

Award Number: W81XWH-11-1-0737

TITLE: Role of Mitochondrial Inheritance on Prostate Cancer Outcome in African American Men

PRINCIPAL INVESTIGATOR: Martha Mims

CONTRACTING ORGANIZATION: Baylor College of Medicine  
Houston, TX 77030-3411

REPORT DATE: November 2016

TYPE OF REPORT: Addendum to Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command  
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;  
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

# REPORT DOCUMENTATION PAGE

*Form Approved*  
*OMB No. 0704-0188*

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

<b>1. REPORT DATE</b> November 2016			<b>2. REPORT TYPE</b> Addendum to Final			<b>3. DATES COVERED</b> 1Oct2016 - 31Aug2016		
<b>4. TITLE AND SUBTITLE</b> Role of Mitochondrial Inheritance on Prostate Cancer Outcome in African American Men						<b>5a. CONTRACT NUMBER</b>		
						<b>5b. GRANT NUMBER</b> W81XWH-11-1-0737		
						<b>5c. PROGRAM ELEMENT NUMBER</b>		
<b>6. AUTHOR(S)</b> Martha Mims  E-Mail: mmims@bcm.edu						<b>5d. PROJECT NUMBER</b>		
						<b>5e. TASK NUMBER</b>		
						<b>5f. WORK UNIT NUMBER</b>		
<b>7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)</b>  Baylor College of Medicine One Baylor Plaza T100 Houston, Texas 77030-3498						<b>8. PERFORMING ORGANIZATION REPORT NUMBER</b>		
<b>9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)</b>  U.S. Army Medical Research and Materiel Command Fort Detrick, Maryland 21702-5012						<b>10. SPONSOR/MONITOR'S ACRONYM(S)</b>		
						<b>11. SPONSOR/MONITOR'S REPORT NUMBER(S)</b>		
<b>12. DISTRIBUTION / AVAILABILITY STATEMENT</b>  Approved for Public Release; Distribution Unlimited								
<b>13. SUPPLEMENTARY NOTES</b>								
<b>14. ABSTRACT</b> We are examining the hypothesis that mitochondrial inheritance plays a significant role in aggressiveness of prostate cancer in African Americans. In the first year of the project we identified 2,000 noncancerous tissues samples from African American men with prostate cancer and we have extracted DNA from ~ 1,500 of them. We have validated a robust new DNA sequencing technique developed by our collaborator using single amplicon long-range PCR that permits deep coverage (10,000-20,000X on average) of the mitochondrial genome. We have sequenced 652 samples derived from frozen fully using this technology. The additional DNA samples derived from paraffin embedded (FFPE) tissue were more challenging, but have now been sequenced. Mapping of DNA variants in our sequenced genomes to mitochondrial genes is nearly complete and examples of the data we have obtained are presented here. After months of testing, we determined that the ethidium bromide methodology for generating prostate cancer cell line cybrids was not effective and we have instead used a Rhodamine-6-G procedure. PNT1A cybrid cell lines have been generated and we have characterized these cells. PC3 cells have been more difficult, but we continue to work on this cell line. We anticipate the first manuscripts relating to the large amount of data generated in this project will be written and submitted in the spring of 2016. To our knowledge this represents the largest collection of mitochondrial sequence data from African Americans in existence.								
<b>15. SUBJECT TERMS</b>  Mitochondrial DNA sequencing, Prostate Cancer, Cybrid Cells, African Americans								
<b>16. SECURITY CLASSIFICATION OF:</b>				<b>17. LIMITATION OF ABSTRACT</b>	<b>18. NUMBER OF PAGES</b>	<b>19a. NAME OF RESPONSIBLE PERSON</b> USAMRMC		
<b>a. REPORT</b>	<b>b. ABSTRACT</b>	<b>c. THIS PAGE</b>	<b>19b. TELEPHONE NUMBER</b> (include area code)					
Unclassified	Unclassified	Unclassified	Unclassified		27			

## Table of Contents

	<u>Page</u>
1. Introduction.....	4
2. Keywords.....	4
3. Accomplishments.....	4
4. Impact .....	16
5. Changes/Problems .....	17
6. Products .....	17
7. Participants .....	18
8. Special reporting Requirements.....	18
9. Appendices .....	18

## 1. Introduction:

African American men are disproportionately affected by prostate cancer with increased lifetime risk, earlier onset of disease and more advanced stage at diagnosis than Caucasians. Prostate cancer is the leading cause of cancer death in African American men, with mortality more than double that observed in Caucasians. There are nearly 20 million African American men in the US, many of whom face significant risk of developing and dying from prostate cancer. Risk for and aggressiveness of prostate cancer in African American men is thought to originate in part from genetic susceptibility. Several nuclear genes and chromosomal regions have been linked to prostate cancer; however, many studies have not included African American men, and no study has linked genetic polymorphisms with clinical outcome. One factor which has not been carefully examined is mitochondrial inheritance which varies significantly between ethnic and racial groups and could explain large differences in disease characteristics. In the present study, we are examining the hypothesis that mitochondrial inheritance plays a significant role in aggressiveness of prostate cancer in African Americans. Further, we predict that adverse clinical outcomes will be reflected in dysregulation of cellular biochemical processes and in alterations in signaling pathways (Akt pathway, apoptosis). The ability to identify mitochondrial variants or haplogroups that contribute to aggressive disease will help to separate patients with indolent disease who can be spared unnecessary intervention from those who need more immediate and aggressive therapy. Identification of cellular pathways involved will help to target treatment strategies for those with cancers predicted to be more aggressive. In Aim 1 we will conduct a study of mitochondrial inheritance in 1,000 African American men with prostate cancer. We will sequence the mitochondrial genome of all 1,000 samples and determine whether particular mitochondrial variants, genes or haplogroups are associated with markers of aggressive disease (age at diagnosis, stage at diagnosis (including bone metastasis), Gleason score, PSA at diagnosis, PSA recurrence and death from disease). Findings will be replicated using an independent set of 1,000 patients from our own tissue resource with linked clinical data. Our strategy will be to sequence the mitochondrial DNA of all 1,000 patients in the discovery phase as well as in the replication group. We will control for population admixture using the Illumina African American Admixture Panel. Using cybrid technology, we will introduce our previously identified mt10398A variant and mitochondrial variants associated with highly aggressive and least aggressive disease identified in our genotyping study into prostate cell lines (derived from normal and cancer cells). Cybrids differing only in their mitochondrial composition will be examined for viability under metabolic stress, cell cycle distribution, production of reactive oxygen species, O<sub>2</sub> consumption, ATP synthesis, respiratory chain activity and capacity to grow in an anchorage independent manner. The effect of mitochondrial variants on nuclear gene expression will be studied using Western blotting and microarrays.

2. **Keywords:** Mitochondrial DNA sequencing, Prostate Cancer, Cybrid Cells, African Americans

## 3. Accomplishments:

In the text to follow, we provide the Aims and original Statement of Work in italics with progress on the project presented in regular text. This is an update of the work to date including this last year which was after funding for the project was completed.

**Aim 1: To examine the association between mitochondrial DNA variants and clinical outcome in African American men with prostate cancer.**

*Task 1: Extraction of all DNA for Initial Study and Validation Set*

*1a. Complete extraction of DNA/ quantitation of DNA for mitochondrial sequencing* *Months 1-9*  
*(Extraction ongoing at present, expect more than 1,000 samples to be prepared before the project is funded. Additional 1,000 to be extracted for validation set IRB protocol already approved)*

We have extracted DNA on more than 1500 Tissue samples from African American Men with Prostate cancer. Work was initially slower than we had initially planned because of difficulties with extracting adequate DNA from some of the FFPE (paraffin-embedded) samples. After significant technical difficulties we have now

extracted adequate quantities of DNA from more than 1000 FFPE samples and 652 Frozen tissue samples and these samples have been sent for sequencing.

*Task 2: Mitochondrial Sequencing*

<i>2a. Order and test overlapping mitochondrial primers with FFPE DNA to be sure that all primers work with FFPE tissue DNA</i>	<i>Ongoing – Month 3</i>
<i>2b. Redesign/test and mitochondrial primers which do not give good PCR results in 2a.</i>	<i>Months 4-6</i>
<i>2c. Establish Database for mitochondrial sequences</i>	<i>Months 1-5</i>
<i>2d. PCR mitochondrial sequences for primary study and validation set</i>	<i>Months 6-20</i>
<i>2e. Sequence mitochondrial PCR products</i>	<i>Months 6-20</i>
<i>(Illumina Admixture Genotyping to be supported by Helis funds, but performed simultaneously)</i>	<i>Months 6-20</i>

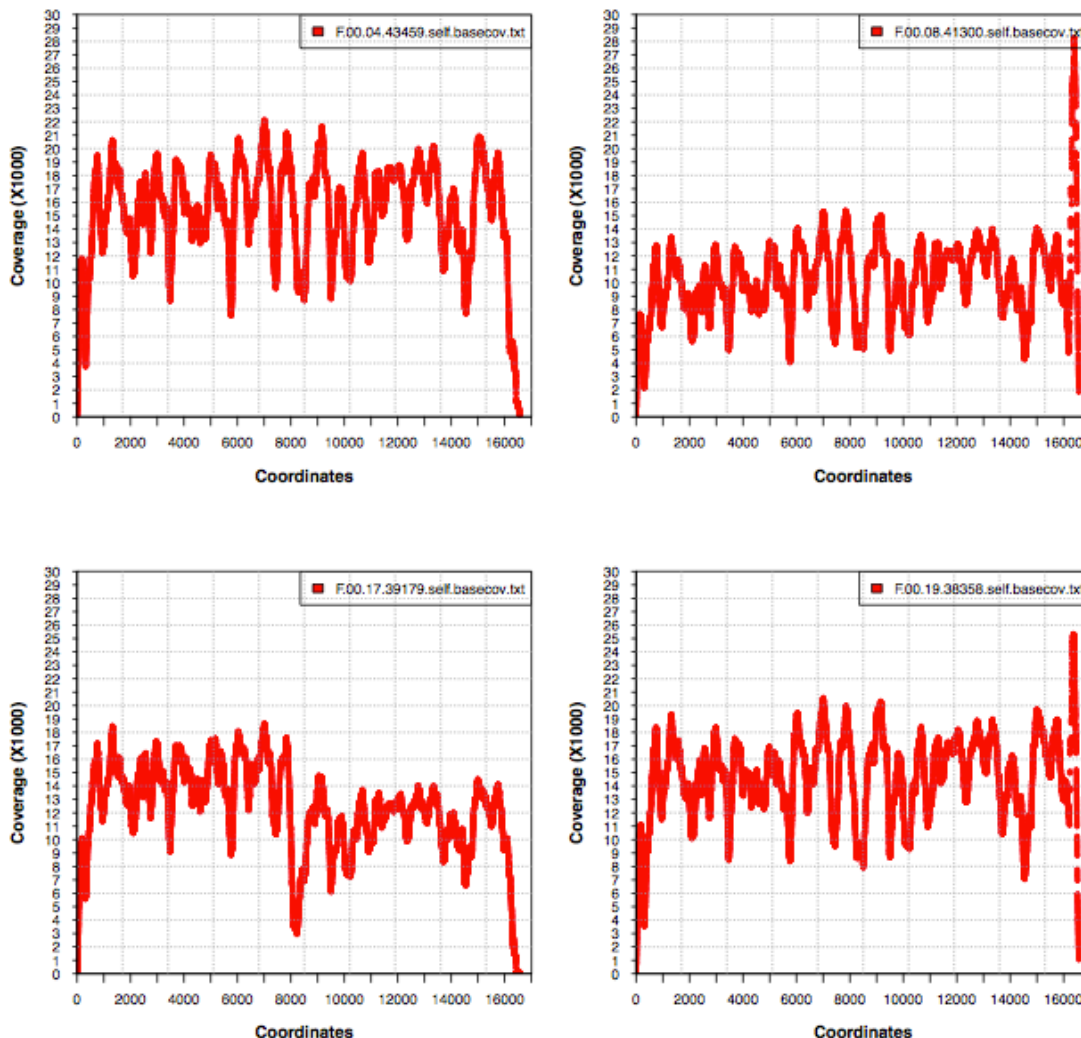
As reported in our initial project proposal, based on published sequences for mitochondrial PCR primers<sup>1, 2</sup> we designed 61 pairs of overlapping primers to amplify the entire 16.6-kb mitochondrial genome. In order to test this technology, we used 15 matched DNA samples derived from 5 patients – frozen tissue, FFPE tissue and whole genome amplified DNA samples derived from FFPE DNA for each patient. Overall 2/3 of the amplicons had reads in both directions on all 15 samples (including WGA DNA). Five of 61 amplicons failed completely and will be redesigned. Thus, we covered 93% of the genome by reads in at least one direction in 90% of the samples, and by reads in both directions in 84% of the mt genome, an excellent outcome for a first pass analysis over a target. Using these results, we identified 165 variants as compared to the published mitochondrial sequence. Within these variants there was 96-100% concordance of calls for each patient across the three sample types. We have continued to optimize the primers for the Sanger sequencing protocols, but as we reported in our initial application, we have also continued to test new technologies for mitochondrial sequencing. As we reported last year, we have obtained excellent results using a new technique developed by one of the collaborators on this project (Dr. Lee-Jun Wong). This technique enriches the entire human mitochondrial genome by a single amplicon long-range PCR followed by massively parallel sequencing.<sup>3</sup> This protocol utilizes less than 100 ng of tissue DNA and makes possible a one- step approach to provide quantitative base calls, exact deletion junction sequences and quantification of deletion heteroplasmy. As we reported last year, this strategy worked well for our frozen tissue DNA and we have now sequenced 652 samples using this technology.

We were not sure that this strategy will be effective in our FFPE DNA samples. Fortunately we had a series of Frozen and FFPE samples derived from the same subject which we could use for comparison purposes. In the figures below we show the result from 4 such Frozen/FFPE pairs. To sequence FFPE samples, we used a sequence-specific capture approach followed by Illumina sequencing. We utilized a custom capture reagent (Roche NimbleGen) targeting the hg19 mitochondrial genome (16.6 Kb). Libraries of ~200 bp insert sizes were prepared from genomic DNA (500 ng) using HGSC Illumina WES protocol and co-captured at 24-48 plex using full-length blockers to capture reads at higher on-target rates. Libraries were pooled and sequenced on a single Illumina HiSeq 2500 lane (~30 Gb). We estimated that each sample would yield 312 Mb of sequence, which roughly would amount to 8000X depth across the target regions. This is a one third less average coverage to the 12000X coverage data currently generated for frozen tissue DNA samples using the long range PCR protocol, where 80 samples are sequenced per lane on Illumina HiSeq. As seen in the figures below, these estimates were relatively accurate with regard to coverage for the FFPE samples. We expected that duplicate rates, which are usually elevated in small target region captures, would not be an issue in this case, as mitochondrial DNA copy number is 1-2 log orders greater than that of the genomic DNA. Non-specific capture of Pseudogenes in the nuclear genome that share high sequence similarity with functional mtDNA genes is another valid concern when using the above described target enrichment approach. However, recent studies have shown that it is possible to bioinformatically determine the limit on heteroplasmy detection due to such contamination and largely eliminate this concern.<sup>4</sup>

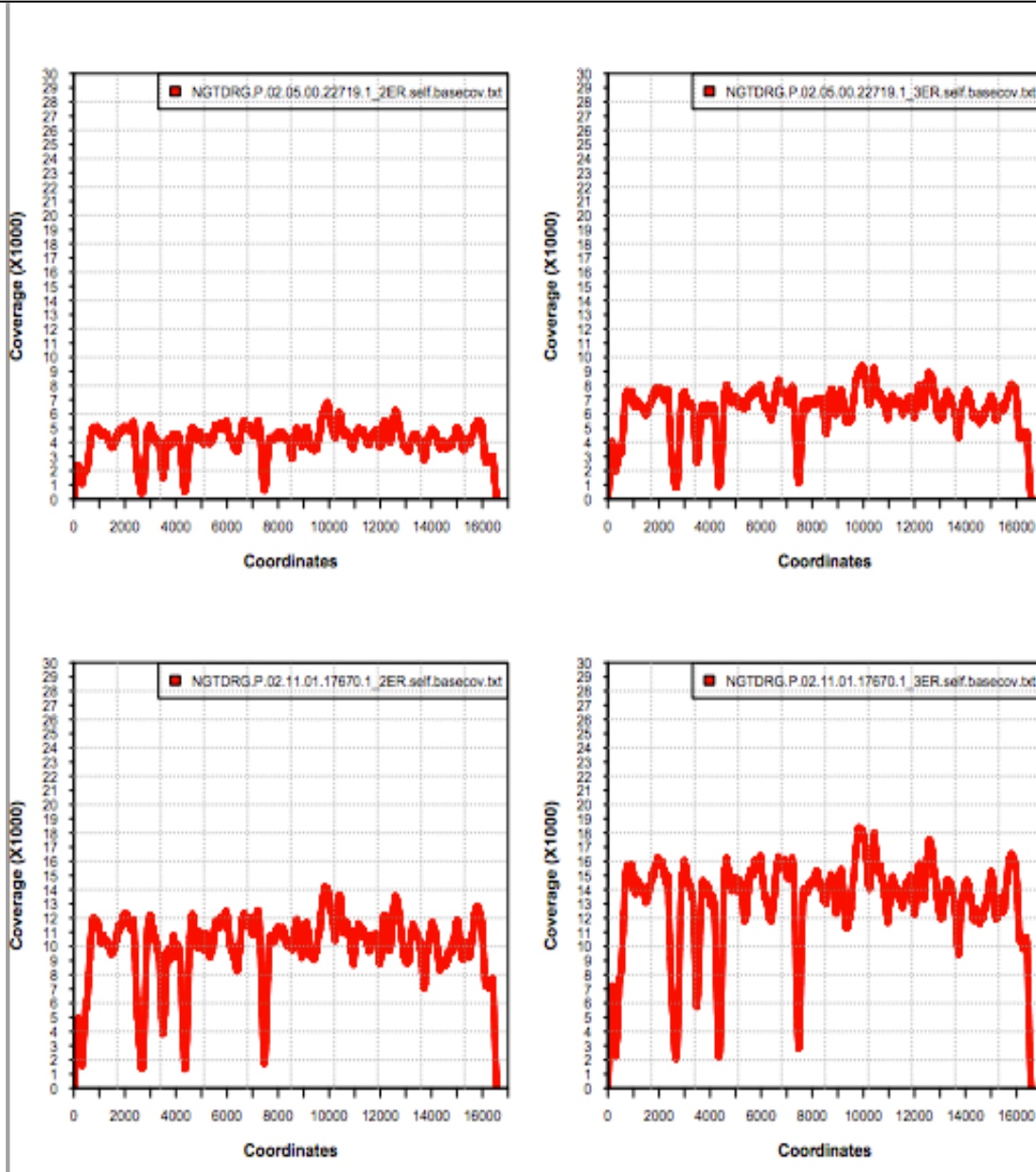
Figure 1A shows results from four samples sequenced by the long range primer technology. We now have sequenced 652 samples using this technology. As seen in figure 1A, a distinct advantage provided by amplification of the entire mitochondrial genome by long range PR with a single primer pair is the uniform coverage. This figure demonstrates that we routinely had coverage depth of 10,000-20,000-fold for the mitochondrial genome. This sequencing strategy has multiple advantages including ease of excluding nuclear DNA sequences which are nearly identical to mtDNA and more uniform coverage of the mitochondrial genome because there is much less risk of having a rare or novel variant at the primer binding site when only one set of primers is used. Deep coverage allows much more accurate base calls and improves the limit of detection for heteroplasmy from about 15% with Sanger sequencing to ~ 1-2% with the long range primer technique. Figure 1B shows sequence data from FFPE DNA samples derived from the same 4 patients; these sequences have been performed in duplicate (data not show) with similar results. Note that as we predicted, coverage depth for the FFPE DNA samples was more like 8000X across the target region. Some question has arisen as to why the depth of coverage is not uniform over the entire mitochondrial genome (sawtooth pattern). We feel this may relate to the GC content of the mitochondrial genome which demonstrates a very similar sawtooth pattern, but we do not feel it will affect our overall results.

**Figure 1: Mitochondrial DNA sequence data**

**Figure 1A: Panels A-D show representative coverage for 4 different DNA samples across the entire 16,569 bp mitochondrial genome. The y axis represents fold coverage at each position (x coordinate) along the mitochondrial genome**



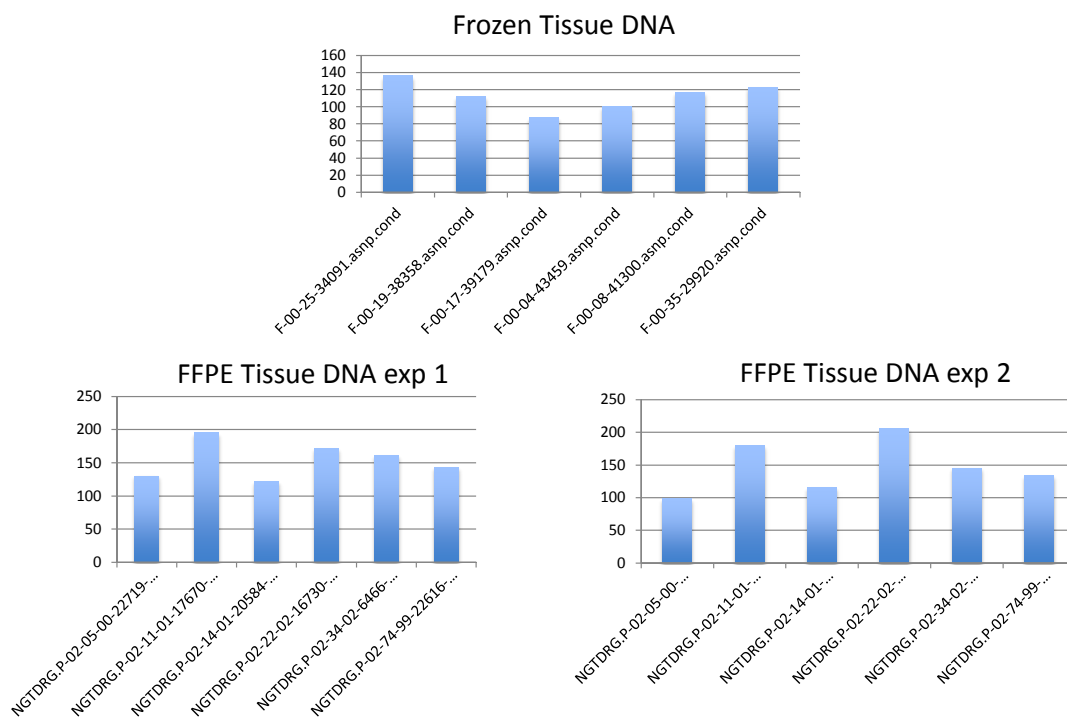
**Figure 1B:** Panels A-D show representative coverage for 4 different FFPE DNA samples across the entire 16,569 bp mitochondrial genome. Note that these 4 samples are matched to the Frozen DNA samples in panel 1A. The y axis represents fold coverage at each position (x coordinate) along the mitochondrial genome.



Using this data we were able to compare SNP counts in Frozen vs FFPE DNA samples. These data (figure 2) demonstrate that similar numbers of SNPs were called in Frozen and FFPE samples. Figure 3 shows that these SNPs were not completely concordant and we are working on this issue as noted below in the updates from this last year .

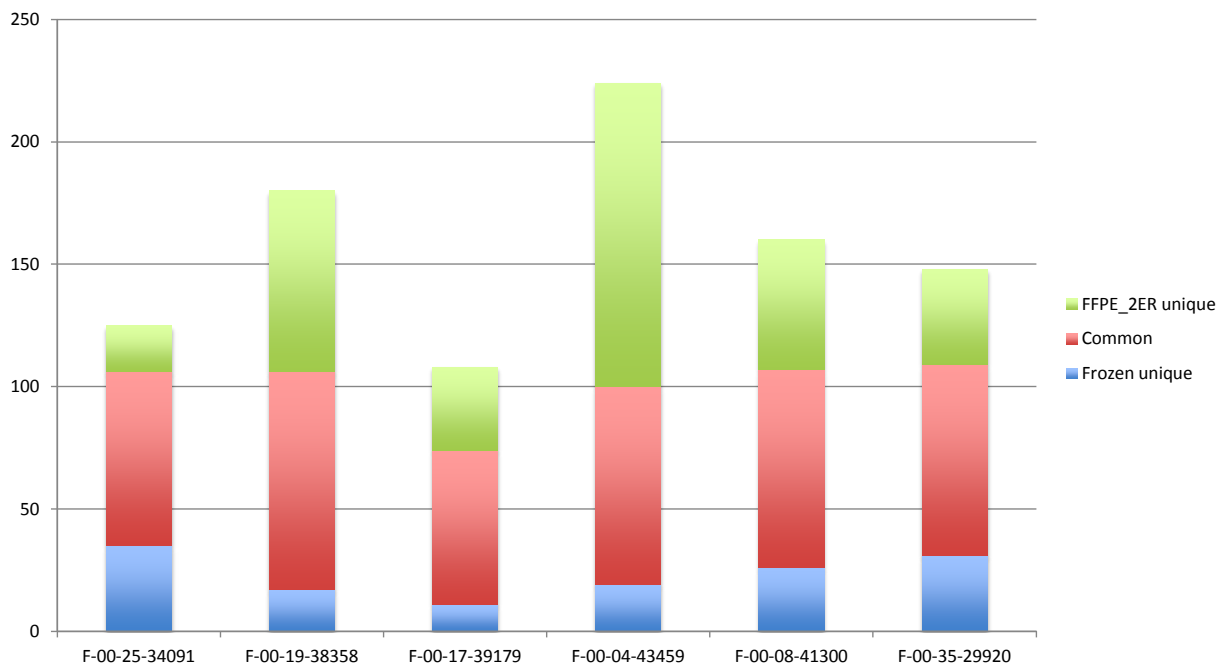
**Figure 2 Mitochondrial SNP counts in Frozen and FFPE DNA samples derived from the same patient**

## SNP counts



**Figure 3 Mitochondrial SNP comparison in Frozen and FFPE DNA samples derived from the same patient**

## SNP comparison (FFPE vs Frozen DNA)



**Task 3: Analysis of Mitochondrial Data**

3a. Analyze mitochondrial sequences and integrate with clinical data

Months 20-28

3b. Admixture analysis based on Illumina data

Months 20-28

3c. Identify mitochondrial variants of interest for cybrid study  
and match with lymphoblast bank

Months 26-28

Figure 2A demonstrates the single nucleotide variant (SNV) calls for the first approximately two hundred subjects. On average each subject has about 150 SNVs with the Cambridge mitochondrial sequence used as reference.<sup>5,6</sup> As shown in our previous report, most of the data at each coordinate clustered at 0 or 1. Since the mitochondrial genome is essentially haploid this suggests there is little contamination with nuclear DNA sequences. Similar data for insertions and deletions is shown in figure 2B.

**Figure 2:** Single nucleotide variants in the mitochondrial genome for the first sequenced subjects.

Figure 2A: Single nucleotide variants in the mitochondrial genome for the first sequenced subjects.

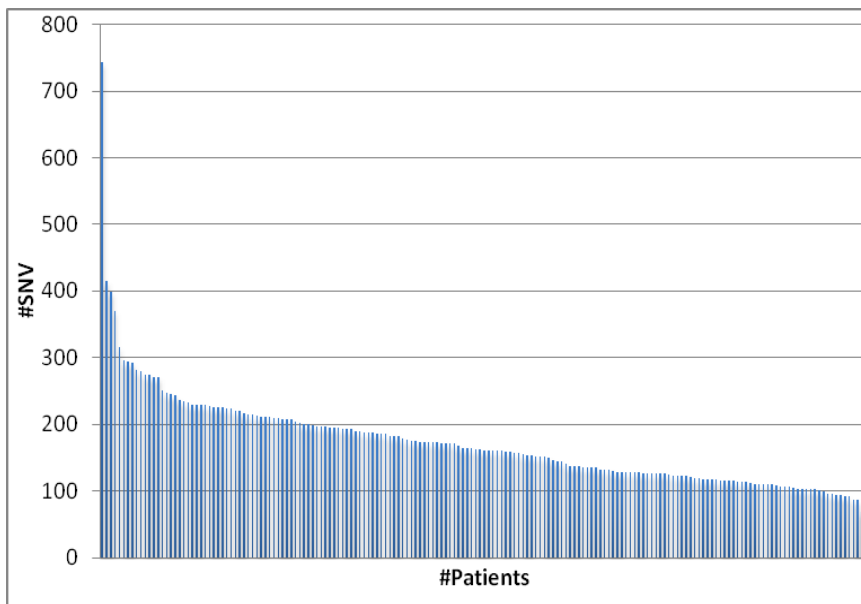
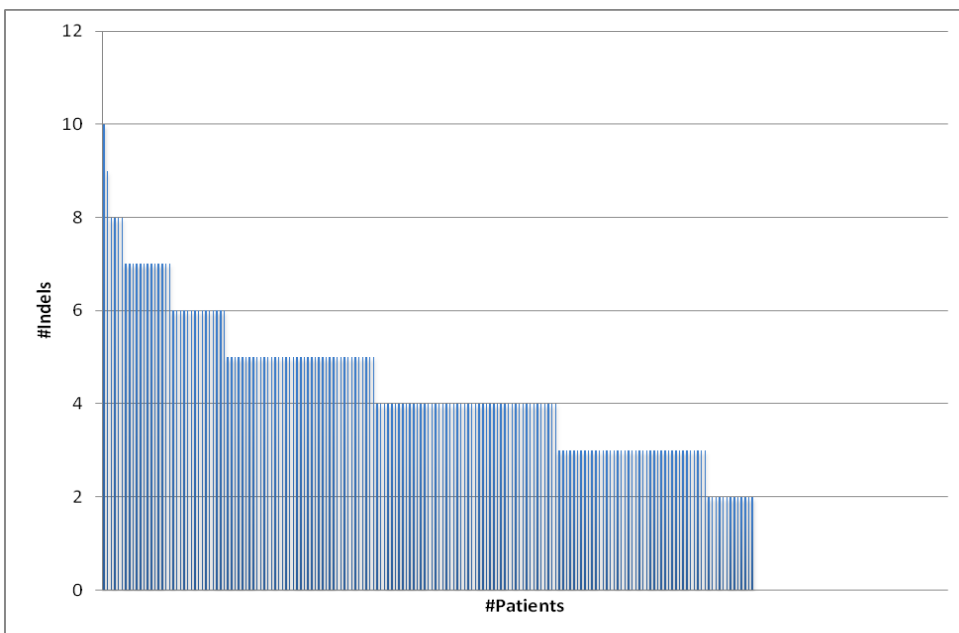


Figure 2B: Insertions and Deletions in the mitochondrial genome for the first sequenced subjects.



We have now created a database of the sequence data which will ultimately include not just mtDNA sequence data, but nuclear SNPs gleaned from the Illumina Race/Ethnicity panel (see below). We have mapped these sequences to the mitochondrial genome (see figure 3 below) to demonstrate the distributions of the mutations found in the various mitochondrial genes. Mutations below the 0.01 variant ratio have been filtered out. Figure 3A shows the SNVs in the mitochondrial genes. As expected the most mutations are seen in the D-loop, but a significant number of additional mutations are found in other regions of the genome as well including the coding regions of for a number of components of the electron transport chain. Similar data in Figure 3B shows the distribution of insertions and deletions across the mitochondrial genes.

**Figure 3:** Distribution of variants found in mitochondrial genes for the initial subjects.

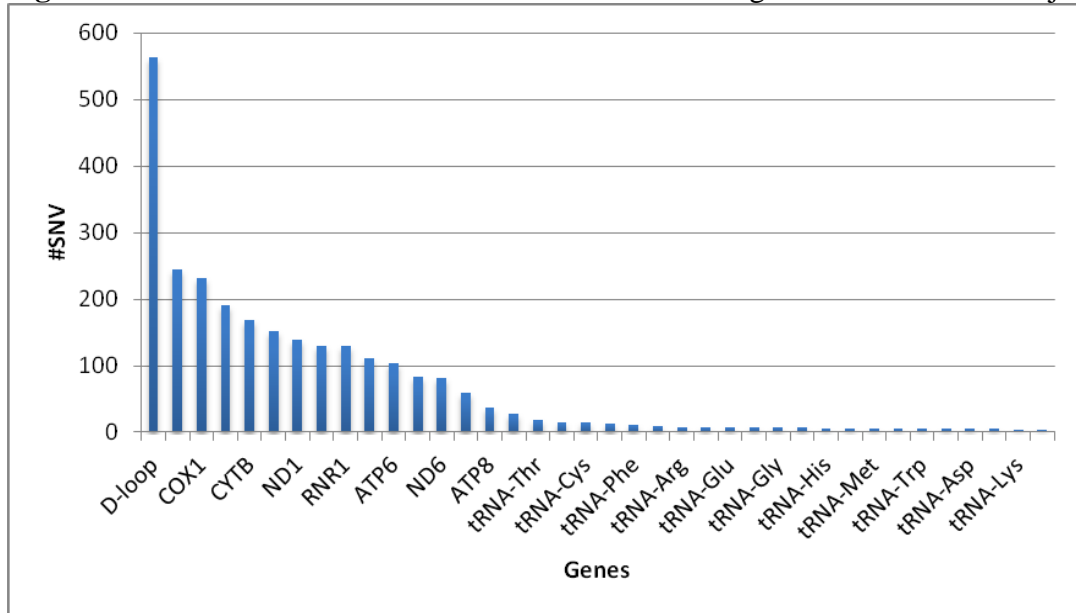


Figure 3A: Distribution of mitochondrial single nucleotide variants found within various mitochondrial genes for the initial subjects as compared to the Cambridge reference sequence. The Y axis represents the absolute number of variants observed within each gene.

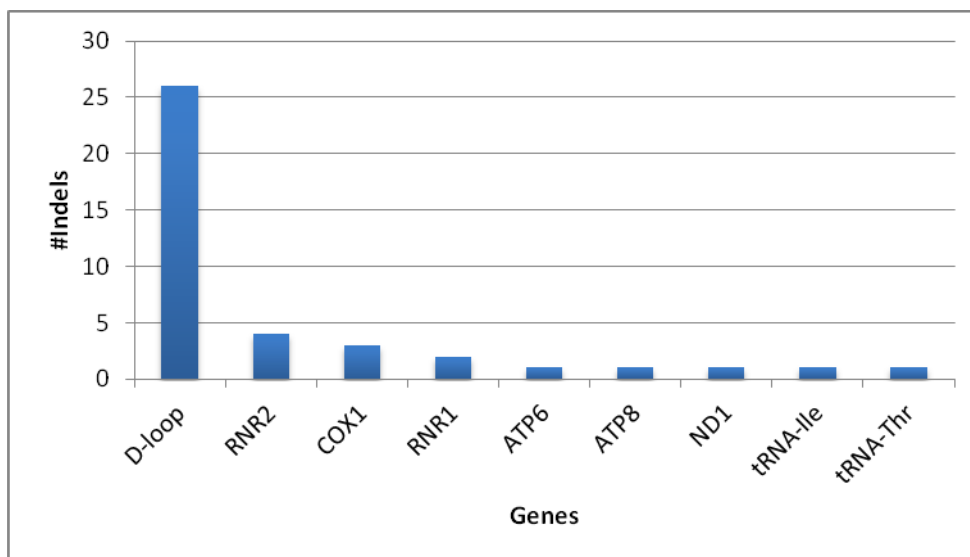


Figure 3B Unique insertions and deletions by gene across the initial patients as compared to the Cambridge reference sequence. The Y axis represents the absolute number of insertions/deletions observed within each gene.

Although not covered by this project, we have also sent more than 300 samples for genotyping on the Illumina African American Admixture panel. Thus far all samples have been successfully genotyped.

***Aim 2: To understand the role of mitochondrial DNA variants in regulation of cellular processes important in the cancer cell including those involved with generation of reactive oxygen species and energy metabolism.***

***Task 1: Creation of Cybrid cell lines***

- |  |                     |
|--|---------------------|
| <i>1a. IRB submission for creation of lymphoblastoid lines from patients</i>   | <i>Months 1-4</i>   |
| <i>1b. Creation of immortalized lymphocyte bank for use in study</i>   | <i>Months 5-24</i>  |
| <i>1c. Order/test primers for variable number tandem repeat studies of nuclear origin</i>                                      | <i>Months 6-12</i>  |
| <i>1d. Deplete prostate cancer cells of mitochondria<br/>3 potential cell lines to be used: PNT1A, LNCaP or PC3</i>            | <i>Months 9-24</i>  |
| <i>1e. Create cytoplasts and fuse with prostate cancer cells devoid of mitochondria to create cybrid cell lines</i>            | <i>Months 20-36</i> |
| <i>1f. Confirm nuclear origin of cybrid cell lines using variable length tandem repeat in insulin receptor and other genes</i> | <i>Months 20-36</i> |

***Task 2: Testing of cybrid cells lines***

- |   |                     |
|---|---------------------|
| <i>2a. Proliferation and cell cycle testing, Oxygen consumption, Production of reactive oxygen species, electron transport chain analysis, analysis of anchorage-independent cell growth.</i> | <i>Months 22-40</i> |
|---|---------------------|

Initially we encountered a number of technical difficulties during the generation of cybrid models using the ethidium bromide (EBr)-mediated mitochondrial DNA (mtDNA) depletion method. During the past year, we continued our original cybrid protocol using the EBr method by varying treatment conditions. Unfortunately, despite our best efforts, quantitative real time PCR (qPCR) analysis suggested a lack of complete mtDNA depletion in these cell lines (rho zero cells). Without the rho zero cells, we could not generate cybrids in prostate cancer models using the original protocol. To overcome these difficulties, we tested multiple alternative approaches including rhodamine-6G (R6G) mediated short-term mitochondrial dysfunction in generating rho zero cells. We used the benign prostate cell line PNT1A and the prostate cancer cell line PC3. Generation of rho zero using another prostate cancer cell line, LNCaP, was not successful due to low adherence of these cells to the bottom of cell culture plates after treatment with mtDNA depletion agents. Lack of adherence has made separation of viable rho zero cells very difficult.

Repeated efforts using R6G treatment finally succeeded in the generation of PNT1A rho zero; we have not yet been successful with the PC3 cells, which did not survive the rho zero condition. Cybrid models were then generated in the PNT1A benign prostate nuclear background. Mitochondria from enucleated metastatic prostate cancer cell line PC3 or benign prostate cell line PNT1A were fused with the rho zero PNT1A nuclear donor to form the cybrids (PC3/PNT1A and PNT1A/PNT1A respectively). We did extensive analysis of these cybrids to confirm their mitochondrial and nuclear identity. We sequenced multiple regions of mtDNA to confirm the mitochondrial identity (Figure 1). Similarly, we sequenced several nuclear SNP markers (Figure 2) as well as short tandem repeat (STR) markers (Figure 3) to confirm the nuclear identity of the cybrids. These analyses confirmed that both PC3/PNT1A and PNT1A/PNT1A cybrids contain PNT1A benign nuclear background but mitochondria from PC3 and PNT1A, respectively.

In order to confirm the role of mitochondrial retrograde regulation of cancer mitochondria in a nuclear background we analyzed the cancer property of these cybrids. We used the soft agar colony formation assay with similar number of PC3 and PNT1A parental cell lines as well as PC3/PNT1A and PNT1A/PNT1A cybrids under uniform experimental conditions. We observed that the while PNT1A/PNT1A cybrid which contain both nucleus and mitochondria from the benign cell line rarely form colonies in the soft agar like PNT1A parental

cells, PC3/PNT1A cybrid with cancer mitochondria could grow anchored dependently in the soft agar (Figure 4).

**Figure-1: Confirmation of mitochondria by sequencing mtDNA of parental cells and cybrids**

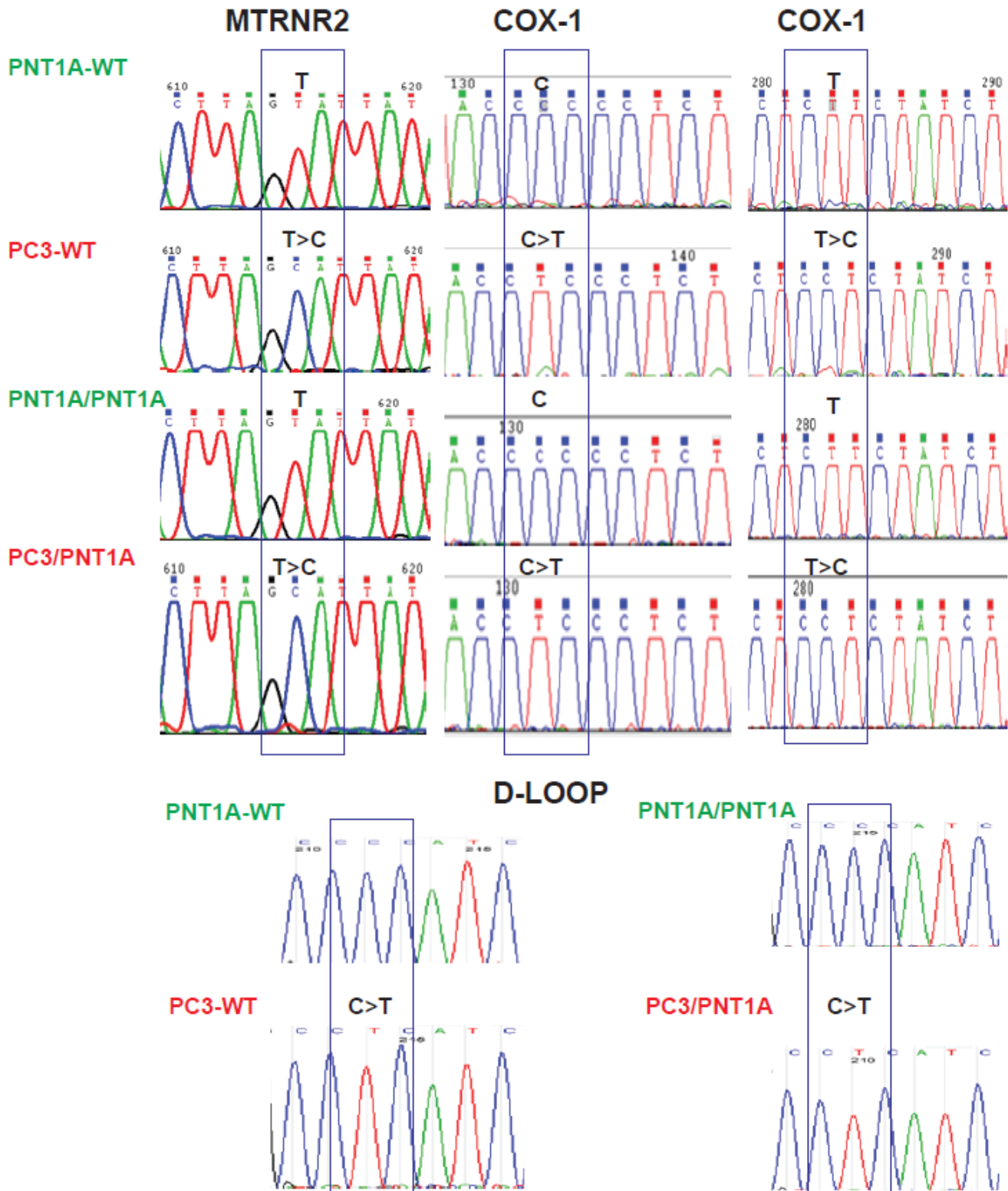


Figure-2: Confirmation of nucleus by sequencing random SNPs in nuclear DNA

PNT1A-WT

No.	Sample File	Dir	Gene	Exon	Start	End	Size	Quality	Mut#	Mutation1
1	PNT1A-InQC-1F_2013-10-22_B01.ab1	1-F	InternalQC-01	ohr1	-55	68	111	14	1	c.2C>GA,p.S1WX
1							111	14	1	100.00%
2	PNT1A-InQC-3F_2013-10-22_B02.ab1	2-F	InternalQC-03	ohr8	-40	68	88	17	1	c.2G>CT,p.R1TM
1							88	17	1	100.00%
3	PNT1A-InQC-5F_2013-10-22_B03.ab1	3-F	InternalQC-05	ohr11	-55	48	104	20	1	c.2G>CT,p.R1TM
1							104	20	1	100.00%
4	PNT1A-InQC-7F_2013-10-22_B04.ab1	4-F	InternalQC-07	ohr13	-55	68	113	31	1	c.2G>CT,p.R1PL
1							113	31	1	100.00%
5	PNT1A-InQC-9F_2013-10-22_B05.ab1	5-F	InternalQC-08	ohr15	-54	61	105	9	1	c.2T>A,p.I1N
1							105	9	1	100.00%
6	PNT1A-InQC-10F_2013-10-22_B06.ab1	6-F	InternalQC-10	ohr19	-10	68	88	10	1	c.2G>CT,p.R1TM
1							88	10	1	100.00%
7	PNT1A-InQC-12F_2013-10-22_B07.ab1	7-F	InternalQC-12	ohr20	-51	65	108	35	1	c.2T>GA,p.L1RQ
1							108	35	1	100.00%

PC3-WT

No.	Sample File	Dir	Gene	Exon	Start	End	Size	Quality	Mut#	Mutation1
1	PC3-InQC-1F_2013-10-22_C01.ab1	1-F	InternalQC-01	ohr1	-55	68	111	20	1	c.2C>GA,p.S1WX
1							111	20	1	100.00%
2	PC3-InQC-3F_2013-10-22_C02.ab1	2-F	InternalQC-03	ohr8	-40	68	88	22	1	c.2G>TC,p.R1MT
1							88	22	1	100.00%
3	PC3-InQC-5F_2013-10-22_C03.ab1	3-F	InternalQC-05	ohr11	-55	48	104	18	1	c.2G>CT,p.R1TM
1							104	18	1	100.00%
4	PC3-InQC-7F_2013-10-22_C04.ab1	4-F	InternalQC-07	ohr13	-55	68	113	28	1	c.2G>CT,p.R1PL
1							113	28	1	100.00%
5	PC3-InQC-10F_2013-10-22_C06.ab1	5-F	InternalQC-10	ohr19	-10	67	87	18	1	c.2G>T,p.R1M
1							87	18	1	100.00%
6	PC3-InQC-12F_2013-10-22_C07.ab1	6-F	InternalQC-12	ohr20	-52	65	107	31	1	c.2T>GA,p.L1RQ
1							107	31	1	100.00%
7	PC3-InQC-8F_2013-10-22_C05.ab1				-1	-178	178	0	-1	Bad Data

PNT1A/PNT1A

No.	Sample File	Dir	Gene	Exon	Start	End	Size	Quality	Mut#	Mutation1
1	PNT1A-InQC-1F_2013-10-22_B01.ab1	1-F	InternalQC-01	chr1	-55	68	111	14	1	c.2C>GA,p.S1WX
1							111	14	1	100.00%
2	PNT1A-InQC-3F_2013-10-22_B02.ab1	2-F	InternalQC-03	chr5	-40	68	98	17	1	c.2G>CT,p.R1TM
1							98	17	1	100.00%
3	PNT1A-InQC-5F_2013-10-22_B03.ab1	3-F	InternalQC-05	chr11	-55	49	104	20	1	c.2G>CT,p.R1TM
1							104	20	1	100.00%
4	PNT1A-InQC-7F_2013-10-22_B04.ab1	4-F	InternalQC-07	chr13	-55	68	113	31	1	c.2G>CT,p.R1PL
1							113	31	1	100.00%
5	PNT1A-InQC-9F_2013-10-22_B05.ab1	5-F	InternalQC-09	chr15	-54	61	106	9	1	c.2T>A,p.I1N
1							106	9	1	100.00%
6	PNT1A-InQC-10F_2013-10-22_B06.ab1	6-F	InternalQC-10	chr19	-10	68	88	10	1	c.2G>CT,p.R1TM
1							88	10	1	100.00%
7	PNT1A-InQC-12F_2013-10-22_B07.ab1	7-F	InternalQC-12	chr20	-51	65	108	35	1	c.2T>GA,p.L1RQ
1							108	35	1	100.00%

PC3/PNT1A

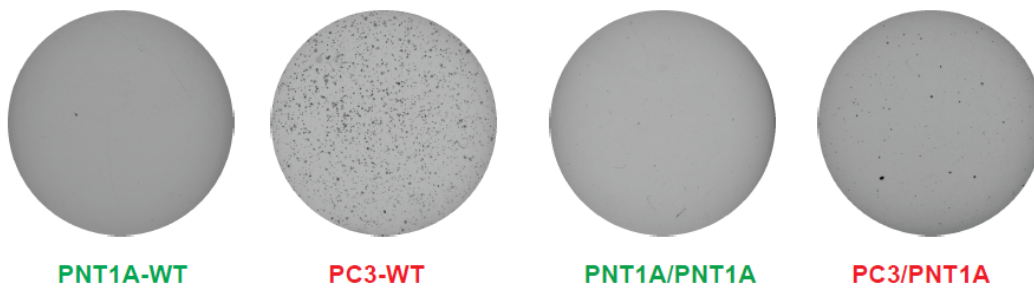
No.	Sample File	Dir	Gene	Exon	Start	End	Size	Quality	Mut#	Mutation1
1	PC_PNT1A-PLK01-InQC-1F_2013-10-22_E01.ab1	1-F	InternalQC-01	ohr1	-55	68	111	28	1	c.2C>GA,p.S1WX
1							111	28	1	100.00%
2	PC_PNT1A-PLK01-InQC-3F_2013-10-22_E02.ab1	2-F	InternalQC-03	ohr8	-40	68	88	23	1	c.2G>TC,p.R1MT
1							88	23	1	100.00%
3	PC_PNT1A-PLK01-InQC-5F_2013-10-22_E03.ab1	3-F	InternalQC-05	ohr11	-55	48	104	21	1	c.2G>CT,p.R1TM
1							104	21	1	100.00%
4	PC_PNT1A-PLK01-InQC-7F_2013-10-22_E04.ab1	4-F	InternalQC-07	ohr13	-52	68	110	28	1	c.2G>CT,p.R1PL
1							110	28	1	100.00%
5	PC_PNT1A-PLK01-InQC-9F_2013-10-22_E05.ab1	5-F	InternalQC-09	ohr15	-60	61	101	22	1	c.2T>A,p.I1N
1							101	22	1	100.00%
6	PC_PNT1A-PLK01-InQC-10F_2013-10-22_E06.ab1	6-F	InternalQC-10	ohr19	-13	67	70	16	1	c.2G>CT,p.R1TM
1							70	16	1	100.00%
7	PC_PNT1A-PLK01-InQC-12F_2013-10-22_E07.ab1	7-F	InternalQC-12	ohr20	-60	65	105	38	1	c.2T>GA,p.L1RQ
1							105	38	1	100.00%

**Figure 3 STR profiles of cybrids and parental cells**

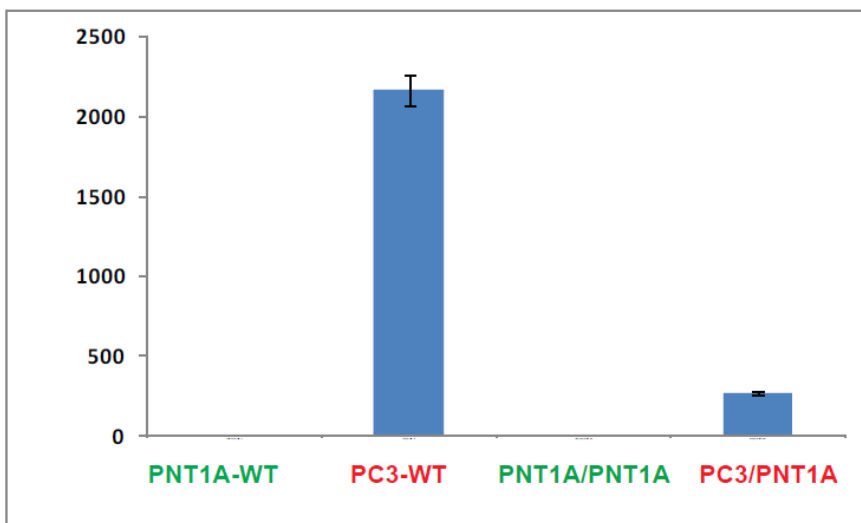
Sample_Name	AMEL	CSF1P O	D13S 317	D16S 539	D18S 51	D21S 11	D3S1 358	D5S81 8	D7S8 20	D8S11 79	FGA	TH01	TPOX	vWA
PNT1A-WT	X,Y	12	11,12	10,11	16	27,28	16	12	10,11	10,12	23,24	6,7,9.3	8,11	17,19
PC3/PNT1A- PLKO1	X,Y	12	11	10,11	16	27,28	16	12	10,11	10,12	24	6,7,9.3	8	17,19
PNT1A/PNT1A- PLKO1	X,Y	12	11,12	10,11	16	27	16	12	10,11	10,12	24	6,7,9.3	8,11	17,19
PC3-ATCC	X	11	11	11				13	8,11			6,7	8,9	17

**Figure-4: Analysis of tumor properties in parentals and cybrids by colony formation assay**

4A: Phase contrast images of soft agar colonies after two weeks of culture



4B: Number of colonies counted using Gelcount colony counter



Overall, we successfully generated cybrid models under benign prostate nuclear background and analyzed its genetic and tumor properties using different mitochondrial donors. We are continuing our efforts to generate new cybrid models with cancer nuclear background and to further characterize the already generated PNT1A cybrids.

### **References:**

1. Chen TJ, Boles R, Wong, L-J C. Detection of mitochondrial DNA mutations by temporal temperature gradient gel electrophoresis. *Clin Chem.* 1999; 45: 1162–1167.
2. Wong LJ, Liang MH, Kwon H, Park J, Bai RK, Tan DJ. Comprehensive scanning of the entire mitochondrial genome for mutations. *Clin Chem.* 2002; 48:1901-12.
3. Zhang W, Cui H, Wong L-J. Comprehensive One-Step Molecular Analyses of Mitochondrial Genome by Massively Parallel Sequencing. *Clin Chem.* 2012; 58: 1322-1331.
4. Falk MJ, Pierce EA, Consugar M, Xie MH, Guadalupe M, Hardy O, Rappaport EF, Wallace DC, LeProust E, Gai X. Mitochondrial disease genetic diagnostics: optimized whole-exome analysis for all MitoCarta nuclear genes and the mitochondrial genome. *Discov Med.* 2012 14:389-99.
5. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F, Schreier PH, Smith AJ, Staden R, Young IG. Sequence and organization of the human mitochondrial genome. *Nature* 1981; 290: 457–465.
6. Andrews RM, Kubacka I, Chinnery PF, Lightowlers RN, Turnbull DM, Howell N "Reanalysis and revision of the Cambridge reference sequence for human mitochondrial DNA. *Nat Genet* 1999; 23: 147.

### **Key Accomplishments Summarized:**

1. Isolation of more than 1500 DNA samples required for the project.
2. Validation of a more robust sequencing technique using single amplicon long-range PCR that permits deep coverage (10,000-20,000X on average) of the mitochondrial genome, to date 652 samples of Frozen Tissue DNA have been sequenced.
3. Sequencing of FFPE DNA samples using Sequence Capture.
4. Mapping of the DNA variants to mitochondrial genes (data shown for ~first 200 samples nearly completed for remaining samples).
5. Cybrid models of benign (PNT1A) prostate cells have been successfully created and characterized. Despite the end of the project funding, we are continuing to work with the PC3 cells with the expectation that we will be able to create PC3 cybrids and characterize them.

### **Opportunities for professional development:**

Though not funded by the DOD directly, several students and postdoctoral fellows have interacted with the project and in the summer of 2015, an undergraduate student from an underserved minority group. This last summer the undergraduate student (under the supervision of one of a postdoctoral fellow in Dr. Wheeler's group) worked on coverage analysis of mitochondrial samples comparing the coverages obtained from FFPE samples to those of fresh frozen samples. He determined that the overall coverage patterns of fresh frozen were different from those of FFPE and that this would require separate analysis of the two approaches. He also identified a region of the FFPE capture that was poorly covered compared with fresh frozen and attempted to map the mitochondrial genes to that area. This area is largely "uncharted territory" and allowed both the postdoctoral fellow and the undergraduate to develop new skills in analyzing sequence data. Post doctoral fellows who were part of the Houston REACH IRACDA, or Houston Research Education and Career Horizon Institutional Research and Academic Career Development Award, K12 Postdoctoral Fellowship Program have also participated in the project. This IRACDA program is a model for postdoctoral training that combines frontier-level biomedical research and training to teach research-oriented courses to undergraduates at partner institutions. The project includes under-represented post-doctoral fellows when possible, who provide excellent role models for the undergraduates.

**Dissemination of data to communities of interest:**

Though data from this work has not yet been published in a peer reviewed journal, portions of the data have been presented in local venues including the Human Genetics Postdoctoral seminar series and the SMART (Summer Medical And Research Training) Program Seminar series aimed at undergraduate students interested in research and medical careers.

**Next Reporting period plans:**

We are excited to report that we have been able to recruit an outstanding postdoctoral fellow – Dr Jing (Cindy) Dong - to work on this project in the next year. Dr. Aaron Thrift, who had intended to work on this project, will serve as one of Dr. Dong’s mentors. Funds for Dr Dong’s salary support will come from the Cancer Prevention and Research Institute of Texas via a training grant held by Dr Margaret Spitz, a BCM Faculty member. Dr. Dong is presently a postdoctoral fellow in Aging and Neuroepidemiology at the Epidemiology branch of the National Institute of Environmental Health Sciences at Research Triangle Park in North Carolina. Dr. Dong has a PhD in Epidemiology and significant experience in genome-wide association studies and genetic and environmental risk factors. Dr. Dong came to Houston in June of 2016 to interview and give a seminar. BCM faculty members were very impressed with her and offered her a position on the spot which has required that she change her visa status (which is taking months unfortunately). During that trip we met at some length to discuss the prostate mitochondrial project – Dr. Dong was extremely enthusiastic to become involved and assist with leading the data analysis. Since that time we have had a series of phone calls to plan data analysis and the decision was to obtain an IRB exemption agreement so that Dr Dong could begin data analysis even prior to joining BCM. That agreement which has just been approved contained the following statements signed by both parties:

***Sender of coded specimens/data:***

**I, Dr. Martha Mims at Baylor College of Medicine, holder of the code-key, cipher and/or identifiers for the shared [Prostate Mitochondrial DNA Data], promise not to release individually identifiable information or about the subjects from whom the [Prostate Mitochondrial DNA Data], derive to Jing Dong, PhD, of NIEHS per the provisions of U.S. Code 45 C.F.R, 46.**

***Recipient of coded specimens/data:***

**I, Jing Dong, PhD, of NIEHS, recipient of the [Prostate Mitochondrial DNA Data], promise not to request individually identifiable information about the subjects from whom the [Prostate Mitochondrial DNA Data], derive, from Dr. Martha Mims at Baylor College of Medicine per the provisions of U.S. Code 45 C.F.R. 46.**

No data has yet been sent to Dr Dong, but our intention is to initiate data analysis intensively in the next year with the expectation that we will produce a high quality publication. Dr Wheeler is fully in support of this plan.

**4. Impact:****Development of the principal discipline:**

This project performed mitochondrial sequencing of more than 1000 subject samples from African American men. To our knowledge this is the largest such database in existence. In concert with Dr. David Wheeler in the Human Genome Sequencing Center and Dr. Lee-Jun Wong, we have implemented a new more rapid and much more powerful technique for sequencing mitochondrial DNA using single amplicon long-range PCR. This technology permitted us to rapidly sequence 652 samples using NGS technology at a depth of coverage of 10,000-20,000X. This technology allows much more accurate base calls and improves the limit of detection for heteroplasmy from about 15% with Sanger sequencing to ~ 1-2% with the long range primer technique. This technology has permitted much more rapid generation of mtDNA sequence with much less DNA and less “contamination” by nuclear DNA. Further, we have optimized sequencing strategies for the FFPE DNA and this portion of the project has just been completed. Two strategies were necessary for sequencing these two

kinds of DNA samples and both of these strategies were either developed or perfected during the course of this project. Publication of the methodology and the data obtained will represent a huge leap forward for the field of mitochondria and cancer. Further, we have created a large database of DNA sequence combined with clinical data and we are in the process of analyzing this data (See above) to produce the first manuscript relating to this work. We envision that this annotated data will reveal that mitochondrial metabolism directly impacts the development of prostate cancer.

**Impact on other disciplines:**

Though we had not foreseen this when the project was first envisioned, we are now working with investigators in metabolomics to examine whether the mitochondrial sequence changes we identify can be correlated with changes in metabolism in vivo (urine or plasma metabolome).

**Impact on Technology Transfer:**

Nothing to report

**Impact on Society:**

The overarching goal of this project is to understand why African American men are more prone to develop prostate cancer than any other racial/ethnic group. We envision that mitochondrial inheritance plays a significant role in aggressiveness of prostate cancer in African American men. Uncovering risk alleles may impact both screening and treatment for these men.

**5. Changes/Problems:**

**Changes in Approach:**

No significant changes in objectives or scope were encountered. As detailed in Aim 1, improvements in technology developed during the project allowed us to gain much deeper sequencing information than originally envisioned. As detailed in Aim 2, the first approach to developing cybrids was not feasible and we utilized an alternative technology which was successful in the PNT1A cells.

**Problems or Delays**

Creation of the cybrid cells proved to be more difficult than we had envisioned at the outset of the project. Ultimately we were able to develop and characterize the PNT1A cybrids, but we are still struggling the PC3 cybrids making slow progress. As noted below, we will be able to share our insights in a methods paper which we hope to publish in the spring of 2017.

**Changes that had a significant impact on expenditures**

None

**Significant changes in humans subjects, vertebrate animals or select agents:**

None

**6. Products:**

**Publications:**

No completed reportable outcomes have been accomplished as yet, however; the mitochondrial database with mtDNA sequence data, African American Admixture Data and Clinical Data is being assembled now. This is a massive amount of data – more than we had originally foreseen. In order to analyze the data, we have recruited Dr. Cindy Dong (See above) to work with Dr. Aaron Thrift, in the Cancer Epidemiology program at the Dan L Duncan Cancer Center. With the help of Drs. Thrift and Dong, we will be in a position to submit the first manuscript on this data – later than we had wanted to do, but we have overcome the technical challenges inherent in sequencing the FFPE DNA samples to create a very large database of information. The cybrid cell lines were unfortunately more difficult to create than we had anticipated, but we have succeeded in creating PNT1A cybrids and we have not given up on the PC3 cybrids.

**Websites:**

None

**Technologies:**

Technology for sequencing mitochondrial DNA from FFPE samples is described above. In concert with Dr. David Wheeler in the Human Genome Sequencing Center and Dr. Lee-Jun Wong, we have implemented a new more rapid and much more powerful technique for sequencing mitochondrial DNA using single amplicon long-range PCR. This technology permitted us to rapidly sequence 652 samples using NGS technology at a depth of coverage of 10,000-20,000X. This technology allows much more accurate base calls and improves the limit of detection for heteroplasmy from about 15% with Sanger sequencing to ~ 1-2% with the long range primer technique. This technology has permitted much more rapid generation of mtDNA sequence with much less DNA and less “contamination” by nuclear DNA. We have optimized sequencing strategies for the FFPE DNA and this portion of the project has just been completed. We will publish this information in a peer reviewed journal along with the actual data gleaned from these samples. The pitfalls of using DNA derived from FFPE samples as compared to DNA derived from frozen tissue will be discussed in detail as will details of how to manage these kinds of data.

**Inventions:**

None

**Other Products:**

We are in the process of creating a very large database of DNA sequence information and clinical data on African American men with prostate cancer. This information will be invaluable in terms of evaluating the role mitochondrial inheritance on prostate cancer aggressiveness. Further, we have developed a library of germline DNA which can be queried for other projects.

**7. Participants**

**Listed below are the participants in the project though no funding for their roles was provided in the last year.**

Martha Mims, MD, PhD, PI No Change  
Lee-Jun Wong, PhD, Co-Investigator No Change  
David Wheeler, PhD, Co-Investigator No Change  
Suzanne Leal, PhD, Co-Investigator No Change  
Myoung Kwon, Research Technician, No Change

**8. Special Reporting Requirements: None****9. Appendices: CV of Dr. Jing Dong proposed postdoctoral fellow**

# CURRICULUM VITAE

## Jing Dong Ph.D.

### CONTACT INFORMATION

---

Epidemiology Branch  
National Institute of Environmental Health Sciences  
BG 101 RM A359 MSC A3-05  
111 T.W. Alexander Drive  
Research Triangle Park, NC 27709

### EDUCATION

---

**Postdoctoral Fellow** in Aging & Neuroepidemiology Oct. 2012 - present  
National Institutes of Health (NIH) / National Institute of Environmental Health Sciences (NIEHS)

- Supervisor: Dr. Honglei Chen
- Research focus: Genetic and environmental risk factors for neurodegenerative disease

**Ph.D. (Five-year Program)** in Epidemiology Sep.2007 – July 2012  
Nanjing Medical University

- Advisor: Prof. Hongbing Shen
- Research focus: Genetic and environmental risk factors for cancers
- Thesis: “Genome-wide association analysis of lung cancer in Han Chinese”

**B.S.** in Prevention Medicine Sep. 2002 – July 2007  
Nanjing Medical University

### FUNDING

---

**Principle Investigator**, National Natural Science Foundation of China, “Fine-mapping and functional analysis for lung cancer susceptibility region 5q31.1 in Chinese” 2014-2016

**Principle Investigator**, Research Innovation Program for College Graduates of Jiangsu Province, “The association between MDM2 polymorphism and non-small cell lung cancer survival” 2010-2012

**Participation**, The National High Technology Research and Development Program of China (863 project), Equivalent to UO1, “Genome-wide association study of lung cancer” 2010-2012

**Participation**, National Science Foundation of China, Equivalent to RO1, “MicroRNA polymorphisms and NSCLC prognosis” 2010-2014

### AWARDS AND HONORS

---

The second-class scholarship from Nanjing Medical University 2003  
The first-class scholarship from Nanjing Medical University 2004

The second-class scholarship from Nanjing Medical University	2005
The second-class scholarship from Nanjing Medical University	2006
The second-class scholarship from Nanjing Medical University	2008
The first-class scholarship from Nanjing Medical University	2009
The first-class scholarship from Nanjing Medical University	2010
The first-class scholarship from Nanjing Medical University	2011
Excellent student in Jiangsu Province	2012
Excellent Graduate of Nanjing Medical University	2012
Excellent Doctoral Dissertation of Jiangsu Province	2013

## **TEACHING AND PROFESSIONAL EXPERIENCE**

**Graduate Teaching Assistant** in Epidemiology 2009-2012  
Nanjing Medical University

**Graduate Research Assistant** 2010-2012  
Department of epidemiology and biostatistics, Nanjing Medical University

## **INVITED PRESENTATIONS**

“Susceptibility Loci for Pigmentation and Melanoma in Relation to Parkinson’s Disease”, Boston, ASHG 2013 annual meeting 2013

“Genome-Wide Association Analysis of the Sense of Smell in U.S. Older Adults: Identification of Novel Risk Loci in African-Americans and European-Americans”, Biltmore, ASHG 2015 annual meeting 2015

## **MEMBERSHIP IN PROFESSIONAL SOCIETIES**

The American Society of Human Genetics 2013-

## **PEER REVIEWER FOR JOURNALS**

Parkinsonism & Related Disorders  
Oncotarget  
Carcinogenesis  
The Pharmacogenomics Journal  
Annals of Surgical Oncology  
The International Journal of Neuroscience  
Lung cancer  
PLOS One  
Tumor Biology  
Mutation Research  
Journal of Bioanalysis & Biomedicine  
Gene  
Frontiers of Medicine

**REVIEW EDITOR FOR JOURNALS**

---

Frontiers in Genetics

**EDITORIAL BOARD**

---

Annals of Translational Medicine  
JSM Biology**PUBLICATIONS (First / Correspondence Author)**

---

1. **Dong J\***, Wyss A, Yang J, Price R, Nicolas A, Nalls M, Tranah G, Franceschini N, Xu Z, Schulte C, Alonso A, Cummings S, Fornage M, Zaykin D, Li L, Huang X, Kritchevsky S, Liu Y, Gasser T, Wilson R, De Jager P, Singleton A, Pinto J, Harris T, Mosley Jr. T, Bennett D, Stephanie London S, Yu L, Chen H. Genome-Wide Association Analysis of the Sense of Smell in U.S. Older Adults: Identification of Novel Risk Loci in African-Americans and European-Americans. (\*correspondence author) (**Submitted**)
2. **Dong J\***, Yang J, Tranah G, Franceschini N, Parimi N, Alkorta-Aranburu G, Xu Z, Alonso A, Cummings S, Huang X, Fornage M, Kritchevsky S, Liu Y, London S, Niu L, Wilson BS, Jager PD, Yu L, Singleton AB, Bennett DA, Harris T, Mosley TH, Pinto JM, Chen H. Meta-Analysis of Genome-Wide Association Studies on the Sense of Smell in US Elderly. *Medicine (Baltimore)* 2015 Nov;94(47):e1892. (\*correspondence author)
3. **Dong J\***, Beard JD, Umbach DM, Park Y, Huang X, Blair A, Kamel F, Chen H. Dietary fat intake and risk for Parkinson's disease. *Mov Disord.* 2014 Nov;29(13):1623-30. (\*correspondence author)
4. **Dong J**, Gao J, Nalls M, Gao X, Huang X, Han J, Singleton AB, Chen H, IPDGC. Susceptibility loci for pigmentation and melanoma in relation to Parkinson's disease. *Neurobiol Aging* 2014 Jun;35(6):1512.e5-1512.e10.
5. Wen Y, Han J, Chen J, **Dong J\***, Xia Y, Liu J, Jiang Y, Dai J, Lu J, Jin G, Han J, Wei Q, Shen H, Sun B, Hu Z. Plasma miRNAs as early biomarkers for detecting hepatocellular carcinoma. *Int J Cancer.* 2015 Oct 1;137(7):1679-90. (\*co-first author)
6. **Dong J**, Hu Z, Wu C, Guo H, Zhou B, Lv J, Lu D, Chen K, Shi Y, Chu M, Wang C, Zhang R, Dai J, Jiang Y, Cao S, Qin Z, Yu D, Ma H, Jin G, Gong J, Sun C, Zhao X, Yin Z, Yang L, Li Z, Deng Q, Wang J, Wu W, Zheng H, Zhou G, Chen H, Guan P, Peng Z, Chen Y, Shu Y, Xu L, Liu X, Liu L, Xu P, Han B, Bai C, Zhao Y, Zhang H, Yan Y, Amos CI, Chen F, Tan W, Jin L, Wu T, Lin D, Shen H. Association analyses identify multiple new lung cancer susceptibility loci and their interactions with smoking in the Chinese population. *Nat Genet.* 2012 Jul 15;44(8):895-9.
7. **Dong J**, Jin G, Wu C, Guo H, Zhou B, Lv J, Lu D, Shi Y, Shu Y, Xu L, Chu M, Wang C, Zhang R, Dai J, Jiang Y, Yu D, Ma H, Zhao X, Yin Z, Yang L, Li Z, Deng Q, Cao S, Qin Z, Gong J, Sun C, Wang J, Wu W, Zhou G, Chen H, Guan P, Chen Y, Liu X, Liu L, Xu P, Han B, Bai C, Zhao Y, Zhang H, Yan Y, Liu J, Amos CI, Chen F, Tan W, Jin L, Wu T, Hu Z, Lin D, Shen H.

- Genome-wide association study identifies a novel susceptibility locus at 12q23.1 for lung squamous cell carcinoma in han chinese. *PLoS Genet.* 2013;9(1):e1003190.
8. Hu Z\*, **Dong J**\*, Wang L, Ma H, Liu J, Zhao Y, Tang J, Chen X, Dai J, Wei Q, Zhang C, Shen H. Serum microRNA profiling and breast cancer risk: the use of miR-484/191 as endogenous controls. *Carcinogenesis.* 2012 Apr;33(4):828-34. (\*co-first author)
  9. **Dong J**, Dai J, Shu Y, Pan S, Xu L, Chen W, Wang Y, Jin G, Ma H, Zhang M, Hu Z, Shen H. Polymorphisms in EGFR and VEGF contribute to non-small-cell lung cancer survival in a Chinese population. *Carcinogenesis.* 2010 Jun;31(6):1080-6.
  10. Zhao C\*, **Dong J**\*, Jiang T\*, Shi Z, Yu B, Zhu Y, Chen D, Xu J, Huo R, Dai J, Xia Y, Pan S, Hu Z, Sha J. Early second-trimester serum miRNA profiling predicts gestational diabetes mellitus. *PLoS One.* 2011;6(8):e23925. (\*co-first author)
  11. **Dong J**, Hu Z, Shu Y, Pan S, Chen W, Wang Y, Hu L, Jiang Y, Dai J, Ma H, Jin G, Shen H. Potentially functional polymorphisms in DNA repair genes and non-small-cell lung cancer survival: A pathway-based analysis. *Mol Carcinog.* 2012 Jul;51(7):546-52.
  12. **Dong J**, Ren B, Hu Z, Chen J, Hu L, Dai J, Jin G, Xu L, Shen H. MDM2 SNP309 contributes to non-small cell lung cancer survival in Chinese. *Mol Carcinog.* 2011 Jun;50(6):433-8.
  13. Lu C\*, **Dong J**\*, Ma H, Jin G, Hu Z, Peng Y, Guo X, Wang X, Shen H. CCND1 G870A polymorphism contributes to breast cancer susceptibility: a meta-analysis. *Breast Cancer Res Treat.* 2009 Aug;116(3):571-5 (\*co-first author)
  14. **Dong J**, Dai J, Zhang M, Hu Z, Shen H. Potentially functional COX-2-1195G>A polymorphism increases the risk of digestive system cancers: a meta-analysis. *J Gastroenterol Hepatol.* 2010 Jun;25(6):1042-50. Review.

#### **PUBLICATIONS (Co-Author)**

---

1. Lesage S, Drouet V, Majounie E, Deramecourt V, Jacoupy M, Nicolas A, Cormier-Dequaire F, Hassoun SM, Pujol C, Ciura S, Erpapazoglou Z, Usenko T, Maura CA, Sahbatou M, Liebau S, Ding J, Bilgic B, Emre M, Erginel-Unaltuna N, Guven G, Tison F, Tranchant C, Vidailhet M, Corvol JC, Krack P, Leutenegger AL, Nalls MA, Hernandez DG, Heutink P, Gibbs JR, Hardy J, Wood NW, Gasser T, Durr A, Deleuze JF, Tazir M, Destée A, Lohmann E, Kabashi E, Singleton A, Corti O, Brice A; French Parkinson's Disease Genetics Study (PDG); **International Parkinson's Disease Genomics Consortium (IPDGC)**; International Parkinson's Disease Genomics Consortium IPDGC. Loss of VPS13C Function in Autosomal-Recessive Parkinsonism Causes Mitochondrial Dysfunction and Increases PINK1/Parkin-Dependent Mitophagy. *Am J Hum Genet.* 2016 Mar 3;98(3):500-13
2. Zheng M, Liu J, Zhu M, Yin R, Dai J, Sun J, Shen W, Ji Y, Jin G, Ma H, **Dong J**, Xu L, Hu Z, Shen H. Potentially functional polymorphisms in PAK1 are associated with risk of lung cancer in a Chinese population. *Cancer Med.* 2015 Nov;4(11):1781-7.
3. Ni B, Lin Y, Sun L, Zhu M, Li Z, Wang H, Yu J, Guo X, Zuo X, **Dong J**, Xia Y, Wen Y, Wu H, Li H, Zhu Y, Ping P, Chen X, Dai J, Jiang Y, Xu P, Du Q, Yao B, Weng N, Lu H, Wang Z, Zhu X, Yang X, Xiong C, Ma H, Jin G, Xu J, Wang X, Zhou Z, Liu J, Zhang X, Conrad DF, Hu Z, Sha J. Low-frequency germline variants across 6p22.2-6p21.33 are associated with non-obstructive azoospermia in Han Chinese men. *Hum Mol Genet.* 2015 Oct 1;24(19):5628-36.

4. Shen W, Yin R, Wang C, Zhu M, Zhou W, Qin N, Sun J, Liu J, **Dong J**, Jin G, Ma H, Hu Z, Shen H, Xu L, Dai J. Polymorphisms in alternative splicing associated genes are associated with lung cancer risk in a Chinese population. *Lung Cancer*. 2015 Sep;89(3):238-42.
5. Sun C, Chu M, Chen W, Jin G, Gong J, Zhu M, Yuan J, Dai J, Wang M, Pan Y, Song Y, Ding X, Du M, **Dong J**, Zhang Z, Hu Z, Wu T, Shen H. Genetic variants of H2AX gene were associated with PM2.5-modulated DNA damage levels in Chinese Han populations. *Mutat Res*. 2015 Aug;778:41-5.
6. Machiela MJ, Hsiung CA, Shu XO, Seow WJ, Wang Z, Matsuo K, Hong YC, Seow A, Wu C, Hosgood HD 3rd, Chen K, Wang JC, Wen W, Cawthon R, Chatterjee N, Hu W, Caporaso NE, Park JY, Chen CJ, Kim YH, Kim YT, Landi MT, Shen H, Lawrence C, Burdett L, Yeager M, Chang IS, Mitsudomi T, Kim HN, Chang GC, Bassig BA, Tucker M, Wei F, Yin Z, An SJ, Qian B, Lee VH, Lu D, Liu J, Jeon HS, Hsiao CF, Sung JS, Kim JH, Gao YT, Tsai YH, Jung YJ, Guo H, Hu Z, Hutchinson A, Wang WC, Klein RJ, Chung CC, Oh IJ, Chen KY, Berndt SI, Wu W, Chang J, Zhang XC, Huang MS, Zheng H, Wang J, Zhao X, Li Y, Choi JE, Su WC, Park KH, Sung SW, Chen YM, Liu L, Kang CH, Hu L, Chen CH, Pao W, Kim YC, Yang TY, Xu J, Guan P, Tan W, Su J, Wang CL, Li H, Sihoe AD, Zhao Z, Chen Y, Choi YY, Hung JY, Kim JS, Yoon HI, Cai Q, Lin CC, Park IK, Xu P, **Dong J**, Kim C, He Q, Perng RP, Kohno T, Kweon SS, Chen CY, Vermeulen RC, Wu J, Lim WY, Chen KC, Chow WH, Ji BT, Chan JK, Chu M, Li YJ, Yokota J, Li J, Chen H, Xiang YB, Yu CJ, Kunitoh H, Wu G, Jin L, Lo YL, Shiraishi K, Chen YH, Lin HC, Wu T, Wong MP, Wu YL, Yang PC, Zhou B, Shin MH, Fraumeni JF Jr, Zheng W, Lin D, Chanock SJ, Rothman N, Lan Q. Genetic variants associated with longer telomere length are associated with increased lung cancer risk among never-smoking women in Asia: a report from the female lung cancer consortium in Asia. *Int J Cancer*. 2015 Jul 15;137(2):311-9.
7. Yu J, Wu H, Wen Y, Liu Y, Zhou T, Ni B, Lin Y, Dong J, Zhou Z, Hu Z, Guo X, Sha J, Tong C. Identification of seven genes essential for male fertility through a genome-wide association study of non-obstructive azoospermia and RNA interference-mediated large-scale functional screening in *Drosophila*. *Hum Mol Genet*. 2015 Mar 1;24(5):1493-503.
8. Nalls MA, Bras J, Hernandez DG, Keller MF, Majounie E, Renton AE, Saad M, Jansen I, Guerreiro R, Lubbe S, Plagnol V, Gibbs JR, Schulte C, Pankratz N, Sutherland M, Bertram L, Lill CM, DeStefano AL, Faroud T, Eriksson N, Tung JY, Edsall C, Nichols N, Brooks J, Arepalli S, Pliner H, Letson C, Heutink P, Martinez M, Gasser T, Traynor BJ, Wood N, Hardy J, Singleton AB; **International Parkinson's Disease Genomics Consortium (IPDGC)**; Parkinson's Disease meta-analysis consortium. NeuroX, a fast and efficient genotyping platform for investigation of neurodegenerative diseases. *Neurobiol Aging*. 2015 Mar;36(3):1605.e7-12.
9. Coffee and Caffeine Genetics Consortium, Cornelis MC, Byrne EM, Esko T, Nalls MA, Ganna A, Paynter N, Monda KL, Amin N, Fischer K, Renstrom F, Ngwa JS, Huikari V, Cavadino A, Nolte IM, Teumer A, Yu K, Marques-Vidal P, Rawal R, Manichaikul A, Wojczynski MK, Vink JM, Zhao JH, Burlutsky G, Lahti J, Mikkilä V, Lemaitre RN, Eriksson J, Musani SK, Tanaka T, Geller F, Luan J, Hui J, Mägi R, Dimitriou M, Garcia ME, Ho WK, Wright MJ, Rose LM, Magnusson PK, Pedersen NL, Couper D, Oostra BA, Hofman A, Ikram MA, Tiemeier HW, Uitterlinden AG, van Rooij FJ, Barroso I, Johansson I, Xue L, Kaakinen M, Milani L, Power C, Snieder H, Stolk RP, Baumeister SE, Biffar R, Gu F, Bastardot F, Kutalik Z, Jacobs DR Jr, Forouhi NG, Mihailov E, Lind L, Lindgren C, Michaëlsson K, Morris A, Jensen M, Khaw KT, Luben RN, Wang JJ, Männistö S, Perälä MM, Kähönen M, Lehtimäki T, Viikari J, Mozaffarian D, Mukamal K, Psaty BM, Döring A, Heath AC, Montgomery GW, Dahmen N, Carithers T, Tucker KL, Ferrucci L, Boyd HA, Melbye M, Treur JL, Mellström D, Hottenga JJ, Prokopenko I, Tönjes A, Deloukas P, Kanoni S, Lorentzon M, Houston DK, Liu Y, Danesh J, Rasheed A, Mason MA, Zonderman AB, Franke L, Kristal BS; **International Parkinson's**

- Disease Genomics Consortium (IPDGC)**; North American Brain Expression Consortium (NABEC); UK Brain Expression Consortium (UKBEC), Karjalainen J, Reed DR, Westra HJ, Evans MK, Saleheen D, Harris TB, Dedoussis G, Curhan G, Stumvoll M, Beilby J, Pasquale LR, Feenstra B, Bandinelli S, Ordovas JM, Chan AT, Peters U, Ohlsson C, Gieger C, Martin NG, Waldenberger M, Siscovick DS, Raitakari O, Eriksson JG, Mitchell P, Hunter DJ, Kraft P, Rimm EB, Boomsma DI, Borecki IB, Loos RJ, Wareham NJ, Vollenweider P, Caporaso N, Grabe HJ, Neuhauser ML, Wolffenbuttel BH, Hu FB, Hyppönen E, Järvelin MR, Cupples LA, Franks PW, Ridker PM, van Duijn CM, Heiss G, Metspalu A, North KE, Ingelsson E, Nettleton JA, van Dam RM, Chasman DI. Genome-wide meta-analysis identifies six novel loci associated with habitual coffee consumption. *Mol Psychiatry*. 2015 May;20(5):647-56.
10. Nalls MA, Pankratz N, Lill CM, Do CB, Hernandez DG, Saad M, DeStefano AL, Kara E, Bras J, Sharma M, Schulte C, Keller MF, Arepalli S, Letson C, Edsall C, Stefansson H, Liu X, Pliner H, Lee JH, Cheng R; **International Parkinson's Disease Genomics Consortium (IPDGC)**; Parkinson's Study Group (PSG) Parkinson's Research: The Organized GENetics Initiative (PROGENI); 23andMe; GenePD; NeuroGenetics Research Consortium (NGRC); Hussman Institute of Human Genomics (HIHG); Ashkenazi Jewish Dataset Investigator; Cohorts for Health and Aging Research in Genetic Epidemiology (CHARGE); North American Brain Expression Consortium (NABEC); United Kingdom Brain Expression Consortium (UKBEC); Greek Parkinson's Disease Consortium; Alzheimer Genetic Analysis Group, Ikram MA, Ioannidis JP, Hadjigeorgiou GM, Bis JC, Martinez M, Perlmutter JS, Goate A, Marder K, Fiske B, Sutherland M, Xiromerisiou G, Myers RH, Clark LN, Stefansson K, Hardy JA, Heutink P, Chen H, Wood NW, Houlden H, Payami H, Brice A, Scott WK, Gasser T, Bertram L, Eriksson N, Foroud T, Singleton AB. Large-scale meta-analysis of genome-wide association data identifies six new risk loci for Parkinson's disease. *Nat Genet*. 2014 Sep;46(9):989-93.
  11. Hu Z, Li Z, Yu J, Tong C, Lin Y, Guo X, Lu F, **Dong J**, Xia Y, Wen Y, Wu H, Li H, Zhu Y, Ping P, Chen X, Dai J, Jiang Y, Pan S, Xu P, Luo K, Du Q, Yao B, Liang M, Gui Y, Weng N, Lu H, Wang Z, Zhang F, Zhu X, Yang X, Zhang Z, Zhao H, Xiong C, Ma H, Jin G, Chen F, Xu J, Wang X, Zhou Z, Chen ZJ, Liu J, Shen H, Sha J. Association analysis identifies new risk loci for non-obstructive azoospermia in Chinese men. *Nat Commun*. 2014 May 23;5:3857. doi: 10.1038/ncomms4857.
  12. Shi Y, Hu Z, Wu C, Dai J, Li H, **Dong J**, Wang M, Miao X, Zhou Y, Lu F, Zhang H, Hu L, Jiang Y, Li Z, Chu M, Ma H, Chen J, Jin G, Tan W, Wu T, Zhang Z, Lin D, Shen H. A genome-wide association study identifies new susceptibility loci for non-cardia gastric cancer at 3q13.31 and 5p13.1. *Nat Genet*. 2011 Oct 30;43(12):1215-8.
  13. Xie K, Liu J, Chen J, **Dong J**, Ma H, Liu Y, Hu Z. Methylation-associated silencing of microRNA-34b in hepatocellular carcinoma cancer. *Gene*. 2014 Jun 10;543(1):101-7.
  14. Zhang R, Chu M, Zhao Y, Wu C, Guo H, Shi Y, Dai J, Wei Y, Jin G, Ma H, **Dong J**, Yi H, Bai J, Gong J, Sun C, Zhu M, Wu T, Hu Z, Lin D, Shen H, Chen F. A genome-wide gene-environment interaction analysis for tobacco smoke and lung cancer susceptibility. *Carcinogenesis*. 2014 Jul;35(7):1528-35.
  15. Lan Q, Hsiung CA, Matsuo K, Hong YC, Seow A, Wang Z, Hosgood HD 3rd, Chen K, Wang JC, Chatterjee N, Hu W, Wong MP, Zheng W, Caporaso N, Park JY, Chen CJ, Kim YH, Kim YT, Landi MT, Shen H, Lawrence C, Burdett L, Yeager M, Yuenger J, Jacobs KB, Chang IS, Mitsudomi T, Kim HN, Chang GC, Bassig BA, Tucker M, Wei F, Yin Z, Wu C, An SJ, Qian B, Lee VH, Lu D, Liu J, Jeon HS, Hsiao CF, Sung JS, Kim JH, Gao YT, Tsai YH, Jung YJ, Guo H, Hu Z, Hutchinson A, Wang WC, Klein R, Chung CC, Oh IJ, Chen KY, Berndt SI, He X, Wu W, Chang J, Zhang XC, Huang MS, Zheng H, Wang J, Zhao X, Li Y, Choi JE, Su WC, Park KH,

- Sung SW, Shu XO, Chen YM, Liu L, Kang CH, Hu L, Chen CH, Pao W, Kim YC, Yang TY, Xu J, Guan P, Tan W, Su J, Wang CL, Li H, Sihoe AD, Zhao Z, Chen Y, Choi YY, Hung JY, Kim JS, Yoon HI, Cai Q, Lin CC, Park IK, Xu P, **Dong J**, Kim C, He Q, Perng RP, Kohno T, Kweon SS, Chen CY, Vermeulen R, Wu J, Lim WY, Chen KC, Chow WH, Ji BT, Chan JK, Chu M, Li YJ, Yokota J, Li J, Chen H, Xiang YB, Yu CJ, Kunitoh H, Wu G, Jin L, Lo YL, Shiraishi K, Chen YH, Lin HC, Wu T, Wu YL, Yang PC, Zhou B, Shin MH, Fraumeni JF Jr, Lin D, Chanock SJ, Rothman N. Genome-wide association analysis identifies new lung cancer susceptibility loci in never-smoking women in Asia. *Nat Genet*. 2012 Dec;44(12):1330-5.
16. Hu Z, Shu Y, Chen Y, Chen J, **Dong J**, Liu Y, Pan S, Xu L, Xu J, Wang Y, Dai J, Ma H, Jin G, Shen H. Genetic polymorphisms in the precursor MicroRNA flanking region and non-small cell lung cancer survival. *Am J Respir Crit Care Med*. 2011 Mar 1;183(5):641-8.
  17. Jin G, Ma H, Wu C, Dai J, Zhang R, Shi Y, Lu J, Miao X, Wang M, Zhou Y, Chen J, Li H, Pan S, Chu M, Lu F, Yu D, Jiang Y, **Dong J**, Hu L, Chen Y, Xu L, Shu Y, Pan S, Tan W, Zhou B, Lu D, Wu T, Zhang Z, Chen F, Wang X, Hu Z, Lin D, Shen H. Genetic variants at 6p21.1 and 7p15.3 are associated with risk of multiple cancers in Han Chinese. *Am J Hum Genet*. 2012 Nov 2;91(5):928-34.
  18. Hu L, Wu C, Zhao X, Heist R, Su L, Zhao Y, Han B, Cao S, Chu M, Dai J, **Dong J**, Shu Y, Xu L, Chen Y, Wang Y, Lu F, Jiang Y, Yu D, Chen H, Tan W, Ma H, Chen J, Jin G, Wu T, Lu D, Christiani DC, Lin D, Hu Z, Shen H. Genome-Wide Association Study of Prognosis in Advanced Non-Small Cell Lung Cancer Patients Receiving Platinum-Based Chemotherapy. *Clin Cancer Res*. 2012 Oct 1;18(19):5507-14.
  19. Chu M, Zhang R, Zhao Y, Wu C, Guo H, Zhou B, Lu J, Shi Y, Dai J, Jin G, Ma H, **Dong J**, Wei Y, Wang C, Gong J, Sun C, Zhu M, Qiu Y, Wu T, Hu Z, Lin D, Shen H, Chen F. A genome-wide gene-gene interaction analysis identifies an epistatic gene pair for lung cancer susceptibility in Han Chinese. *Carcinogenesis*. 2014 Mar;35(3):572-7.
  20. Dai J, Hu Z, Jiang Y, Shen H, **Dong J**, Ma H, Shen H. Breast cancer risk assessment with five independent genetic variants and two risk factors in chinese women. *Breast Cancer Res*. 2012 Jan 23;14(1):R17.
  21. Li N, **Dong J**, Hu Z, Shen H, Dai M. Potentially functional polymorphisms in ESR1 and breast cancer risk: a meta-analysis. *Breast Cancer Res Treat*. 2010 May;121(1):177-84.
  22. Dai J, Hu Z, **Dong J**, Xu L, Pan S, Jiang Y, Jin G, Chen Y, Shen H. Host immune gene polymorphisms were associated with the prognosis of non-small-cell lung cancer in Chinese. *Int J Cancer*. 2012 Feb 1;130(3):671-6.
  23. Dai J, Jin G, **Dong J**, Chen Y, Xu L, Hu Z, Shen H. Prognostic significance of survivin polymorphisms on non-small cell lung cancer survival. *J Thorac Oncol*. 2010 Nov;5(11):1748-54.
  24. Han J, Jiang T, Bai H, Gu H, **Dong J**, Ma H, Hu Z, Shen H. Genetic variants of 6q25 and breast cancer susceptibility: a two-stage fine mapping study in a Chinese population. *Breast Cancer Res Treat*. 2011 Oct;129(3):901-7.
  25. Zhang M, Hu Z, Huang J, Shu Y, Dai J, Jin G, Tang R, **Dong J**, Chen Y, Xu L, Huang X, Shen H. A 3'-untranslated region polymorphism in IGF1 predicts survival of non-small cell lung cancer in a Chinese population. *Clin Cancer Res*. 2010 Feb 15;16(4):1236-44.
  26. Chen X, Han S, Wang S, Zhou X, Zhang M, **Dong J**, Shi X, Qian N, Wang X, Wei Q, Shen H, Hu Z. Interactions of IL-12A and IL-12B polymorphisms on the risk of cervical cancer in Chinese women. *Clin Cancer Res*. 2009 Jan 1;15(1):400-5.

27. Chen W, Wang S, Tian T, Bai J, Hu Z, Xu Y, **Dong J**, Chen F, Wang X, Shen H. Phenotypes and genotypes of insulin-like growth factor 1, IGF-binding protein-3 and cancer risk: evidence from 96 studies. *Eur J Hum Genet*. 2009 Dec;17(12):1668-75.
28. Zhou X, Han S, Wang S, Chen X, **Dong J**, Shi X, Xia Y, Wang X, Hu Z, Shen H. Polymorphisms in HPV E6/E7 protein interacted genes and risk of cervical cancer in Chinese women: a case-control analysis. *Gynecol Oncol*. 2009 Aug;114(2):327-31.
29. Qian N, Chen X, Han S, Qiang F, Jin G, Zhou X, **Dong J**, Wang X, Shen H, Hu Z. Circulating IL-1beta levels, polymorphisms of IL-1B, and risk of cervical cancer in Chinese women. *J Cancer Res Clin Oncol*. 2010 May;136(5):709-16.
30. Ma H, Li H, Jin G, Dai J, **Dong J**, Qin Z, Chen J, Wang S, Wang X, Hu Z, Shen H. Genetic variants at 14q24.1 and breast cancer susceptibility: a fine-mapping study in Chinese women. *DNA Cell Biol*. 2012 Jun;31(6):1114-20.
31. Wu W, Hu Z, Qin Y, **Dong J**, Dai J, Lu C, Zhang W, Shen H, Xia Y, Wang X. Seminal plasma microRNAs: potential biomarkers for spermatogenesis status. *Mol Hum Reprod*. 2012 Oct;18(10):489-97.
32. Zhang M, Hu L, Shen H, **Dong J**, Shu Y, Xu L, Jin G, Tian T, Hu Z, Shen H. Candidate variants at 6p21.33 and 6p22.1 and risk of non-small cell lung cancer in a Chinese population. *Int J Mol Epidemiol Genet*. 2010;1(1):11-8.
33. Liu JB, Qiang FL, **Dong J**, Cai J, Zhou SH, Shi MX, Chen KP, Hu ZB Plasma DNA methylation of Wnt antagonists predicts recurrence of esophageal squamous cell carcinoma. *World J Gastroenterol*. 2011 Nov 28;17(44):4917-21.
34. Zhang XW, Pan SD, Feng YL, Liu JB, **Dong J**, Zhang YX, Chen JG, Hu ZB, Shen HB. [Relationship between genetic polymorphism in microRNAs precursor and genetic predisposition of hepatocellular carcinoma]. *Zhonghua Yu Fang Yi Xue Za Zhi*. 2011 Mar;45(3):239-43. Chinese.
35. Qian X, Cao S, Yang G, **Dong J**, Jin G, Shen Y, Hu Z. Variant genotypes of MDR1 C3435T increase the risk of leukemia: evidence from 10 case-control studies. *Leuk Lymphoma*. 2012 Jun;53(6):1183-7.
36. Wu W, Qin Y, Li Z, Dong J, Dai J, Lu C, Guo X, Zhao Y, Zhu Y, Zhang W, Hang B, Sha J, Shen H, Xia Y, Hu Z, Wang X. Genome-wide microRNA expression profiling in idiopathic non-obstructive azoospermia: significant up-regulation of miR-141, miR-429 and miR-7-1-3p. *Hum Reprod*. 2013 Jul;28(7):1827-36.

## **RESAERCH EXPERIENCE**

---

1. **Epidemiological investigations:** questionnaire design, face-to-face interviews, bio-samples collection and conservation and follow-up.
2. **Laboratory training:**
  - (1) Genomic experiment, such as extraction and purification of DNA in tissue, blood and serum/plasma, and SNPs screening and genotyping techniques (including PCR-RFLP, TaqMan and Openarray);
  - (2) Expression experiment, including extraction and purification of RNA (in tissue and plasma/serum) and protein, Expression analysis of mRNA/microRNA/protein by RT-PCR, Real-Time PCR and TaqMan Low Density Array or EMSA;
  - (3) Molecular biological techniques, including cell culture, reporter gene assay and MTT and

trans-well assay;

(4) Tissue/serum/plasma DNA methylation analysis, such as MSP.

3. **SNP data mining:** primer design and sequencing analysis, tagging SNPs selection using haploview and tagger, dissection of LD pattern and haplotype construction using LDA, PHASE, SAS, R and haploview, and SNP functional prediction using TRANSFAC, SIFT, UCSC, SNPinfo, PolyPen and PupaSuite.
4. **MicroRNA data mining:** primer design and sequencing analysis, and target gene searching using PicTar, miTarget, TargetsCan and miRTarBase.
5. **Statistical analysis:** general statistical analysis for SNP and expression data, genome-wide association analysis, gene-gene/gene-environmental interaction analysis, meta-analysis, cis-eQTL analysis, etc.
6. **Software:** statistical software (including EPIDATA, SPSS, SAS, Stata, R and Plink, etc.) and Microsoft Offices (including Word/Excel/PowerPoint and Adobe's Reader, etc.).
7. **Professional Training:** 2010 IARC summer school in Cancer Epidemiology

### **PATENTS IN CHINA**

---

1. Non-small cell lung cancer prognosis related serum/plasma miRNA markers and their applications (2009.08.28-2029.08.28 ZL200910184518.4)
2. Breast Cancer related serum/plasma miRNA markers and their applications (2010.09.08-2030.09.08 ZL 201010275557.8)
3. Cervical Cancer and CIN related serum/plasma miRNA markers and their applications (2010.09.08-2030.09.08 ZL 201010275552.5)