Children's Hospital of Philadelphia

Annual Progress Report: 2012 Formula Grant

Reporting Period

July 1, 2014 – June 30, 2015

Formula Grant Overview

The Children's Hospital of Philadelphia received $3,713,220 in formula funds for the grant award period January 1, 2013 through December 31, 2016. Accomplishments for the reporting period are described below.

Research Project 1: Project Title and Purpose

Functional Follow up of Genetic Commonalities to Diabetes and Cancer – Diabetes affects 18 million adults in the United States, with approximately 90 to 95 percent of those affected having type 2 diabetes (T2D). It is clear from recent genomic study observations that there is a specific yin and yang between cancer and T2D at the genetic level, including for the most strongly associated T2D gene, TCF7L2, which was discovered by the principal investigator. This has motivated us to systematically investigate the relationship between loci uncovered by genome wide association studies (GWAS) for cancer, leading us to ultimately carry out islet proliferation studies in mice, a mechanism which still largely eludes the diabetes research community but could revolutionize the way diabetes is treated if successful.

Anticipated Duration of Project

1/1/2013 – 12/31/2016

Project Overview

The repertoire of genes already established to play a role in the pathogenesis of type 2 diabetes (T2D) has grown substantially due to recent genome wide association studies (GWAS). In 2006, the P.I. on this application discovered the strong association of variants in the transcription factor 7 like 2 (*TCF7L2*) gene with T2D. Other investigators have already independently replicated this finding in different ethnicities and, interestingly, from the first GWAS of T2D in Caucasians, the strongest association was indeed with *TCF7L2*; this is now considered the most significant genetic findings in T2D to date.

Interestingly, there is also a very strong connection between *TCF7L2* and cancer. The key 8q24 locus found to be the most strongly associated genomic region with a number of cancers through GWAS contributes to the disease pathogenesis through mutation of an upstream TCF7L2-

Note: This project is funded, in part, under a grant with the Pennsylvania Department of Health. The Department specifically disclaims responsibility for any analyses, interpretations or conclusions.

binding element driving the transcription of the *MYC* gene. Indeed, it has been known for some years that *TCF7L2* harbors specific mutations that strongly influence colorectal cancer risk plus genomic sequencing of colorectal adenocarcinomas identified a recurrent *VTI1A-TCF7L2* gene fusion. Furthermore, many of the T2D GWAS-derived risk conferring alleles have been shown to protect against prostate cancer; in addition, *THADA, JAZF1* and *TCF2* are loci that have been strongly detected in separate GWAS analyses of prostate cancer and T2D. Thus, *TCF7L2* and other T2D associated genes also appear to be key players in cancer pathogenesis; however, this mechanism is still far from understood.

We previously performed ChIP-seq with this transcription factor to elucidate its binding repertoire genome wide. Unexpectedly, and despite employing a carcinoma cell line, the genes with TCF7L2 binding sites are strongly enriched in pathway categories related to metabolic-related functions and traits, further suggesting a role for metabolism in cancer.

We are taking forward the loci that are common to T2D and cancer GWAS outcomes to investigate their impact on cell proliferation with the ultimate goal of testing their role in beta-cell proliferation in mice, a mechanism which still largely eludes the diabetes research community.

Aim 1: Oligo-pull down / mass spec – characterize the set of proteins binding across rs7903146 and any allele differences in affinity.
Aim 2: Over-expression through constructs and under-expression through siRNA of *TCF7L2*, *THADA*, *JAZF1* and *TCF2* (also known as *HNF1B*) plus two other genes (based on expected finds from subsequent literature) in selected cell lines (colorectal and prostate, including HCT116) plus murine derived colon cells in order to assess influence of proliferation. Further we aim to also carry this out in an L-cell gut cell line, the human EndoC-βH1 cell line and a rodent beta-cell line to further explore influence on proliferation.
Aim 3: Somatic Gene Targeting (SGT) of *TCF7L2* in selected cell lines, both at the single rs7903146 SNP level and at the whole gene level - carried out internally with vendor generated constructs. Subsequent RNA-seq, proteomic studies and proliferation of SGT generated cell lines to fully investigate *TCF7L2* allele and gene effects in the cell lines altered. Also assess the influence of using medium from alpha, gamma and other cells in this context.
Aim 4: siRNA experiments with key genes resulting from pathway analyses to investigate how their perturbation influences gene expression in cancer, adipocyte and pancreatic derived cell lines.
Aim 5: Generation of mouse for best gene coming out of proliferation studies in year 1-3 to assess islet proliferation in this model system.

Principal Investigator

Struan F.A. Grant, PhD
Associate Professor, Department of Pediatrics
Children's Hospital of Philadelphia
Rm 1216F, 3615 Civic Center Boulevard
Philadelphia PA 19104-4318

Other Participating Researchers

Andrei Thomas-Tikhonenko, PhD; Adam Resnick, PhD; Qianghua Xia, PhD; Vanessa Guy, MS- employed by Children's Hospital of Philadelphia
Klaus Kaestner, PhD; Morris Birnbaum, MD, PhD; Tia Banks-Bernard; Jia Zhang, PhD – employed by University of Pennsylvania

Expected Research Outcomes and Benefits

The T2D *TCF7L2* association has an advantage, as the causal variant is widely thought to be identified, i.e., the work with multiple ethnicities has distilled down the association to a single variant, rs7903146, in intron 3. We will be applying the cutting edge approaches of somatic gene targeting in pre-selected cell lines and oligo pull-down combined with mass spectrophotometry in order to elucidate which proteins are binding across the SNP (plus look for allelic-specific binding differences) and determine what happens when you perturb the gene and even the actual base specifically in a cell model setting – this will in turn shed light on the mechanism by which TCF7L2 exerts its effect on T2D risk, giving us crucial insights in to the genetic architecture of the disease.

siRNA experiments with key genes resulting from pathway analyses related to Akt will be used to investigate how their perturbation influences gene expression in cancer, adipocyte and pancreatic derived cell lines to understand more precisely the classical pathways for cancer and diabetes.

Over-expression of *TCF7L2*, *THADA*, *JAZF1* and *TCF2* in key cell lines will help us determine which influence proliferative potential which in turn will inform us as to which gene(s) to take forward in to a mouse model to check for beta-cell proliferation. If successful this would have a fundamental impact on the way type 2 diabetes is both treated and prevented in the future. Furthermore, it will inform us on how metabolism contributes to the etiology of cancer.

Summary of Research Completed

MILESTONE(S) FOR 7/1/2014-6/30/2015:
Aim 1: COMPLETED LAST YEAR i.e. Oligo-pull down / mass spec – characterize the set of proteins binding across rs7903146 and any allele differences in affinity.

Aim 2: Completed further over-expression and under-expression experiments of the same genes in an L-cell gut cell line, the EndoC-βH1 cell line and a rodent beta-cell line to further explore influence on proliferation.

Cell culture and siRNA transfections
Human NCM460 cells were obtained through a Material Transfer Agreement with INCELL and grown in M3Base Medium supplemented with 10% fetal bovine serum, 100 U/ml penicillin, and 100 µg/ml streptomycin. EndoC-βH1 cells were obtained through a Material Transfer

Agreement with ENDOCELLS and grown in low glucose DMEM with 2% BSA Fraction V, 50 µM beta-mercaptoethanol, 10 mM nicotinamide, 5.5 µg/ml transferrin, 6.7 ng/ml sodium selenite, and 100 µg/ml streptomycin. Cell culture dishes used to grow EndoC-βH1 cells were coated with 10 µg fibronectin and 500 µg extracellular matrix gel per 10 cm dish. Cells were transfected with siRNAs at a final concentration of 75 nmol/L using Lipofectamine RNAiMAX (Invitrogen). Nontargeting control and JAZF1 smartpool siRNAs were purchased from Dharmacon.

Expression of target genes in EndoC-βH1 cells
Before manipulating *TCF7L2*, *THADA*, *HNF1B* and *JAZF1*, we chose to assess levels of their expression in a non-carcinoma colorectal cell line (NCM460) and a human beta cell line (EndoC-βH1). As shown in Figure 1 Left Panel, *TCF7L2* is highly expressed in both NCM460 cells, while *THADA*, *HNF1B* and *JAZF1* are highly expressed in EndoC-βH1 cells.

Manipulation of T2D/cancer implicated genes
Since *JAZF1* is highly expressed in EndoC-βH1 cells and has been previously implicated T2D GWAS, we decided to deplete its expression using a distinct siRNA smartpool in order to assess its effect on gene expression. As shown in Figure 1 Right Panel, the JAZF1 smartpool siRNAs significantly depleted expression of *JAZF1* as detected by q-RT-PCR using two distinct primer pairs to assess relative expression of *JAZF1* mRNA. RNA samples from *JAZF1* knockdown samples are queued for gene expression analysis by microarray. Once we have data back from this order, we will proceed with the other genes in a similar manner.

Aim 3: Completed Somatic Gene Targeting (SGT) of TCF7L2 in an L-cell gut cell line and in the EndoC-βH1cell line, both at the single rs7903146 SNP level and at the whole gene level.

Regulation of TCF7L2 expression via the genomic region immediately spanning rs7903146
The gut cell line, HCT116, is a C/T heterozygote for SNP rs7903146. In order to investigate how rs7903146 may influence TCF7L2 function and general gene expression, we elected to carry out CRISPR/cas9 procedures by working with sgRNAs that targeted a genomic region spanning 700bp upstream and downstream of rs7903146 in order to delete either allele. We successfully identified three clones, namely 'hap-T' (haploid T allele), 'hap-C' (haploid C allele) and 'dip-del' (diploid knock-out, both alleles deleted).

TCF7L2 protein levels were dramatically reduced in the hap-C cells compared to the wild type heterozygous HCT116 (C/T) cells; in contrast, TCF7L2 protein levels in the hap-T allele cells were higher than for the wild type cells. TCF7L2 protein levels in the homozygote deletion (dip-del) were slightly decreased compared to the wild-type cells, but higher than in hap-C cells (Figure 2 Left Panel).

Furthermore, it was extensively reported that TCF7L2 binds to the extreme upstream enhancer region of the *MYC* oncogene in colorectal and prostate cancer. Indeed, consistent with TCF7L2 levels, MYC protein levels were reduced in the hap-C cells but slightly increased in the hap-T cells (Figure 2 Left Panel).

To validate our overall findings, we generated cDNA from both the control and targeted cells and carried out quantitative PCR to determine the respective levels of *TCF7L2* mRNA in these

cells. Using probes from each individual exon along *TCF7L2*, we observed that deletion of the T allele strikingly reduced expression of the vast majority of the exons (two-tailed $P<0.01$), while conversely deletion of the C allele slightly increased expression of most exons, especially those located 3' to rs7903146 (Figure 2 Right Panel). To further evaluate the respective effects of the C and T alleles on *TCF7L2* expression levels, we normalized expression levels for each exon based on the wild type C/T cells. The haploid T allele led to an increase in *TCF7L2* gene expression while the haploid C allele decreased expression. Interestingly, the expression changes varied between exons, with the most obvious difference being when moving between exon 4 and exon 5.

Dipeptidyl peptidase-4 (DPP-4) is an enzyme that plays a fundamental role in glucose metabolism via gut signaling and a key target for multiple pharmacological programs in the T2D field. The two major incretins, glucagon-like peptide-1 (GLP-1) and gastric inhibitory peptide (GIP), are both degraded by DPP-4. Given the known TCF7L2 link with proglucagon gene regulation, we queried our data for a correlation with *DPP-4* gene expression levels. Hap-T yielded higher DPP-4 levels than with hap-C (two-tailed $P<0.01$).

The region harboring rs7903146 regulates chromatin accessibility

FAIRE (formaldehyde-assisted isolation of regulatory elements) is an approach to study chromatin accessibility and to identify regulatory elements. In a recent study from human islet samples, the risk T allele showed greater abundance than the non-risk C allele in the open chromatin fraction. Taking advantage of our successfully isolated haploid C or T allele clones in HCT116 cells, we used FAIRE to study chromatin accessibility around this region. As shown in Figure 3, hap-C exhibited greater abundance in the open chromatin fraction than hap-T. To study the activity and regulation of promoters, we cloned the 1.8kb elements with the C or T allele and 0.4 kb element upstream of the luciferase reporter gene. By dual luciferase assay, we observed that the T allele yielded slightly higher (albeit not statistically significant) activity than the C allele in forward and reverse orientation.

The rs7903146 C allele also has an inhibitory effect on TCF7L2 expression in Miapaca2 cells

The human pancreatic EndoC-βH1 cell line grows and divides very slowly and thus making it extremely challenging for CRISPR-cas9 approaches. In addition, it is becoming increasingly clear in the community that the TCF7L2 protein is not detectable in human pancreatic β cells. We have therefore turned our focus to other key tissues instead. Given that proglucagon is produced by only two cell types – enteroendocrine L cells and pancreatic alpha cells - we elected to follow up our observations in the gut HCT116 cells with a similar investigation in Miapaca2 cells, which are pancreatic carcinomas and are considered to 'alpha cell like', and have already been successfully used to study insulin and glucagon production previously. Miapaca2 is homozygous for C allele, so we utilized the CRIPR-cas9 approach to target the cells and successfully identified a clone namely M5 in which both C alleles were removed. Interestingly, M5 cells showed increased TCF7L2 at both the RNA and protein levels with real time PCR and Western blotting, respectively (Figure 4). It should be noted that these results are consistent with those observed in HCT116 cells described above.

The region harboring rs7903146 forms long-range functional connections with ACSL5
GWAS only reports genomic signals associated with a given trait and not necessarily the precise localization of culprit genes. Chromatin conformation capture techniques can aid in the identification of causal genes by characterizing genomic regions that make physical contact with a GWAS-implicated locus. Given that we previously successfully identified in Aim 1 a PARP-1 related protein complex binding across rs7903146, we elected to leverage the CRISPR-cas9 data and intersect it with chromatin conformation capture data in order to get a better insight in to the downstream effects of this SNP is and on which targets PARP-1 may be principally impacting.

As such, we carried out 4C-seq and Capture C in parallel libraries using the immediate sequence harboring rs7903146 as the bait to elucidate the genomic regions it interacted with in the NCM460 gut-derived cell line. When overlapping both sets of data, the bait region consistently interacted with 5 promoters, of which 4 resided in the same topologically associating domain (TAD) as *TCF7L2*, namely - and in order of peak score strength - *ACSL5, HABP2, LOC143188* and *TDRD1* (the latter being in a separate sub-TAD); furthermore, we observed interaction within *TCF7L2* itself. An additional promoter was observed on chromosome 6 corresponding to *MMS22L*.

Leveraging CRISPR/Cas9-generated 'dip-del' HCT116 cell line described above, we observed a particularly dramatic impact on *ACSL5* gene expression levels (approx. 30x decrease), and subsequent Western blotting revealed that protein levels were almost entirely ablated. Furthermore, the gene expression changes for *HABP2* were also very notable. Interestingly, *ACSL5* encodes an enzyme known to play a role in mammalian fatty acid metabolism while *HABP2* has not been strongly implicated in metabolic processes previously.

As such, our data point to the immediate genomic location harboring rs7903146 as being a putative locus control region for a number of genes playing a role in the pathogenesis of T2D, including *TCF7L2, ACSL5* and *HABP2.*

Aim 4: COMPLETED LAST YEAR i.e. siRNA experiments with key genes resulting from pathway analyses to investigate how their perturbation influences gene expression in cancer, adipocyte and pancreatic derived cell lines.

Note: Aim 5 not due to start until 7/1/2015 (Rodent-related work)

Figure 1: Left Panel: Expression of target genes in colorectal and beta cell lines. Total cell lysates were obtained from NCM460 and EndoC-βH1 cells. Cell lysates were analyzed by western blotting as indicated; Right Panel: Manipulation of T2D/cancer implicated genes in human beta cells. EndoC-βH1 cells were transfected with non-targeting control siRNA (Ctrl) or JAZF1 smartpool siRNA sequences. After 96 hours, cells were lysed, and lysates were analyzed by q-RT-PCR as indicated.

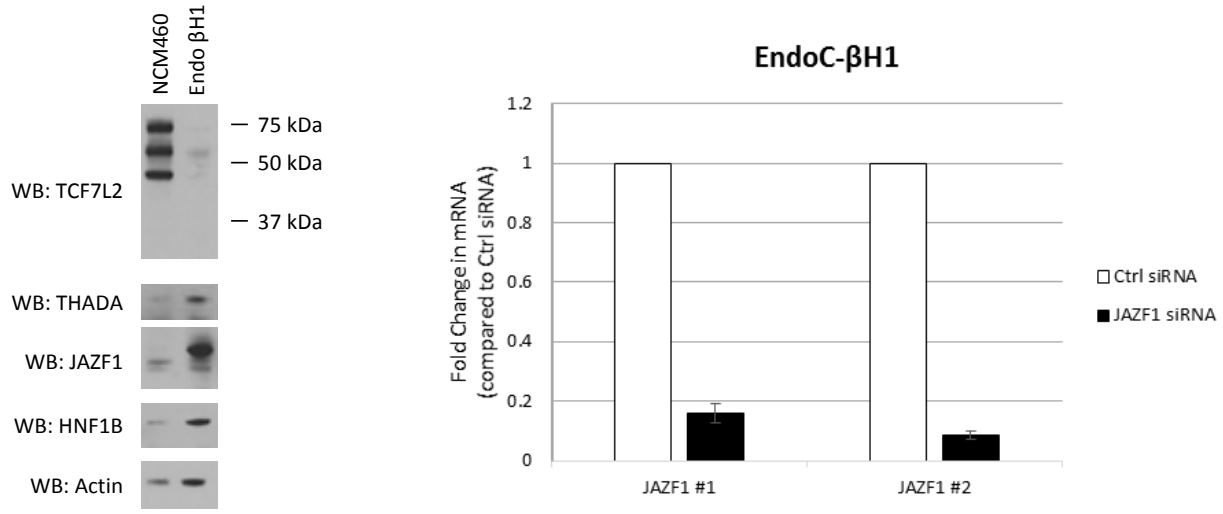


Figure 2. Expression analysis following gene editing of the genomic region harboring rs7903146. Left Panel: Expression of TCF7L2 and c-Myc protein in the control and targeted cells determined by Western blot; Right Panel: Relative levels of individual *TCF7L2* exon expression in the control and targeted cells, determined by quantitative PCR with GAPDH normalization.

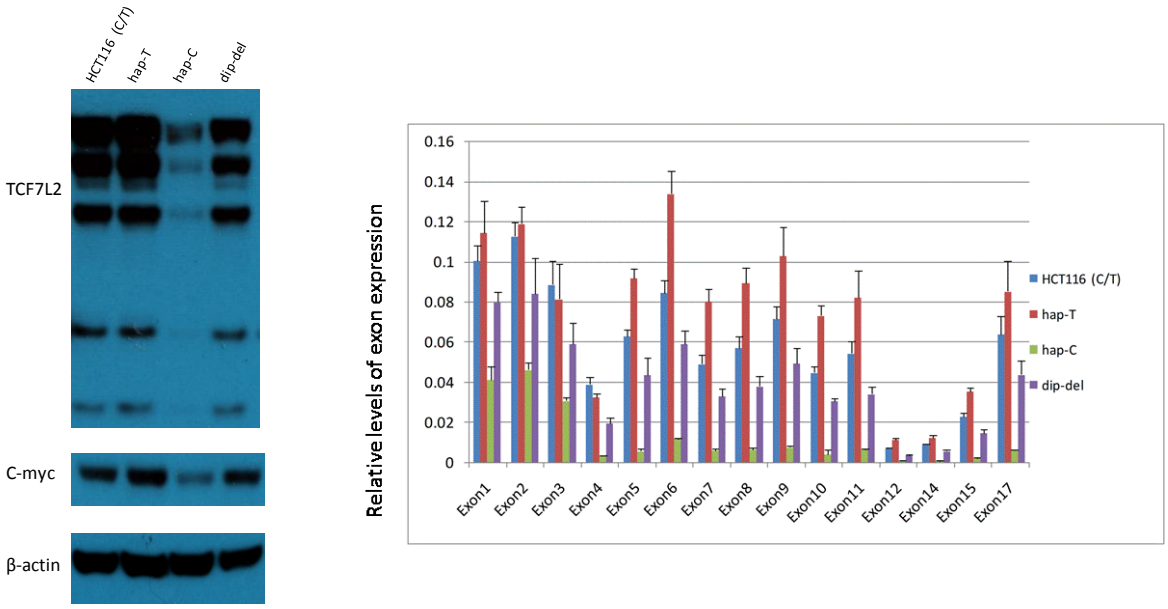


Figure 3. SNP rs7903146 regulates chromatin accessibility. PCR analysis of FAIRE-processed chromatin from control and targeted cells

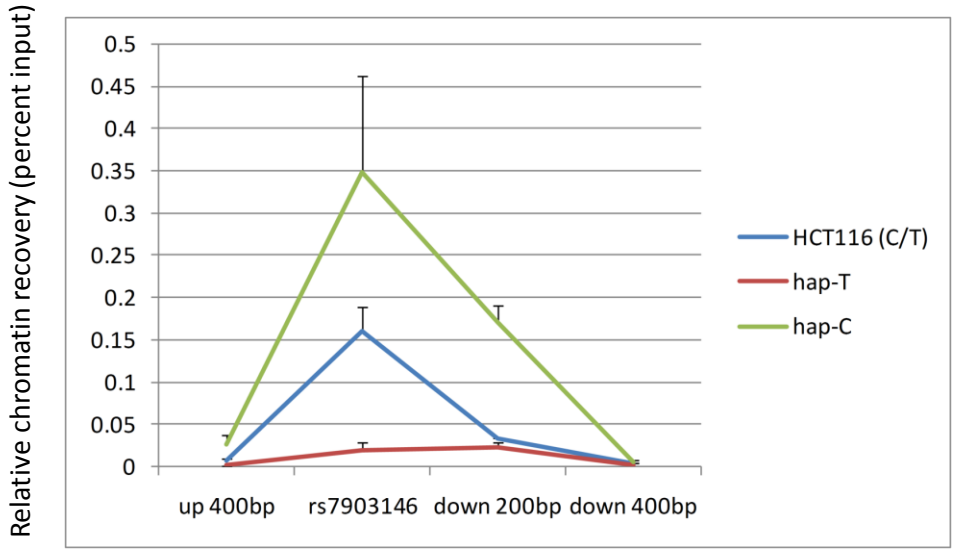
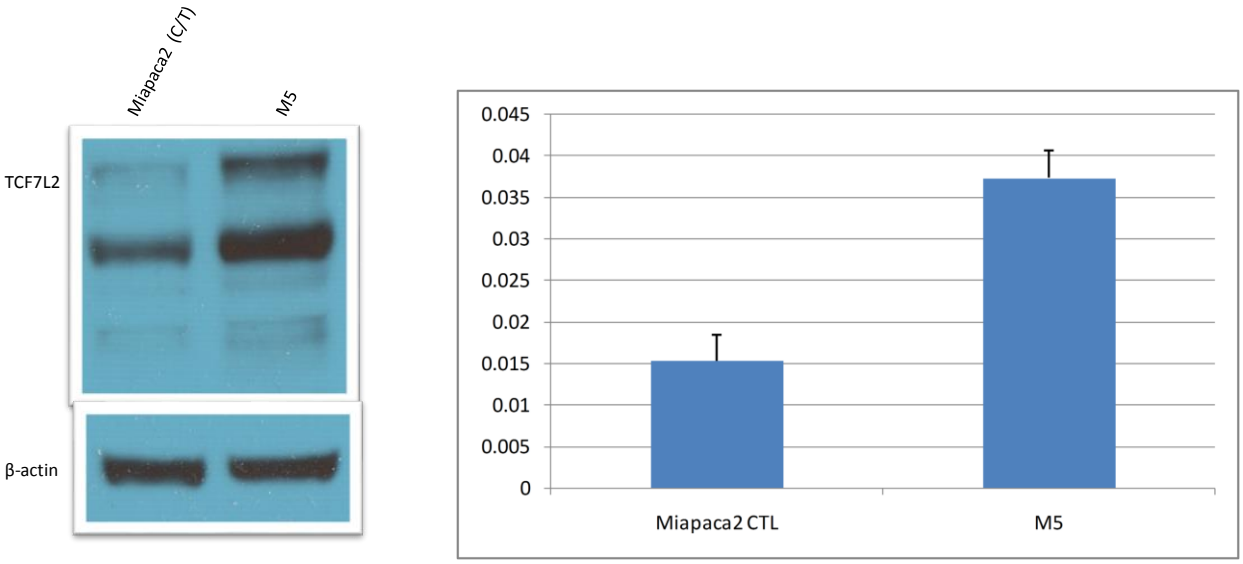


Figure 4. Expression analysis following gene editing of the genomic region harboring rs7903146 in Miapaca2. Left Panel: Expression of TCF7L2 in the control and targeted cells determined by Western blot; Right Panel: Relative levels of individual *TCF7L2* exon expression in the control and targeted cells, determined by quantitative PCR with GAPDH normalization.



Research Project 2: Project Title and Purpose

Comparative Effectiveness of Developmental-Behavioral Screening Instruments in Children – Evidence of the advantages of using the Survey of Wellbeing of Young Children (SWYC) to screen for childhood behavioral-developmental disorders is accumulating. Findings are limited to English speaking, well-educated, predominantly U.S.-born families. Whether the findings are generalizable to other children from two groups rapidly growing in size in Pennsylvania-- Hispanic families lacking English language proficiency and African-born Black women-- is unclear. Consequently we propose to recruit a sample of Hispanic children and a sample of children of African-born Black women and conduct a systematic comparison study of the SWYC to other instruments widely used to identify behavioral-developmental disorders in children.

Anticipated Duration of Project

1/1/2013 – 12/31/2016

Project Overview

The overarching objective of the proposed project is to build upon existing research which systematically compares the accuracy, feasibility, and effectiveness of pediatric behavioral and developmental screening instruments. In brief, we will extend and deepen this research by recruiting families from linguistically/culturally distinct populations (of particular importance to pediatric providers and public health officials in Pennsylvania) which heretofore have *not* been included in this research. The first specific aim, is to assess the accuracy (specificity; sensitivity; positive and negative predictive values) of the Ages and Stages Questionnaires (ASQ-SE and ASQ-3), M-CHAT and the SWYC set of screening instruments for the above mentioned sub-populations. The second specific aim is to integrate all screening data collected with relevant, supplemental self-reported survey data (behavioral, socioeconomic and psychosocial) in order to *develop and evaluate a predictive model related to the detection of developmental and behavioral disorders in the study population(s*). The essential components of the research design are as follows: 1) recruit parent/child pairs across 9 child age categories (9 to 60 months) at pediatric/prenatal providers and community centers throughout the city of Philadelphia; 2) obtain informed consent, administer the Ages and Stages, M-CHAT, and the SWYC screening instrument components and collect sociodemographic, psychosocial and behavioral data; 3) conduct complete clinical behavioral and developmental evaluations on at least 356 of the enrollees providing the 'gold standard' for assessing the accuracy of the screening instruments described above; and 4) use advanced statistical techniques on all data (screening, clinical evaluation and survey), in order to identify the best predictive model of pediatric behavioral and developmental disorders.

Principal Investigator

Jennifer F Culhane, PhD, MPH
Senior Investigator
The Children's Hospital of Philadelphia
34th and Civic Center Blvd
CHOP North-Suite 880
Philadelphia, PA 19104

Other Participating Researchers

Sandra R Wadlinger, MS, RRT; Marsha Gerdes, PhD; Leny Mathew, MS; David Webb, PhD; Ximena Cuellar, BS, Karen B. Friedman, PhD; Sherry Winston, BA; Delitza Hernandez, BS – employed by The Children's Hospital of Philadelphia

Expected Research Outcomes and Benefits

The number of professional organizations and state governments that are recommending or mandating systematic screening for pediatric developmental and behavioral problems is growing. The use of standardized screening instruments to detect these disorders is thus rapidly increasing in pediatric care settings. The proportion of pediatricians who reported using a standardized instrument to screen for developmental problems rose from 23% to 47.7% over a five year period between 2002 and 2009. Unfortunately, because of limited research pediatricians make decisions about screening from a fairly weak base of evidence to select an appropriate tool that is inexpensive, brief, and easy to administer and score. However, financial considerations and ease of use should not trump considerations of the accuracy and validity of screening tools, since these choices will have significant implications for children and their families. Additionally, screening tool selection could have profound implications for public health costs and policies as well as individual child life-course outcomes. If only a small fraction of pediatricians were to use a screen that has slightly better accuracy, the number of children correctly identified with a developmental-behavioral disability would increase considerably, facilitating early intervention and this improved prognosis, while the number of incorrectly identified children (false positives) would be greatly diminished.

Summary of Research Completed

Project Milestones (for period 7/1/14-6/30/15):
1. Achieve total enrollment of 840 subjects.
2. Enduring Executive Steering Committee Data Quality Meetings

<u>Specific Aim 1</u>

Specific Aim 1 is to assess the accuracy (specificity; sensitivity; positive and negative predictive values) of the Ages and Stages Questionnaire-Social-Emotional (ASQ-SE) and Ages and Stages Questionnaire-Third Edition (ASQ-3), the Modified Checklist for Autism in Toddlers (M-

CHAT) and the Survey of Well-Being of Young Children (SWYC) screening instruments in a sample of Hispanic children and children of African-born Black women. Progress towards this aim was achieved during this project period by continuing a robust recruitment campaign, successfully screening and enrolling eligible study subjects from a variety of community locations and the completion of study evaluations in these children of Spanish speaking and African-born Black parents. Study progress and data quality was closely monitored through regular Executive Steering Committee Data Quality meetings.

Recruitment, Screening and Evaluation of Children

During this reporting period, we added multiple recruitment locations including three Women, Infant and Children (WIC) Offices in Philadelphia, the KenCrest Early Learning Center in South Philadelphia, the Norris Square Community Alliance (NSCA) and Casa del Carmen Family Service Center in North Philadelphia, and Women and Children's Health Services in Center City Philadelphia. This expansion of recruitment locations has allowed us to exceed our recruitment milestone for this project period while maintaining strong support for this scientific endeavor in the community.

To date, the Study Team has approached 1140 families, 712 in this project period. Of the total enrolled, 151 (13%) were ineligible (based on language, child age requirements and/or previous participation in the study) and 143 (13%) indicated that they did not wish to participate. The most common concern expressed by families who declined to participate was the inability to commit the amount of time needed to complete study visits and none raised an objection to the nature of the project itself. Of the remaining 989 eligible families, 846 (86%) agreed to participate, completed the initial study visit and completed the screening questionnaires (ASQ-SE, ASQ-3, SWYC and M-CHAT) and the supplemental family survey.

We have exceeded our stated milestone to consent and enroll 840 participants by June 30, 2015. Of the participants who completed the first study visit, 394 were invited for a second visit, and 296 (83%) have completed this clinical evaluation. An additional 36 are scheduled for this evaluation in the next three weeks. Given these results and the successful inroads we have made into some of the major community centers in the Hispanic community, and the remarkable support and interest we have received from the Hispanic community in and around Philadelphia, we expect to achieve our goal of recruiting 1259 total subjects.

Executive Steering Committee Meetings

Dr. Marsha Gerdes was added to the Executive Steering Committee at the beginning of the project period. The full Committee (Culhane, Webb, Gerdes, Mathew and Cuellar) continued to meet biweekly during this fiscal year and reviewed recruitment metrics, data quality, data accuracy and data collection strategies in addition to developing strategies to assure that all study milestones are achieved. The Committee carefully reviews the outreach and recruitment strategies deployed and examines these activities relative to the identification of eligible families, study participation and refusal rates and the successful retention of study participants for the duration of study activities. As appropriate, preliminary study specific data are reviewed to compare results obtained at the initial screening and data collection assessment with those data

obtained as part of the second clinical evaluation visit. As a result of these meetings, our procedures have been amended to assure continued success of the project.

Revisions to Study Protocol

Based on preliminary data and after gaining approval from The Children's Hospital of Philadelphia (CHOP) Institutional Review Board (IRB), the study protocol was revised during the current reporting period. Changes included the addition of a follow up phone call for those families that had children identified as being at risk for developmental and/or behavioral problems at the clinical evaluation visit, an increase in both the number of children eligible to receive the clinical evaluation and the total number of subjects enrolled in the study. A detailed discussion of these changes is below:

- Additional Follow Up Phone Call

  This supplementary follow up procedure, which was added in July 2014, adds a phone call approximately 3 months after the completion of the second visit feedback phone call for participants deemed to be at risk for developmental delay and/or behavioral problems. This call is performed by the Early Childhood Assessor that completed the clinical evaluation with the child and provided the initial feedback to the parent after the visit. As a result, the maximum time that an individual study subject would potentially participate in the study was increased from 4 weeks to 6 months.

  Contacting the parent approximately 3 months after clinical findings are communicated will allow us to evaluate the efficacy of our initial feedback procedures, assess whether the participant completed any additional follow-up steps as a result of the feedback received and to inform researchers concerning factors that may facilitate or present barriers when seeking specific developmental and behavioral health follow-up in this population. Using a standardized questionnaire, we collect data about the steps taken to initiate further evaluation and/or treatment for their child, which specific services were sought and/or are being provided (ie. early childhood intervention services) and if a medical diagnosis was reported to the participant.

- Increase in Study Subject Enrollment

  After careful examination of key study metrics and a review of data collected, we have updated both our total study enrollment goal and the number of children that will receive the second clinical evaluation visit. Data regarding the proportion of children who had screened positive (failure rate) on at least one assessment measure for behavioral problems/and or development delays at the first study visit showed that approximately 60% of all children had failed at least one screening instrument, a rate considerably higher than we had initially estimated (35%). Conversely, the data showed a lower number of children that failed no instruments than we had originally estimated. These findings caused us to reassess the number of clinical evaluations warranted to best compare, based on both the *specificity* and *sensitivity* criteria, the results of the screening instruments used at the first evaluation visit to

those used as part of the second clinical evaluation. Given our ability to accommodate a higher number of clinical evaluation visits than had been originally planned, and our goal to obtain reliable and accurate estimations of the *sensitivity* and *specificity* of the instruments used in the study, we increased our enrollment goal from 980 to 1259 children.

We carefully considered our capacity to accommodate an increased number of clinical evaluations and concluded that we could successfully complete 655 clinical evaluations over the lifetime of the project, an increase over the originally planned 356. We have completed 296 clinical evaluation visits to date with an additional 36 scheduled. Based on our current average of completing 34 visits per month, we will finish all visits by early May 2016, allowing us to complete all data collection activities, data quality checks, data entry and analyses within the overall project period.

Specific Aim 2

The second Specific Aim is to integrate all screening data collected with relevant, supplemental self-reported survey data (behavioral, socioeconomic and psychosocial) in order to *develop and evaluate a predictive model related to the detection of developmental and behavioral disorders in young children who enrolled in the study*. Significant progress toward this aim was achieved by successfully consenting and administering first visit screening instruments inclusive of the parental survey questionnaire to a total of 846 children and their mothers. (Table 1). Data from the supplemental interviews completed to date demonstrates that 95% of the mothers are foreign born (4% born in Puerto Rico) and undocumented. Of the 16 sending countries, the top four are Mexico (77%), Dominican Republic (5%), Guatemala (3%), and Honduras (2%). These women are married (87%), are multiparous (72.2%) and report low educational attainment (59% less than HS). Over 75% reported using WIC, food stamps, and Medicaid services and a third reported receiving free school lunch. Mothers also report significant trouble reading, writing, and understanding English (75%) high stress associated with interactions with service providers or government officials (30.3%) and grave concerns over deportation (44.3%). Less than 40% evaluate their current physical and mental health as excellent or very good, and 20% are possibly depressed as measured by the Center for Epidemiological Studies (CES-D) with a score greater than 16. In sum, children enrolled in this project are vulnerable with undocumented immigrant mothers with uniformly low levels of SES and educational attainment and high rates of self-reported poor health and stress (Table 2). Once enrollment is complete, we will have the opportunity for further analyses in order to examine these factors in relation to the clinical evaluation results and identify the best predictive model relative to developmental and behavioral disorders in this study population.

Table 1. Study Subject Age Distribution

Age of child (Months)	9-11	12-14	15-17	18-23	24-29	30-35	36-47	48-60	(Total)
N (%)	102 (12.1)	47 (5.6)	42 (5.0)	89 (10.5)	115 (13.6)	84 (9.9)	182 (21.5)	185 (21.8)	846

Table 2: Characteristics of the Families

Select Maternal Characteristics	*%*
Household income per year <30,000	57.7
Family size supported by income: median(range)	4 (1,12)
Father's education: less than high school	64.7
Child care in day care centers	27.8
Health care provider reported special needs	16.7
High discrimination: treated with less courtesy	15.5
High discrimination: poor service in stores/restaurants	10.5
High discrimination: low public perception of smartness	14.6
High discrimination: low public perception of social status	23.5
High stress: difficulty to find work due to immigrant status	43.7
High stress: Worry about legal status	31.6
High stress: Worry about deportation	44.3

Changes in Study Subcontract Agreements

The subcontract agreement to the University of Pennsylvania for efforts by Dr. Irma Elo and Dr. Emilio Parrado was terminated during this reporting period. Both Dr. Elo and Dr. Parrado completed the work reflected in the Statement of Work (SOW) relative to this agreement. Their work was related to finalizing data collection instruments and analytic plans in addition to helping with community outreach to facilitate the recruitment of study participants. The instrument and analytic plans are complete and the remaining outreach and recruitment work is being successfully accomplished by CHOP staff members. There is no change to the overall scope of work for the study as a result of this termination and the project will continue to proceed as originally planned.