionals were involved in the research project?

___0___ Number of hospital and health care professionals involved in the research project

18(D) How many subjects were included in the study compared to targeted goals?

___100___ Number of subjects originally targeted to be included in the study
___31___ Number of subjects enrolled in the study

Note: Studies that fall dramatically short on recruitment are encouraged to provide the details of their recruitment efforts in Item 17, Progress in Achieving Research Goals, Objectives and Aims. For example, the number of eligible subjects approached, the number that refused to participate and the reasons for refusal. Without this information it is difficult to discern whether eligibility criteria were too restrictive or the study simply did not appeal to subjects.

We overestimated the patient volume in Pittsburgh area. There is approximately half the volume of real patients that can be enrolled. In addition, male and female patients are seen in two separate UPMC clinics and we were unable to fund the position for second genetic counselor to enroll female patients in female infertility clinic.

18(E) How many subjects were enrolled in the study by gender, ethnicity and race?

Gender:
18(F) Where was the research study conducted? (List the county where the research study was conducted. If the treatment, prevention and diagnostic tests were offered in more than one county, list all of the counties where the research study was conducted.)

Magee Womens Hospital, Pittsburgh, Allegheny County, PA

19. **Human Embryonic Stem Cell Research.** Item 19(A) should be completed for all research projects. If the research project involved human embryonic stem cells, items 19(B) and 19(C) must also be completed.

19(A) Did this project involve, in any capacity, human embryonic stem cells?
   - Yes
   - X No

19(B) Were these stem cell lines NIH-approved lines that were derived outside of Pennsylvania?
   - Yes
   - No

19(C) Please describe how this project involved human embryonic stem cells:

20. **Articles Submitted to Peer-Reviewed Publications.**

20(A) Identify all publications that resulted from the research performed during the funding period and that have been submitted to peer-reviewed publications. Do not list journal abstracts or presentations at professional meetings; abstract and meeting presentations should
be listed at the end of item 17. **Include only those publications that acknowledge the Pennsylvania Department of Health as a funding source** (as required in the grant agreement). List the title of the journal article, the authors, the name of the peer-reviewed publication, the month and year when it was submitted, and the status of publication (submitted for publication, accepted for publication or published.). Submit an electronic copy of each publication or paper submitted for publication, listed in the table, in a PDF version 5.0.5 (or greater) format, 1,200 dpi. Filenames for each publication should include the number of the research project, the last name of the PI, and an abbreviated title of the publication. For example, if you submit two publications for Smith (PI for Project 01), one publication for Zhang (PI for Project 03), and one publication for Bates (PI for Project 04), the filenames would be:

- Project 01 – Smith – Three cases of isolated
- Project 01 – Smith – Investigation of NEB1 deletions
- Project 03 – Zhang – Molecular profiling of aromatase
- Project 04 – Bates – Neonatal intensive care

If the publication is not available electronically, provide 5 paper copies of the publication.

**Note:** The grant agreement requires that recipients acknowledge the Pennsylvania Department of Health funding in all publications. Please ensure that all publications listed acknowledge the Department of Health funding. If a publication does not acknowledge the funding from the Commonwealth, do not list the publication.

<table>
<thead>
<tr>
<th>Title of Journal Article:</th>
<th>Authors:</th>
<th>Name of Peer-reviewed Publication:</th>
<th>Month and Year Submitted:</th>
<th>Publication Status (check appropriate box below):</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.None</td>
<td></td>
<td></td>
<td></td>
<td>Submitted ☐ Accepted ☐ Published ☐</td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
<td></td>
<td>Submitted ☐ Accepted ☐ Published ☐</td>
</tr>
</tbody>
</table>

20(B) Based on this project, are you planning to submit articles to peer-reviewed publications in the future?

**Yes** ☒ **X** ☐ **No** ☐

If yes, please describe your plans:
We are in the process of preparation of one manuscript related to array CGH screening in azoospermic infertile males and one manuscript related to whole exome sequencing. We plan to perform and publish results of replication CGH and WES studies as well.

Describe the outcome, impact, and effectiveness of the research project by summarizing its impact on the incidence of disease, death from disease, stage of disease at time of diagnosis, or other relevant measures of outcome, impact or effectiveness of the research project. If there were no changes, insert “None”; do not use “Not applicable.” Responses must be single-spaced below, and no smaller than 12-point type. DO NOT DELETE THESE INSTRUCTIONS. There is no limit to the length of your response.

Our results after validation in larger cohorts of azoospermic male patients will improve existing diagnostic procedure of underlying genetic causes of male infertility.

22. Major Discoveries, New Drugs, and New Approaches for Prevention Diagnosis and Treatment. Describe major discoveries, new drugs, and new approaches for prevention, diagnosis and treatment that are attributable to the completed research project. If there were no major discoveries, drugs or approaches, insert “None”; do not use “Not applicable.” Responses must be single-spaced below, and no smaller than 12-point type. DO NOT DELETE THESE INSTRUCTIONS. There is no limit to the length of your response.

When validated, new CGH and WES genomic protocols will allow the performance of a one-time all inclusive genetic test that will predict fertility outcome. The new CGH and WES genomic protocols will likely be implemented in clinical practice of diagnosis and treatment of genetic forms of male and female infertility.

23. Inventions, Patents and Commercial Development Opportunities.

23(A) Were any inventions, which may be patentable or otherwise protectable under Title 35 of the United States Code, conceived or first actually reduced to practice in the performance of work under this health research grant? Yes_______ No____ X____

If “Yes” to 23(A), complete items a – g below for each invention. (Do NOT complete items a - g if 23(A) is “No.”)

a. Title of Invention:

b. Name of Inventor(s):

c. Technical Description of Invention (describe nature, purpose, operation and physical, chemical, biological or electrical characteristics of the invention):
d. Was a patent filed for the invention conceived or first actually reduced to practice in the performance of work under this health research grant?
   Yes_______ No_____
   If yes, indicate date patent was filed:

e. Was a patent issued for the invention conceived or first actually reduced to practice in the performance of work under this health research grant?
   Yes_______ No_____
   If yes, indicate number of patent, title and date issued:
   Patent number:
   Title of patent:
   Date issued:

f. Were any licenses granted for the patent obtained as a result of work performed under this health research grant? Yes_______ No_____
   If yes, how many licenses were granted?___________

g. Were any commercial development activities taken to develop the invention into a commercial product or service for manufacture or sale? Yes_____ No_____ 
   If yes, describe the commercial development activities:

23(B) Based on the results of this project, are you planning to file for any licenses or patents, or undertake any commercial development opportunities in the future?
   Yes_______ No____X____
   If yes, please describe your plans:

24. **Key Investigator Qualifications.** Briefly describe the education, research interests and experience and professional commitments of the Principal Investigator and all other key investigators. In place of narrative you may insert the NIH biosketch form here; however, please limit each biosketch to 1-2 pages. *For Nonformula grants only – include information for only those key investigators whose biosketches were not included in the original grant application.*
BIOGRAPHICAL SKETCH

NAME
Alexander N. Yatsenko, M.D., Ph.D.

POSITION TITLE
Assistant Professor

eRA COMMONS USER NAME: yatsenko

EDUCATION / TRAINING (Begin with baccalaureate or other initial professional education, such as nursing, and include postdoctoral training.)

<table>
<thead>
<tr>
<th>INSTITUTION AND LOCATION</th>
<th>DEGREE (if applicable)</th>
<th>YEAR(s)</th>
<th>FIELD OF STUDY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Russian State Medical University, Moscow, Russia</td>
<td>B.S.</td>
<td>1990</td>
<td>Biochemistry</td>
</tr>
<tr>
<td>Russian State Medical University, Moscow, Russia</td>
<td>M.D.</td>
<td>1993</td>
<td>Medicine</td>
</tr>
<tr>
<td>Medical Genetics Research Center, Moscow, Russia</td>
<td>Ph.D.</td>
<td>1996</td>
<td>Genetics</td>
</tr>
<tr>
<td>Baylor College of Medicine, Mol. Human Genetics</td>
<td>Post Doc</td>
<td>1999 – 2004</td>
<td>Molecular Genetics</td>
</tr>
<tr>
<td>Baylor College of Medicine, Pathology, Houston, TX</td>
<td>Post Doc</td>
<td>2004 – 2007</td>
<td>Reproductive Genetics</td>
</tr>
</tbody>
</table>

A. Personal Statement
I have more than a decade of experience deciphering the effects of DNA and mRNA mutations on proteins and cell function. Also, I am ABMG-board certified in clinical molecular genetics. My scientific focuses on the genetics of male infertility, particularly oligozoospermia and azoospermia. I hypothesized that genetic defects play an important role in idiopathic human male infertility. Using modern genetic approaches we identified genetic defects that associate with semen deficiencies in infertile males; we have discovered that KLHL10 and UBE2B genes responsible for oligozoospermia semen defects and male infertility, and ZPBP1 associated with teratozoospermia. The international clinical research database “Faculty of 1000 Medicine”, selected my 2006 publication in Molecular Human Genetics (publication #1) as highly influential in clinical research. I have also participated in the discovery of chromosome aberrations role in semen deficiencies. My laboratory performs clinical studies collecting and genotyping semen and blood samples from infertile men. We propose that a spectrum of genomic aberrations and single nucleotide mutations is associated with azoospermia and male infertility. Using high-throughput genomic approaches (genomic arrays and Exome sequencing) we will identify micro aberrations and nucleotide mutations in all coding exons in azoospermic males. These studies will bring us to unprecedented scale of information and we will perform combined whole genome association study of the genetic defects with azoospermia and male infertility. Ultimately, the study will improve our knowledge of the genetics of male infertility and will help to translate it into the clinical diagnostics.

B. Positions

1996-97 Scientist, Laboratory of Clinical Cytogenetics, Medical Genetics Research Center, Moscow, Russia.
1997-99 Scientist, Laboratory of Medical Biotechnology, Institute of Biomedical Chemistry, Moscow,
1999-04 Postdoctoral Associate, Department of Molecular & Human Genetics, Baylor College of Medicine, Houston, TX.
2004-07 Research Associate, Department of Pathology, Baylor College of Medicine, Houston, TX
2007-09 Clinical Fellow, ABMG fellowship in Clinical Molecular Genetics, Department of Molecular & Human Genetics, Baylor College of Medicine, Houston, TX.
2009-10  Research Associate, Department of Pathology, Baylor College of Medicine, Houston, TX
2010  Assistant Professor, Pathology & Immunology, Molecular & Human Genetics, Baylor College of Medicine, Houston, TX
2011  Assistant Professor, Magee-Womens Research Institute and Department of OBGY N and Reproductive Science at the University of Pittsburgh School of Medicine
2011  Assistant Professor (Secondary Appointment) Department of Pathology, University of Pittsburgh School of Medicine

C. Selected Peer-reviewed Publications (Selected from 19 peer-reviewed publications)

D. Research Support
**Ongoing research support**
5 K08 HD058073-04 (Yatsenko AN) 05/06/2010 - 04/30/2015
NIH/NICHD $241,674 DC
Genetic Basis of Oligozoospermia in Infertile Men

This study proposes to identify specific gene mutations that associate with oligozoospermia that would lead to the development of novel diagnostic procedures and will help with prognosis in assisted reproductive medicine. Since oligozoospermia is a makor clinical category that accounts for nearly 30% of all semen abnormalities, the test for mutations responsible for oligozoospermia in humans will be a major improvement to health care practice of reproductive and fetal defects. Role: PI