Next Generation Sequencing of autism genes in the Lebanese population and their functional evaluation

Jihane Soueid[1], Jamal Al Ali[1], Dikran R. Guisso[1], Maria Charhrour[2], Rose-Mary Boustany[1,3]

[1]Department of Biochemistry and Molecular Genetics, American University of Beirut Medical Center,. [2] McDermott Center for Human Growth and Development / Center for Human Genetics, Department of Neuroscience, Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, Texas. [3]Neurogenetics Program and Division of Pediatric Neurology, Department of Pediatrics and Adolescent Medicine, American University of Beirut Medical Center.

Introduction

Autism spectrum disorders (ASDs) are a group of neurodevelopmental disorders characterized by marked difficulties in social and verbal communication, and behavior. In Lebanon, ASD prevalence is 1/68 children. Recent findings by our group support a heterogeneous genetic etiology, including rare *de novo* and inherited mutations, chromosomal rearrangements, as well as double hit mutations. Only a fraction of autism genes have been discovered world-wide. Applying whole exome sequencing (WES) to Lebanese families in order to identify mutations is extremely useful in identifying recessive causes of autism.

AIMS:

- **Uncover novel autism risk genes in the Lebanese population**
- **Characterize functional impact of rare variants in candidate genes**
- **Investigate effect of a novel frameshift mutation within the Ubiquitin-like domain-containing C-terminal domain phosphatase 1 (*UBLCP1*) gene.**

Methods

Materials:

Eight Lebanese families were chosen because of consanguinity or having 2 affected children. These families did **not** show any pathogenic aberrations with Affymetrix microarray genechip analysis.

Whole exome sequencing and validation:

Whole-exome sequence (Illumina HiSeq) was obtained with a mean target coverage of 90% at ≥20× and a mean read depth of 106X. Variants were validated by **Sanger sequencing** in the patient and the available family members and in normal Lebanese controls (n=104).

Quantitative real-time PCR:

qPCR was performed in triplicate using specific primers and the iQTM SYBR® Green Supermix (BioRad). Melt curve analysis was applied and all results were normalized to GAPDH level and calculated using the ΔΔ CT method.

UBLCP1 subcloning:

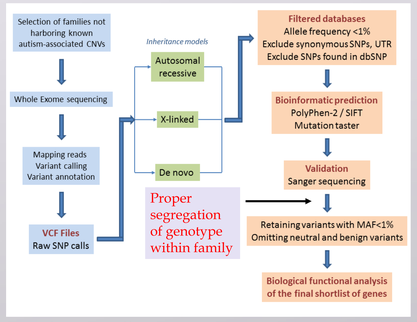
Normal and mutated UBLCP1 were each subcloned into a pCDNA3.1-myc-HisA vector, containing CMV promoter, a neomycin selection marker, and a C-terminal tag encoding a polyhistidine metal-binding peptide for rapid purification on nickel-chelating resin.

Cell culture and transfection:

PC12 cells were grown in DMEM supplemented with 10% FBS, 5% horse serum, 1% sodium pyruvate in 5% CO2 at 37 °C. Transfections were performed with attractene reagent (Qiagen). Stably transfected cell lines were selected by culturing transiently transfected cells in the presence of 500 µg/µl of G418 for 14 days.

WES analysis

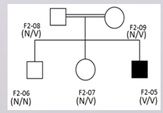
Our lab conducted the first genome-wide CNV association study in autism patients from 36 Lebanese ASD families screening the genome for microdeletions and microduplications. Families that showed no evidence of any pathogenic aberrations with Affymetrix microarray genechip analysis are subject to further investigation. Our focus shifted towards investigating the role of inherited and *de novo* point mutations. These more subtle variations in the genome can be detected by Whole Exome Sequencing (WES).



Schematic illustrating the steps of WES data analysis pipeline employed in this study.

In a preliminary study, we performed WES in 8 out of 36 families to identify risk genes and rare autism susceptibility variants. Preliminary WES analysis led to short-listed susceptibility gene candidates for each family.

The case of family F2



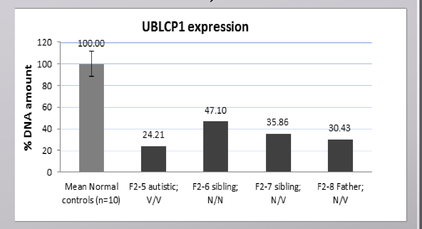
Family F2 pedigree showing genotypes of UBLCP1 mutation where N is the reference allele and V is the alternate allele harboring a deletion within the exon10 of UBLCP1 (g. 158710261CAAAG>C) that leads to a premature stop codon..

Assuming a recessive model of inheritance, we found 5 genes (*ANXA6, GABRB2, ERCC2, HDAC9, PADI1*) exhibiting intronic variants, and 2 genes (*UBLCP1* and *WNK1*) exonic variants. **Ubiquitin-like domain-containing CTD phosphatase 1 (UBLCP1) is of particular interest**.

UBLCP1 dephosphorylates 26S nuclear proteasomes, preventing assembly of the core and regulatory particles into mature 26S proteasome, thereby decreasing their proteolytic activity.

Schematic representation of UBLCP1 showing the predicted stop codon within the phosphatase domain.

UBLCP1 expression by Real-time PCR.
N: reference allele; V: alternate allele



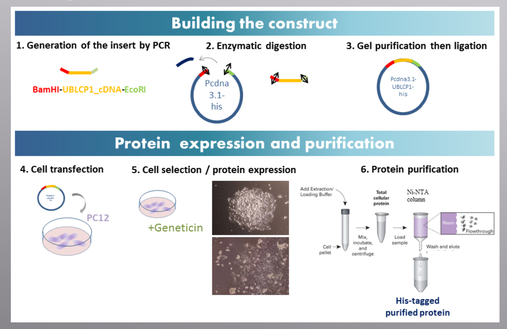
Investigating gene expression levels in peripheral blood by real-time PCR, we observed a decrease in UBLCP1 expression in the autistic patient in comparison with family members and with normal controls.

Effect of the novel mutation within UBLCP1

The novel UBLCP1 gene mutation is predicted to generate a stop codon within the sequence encoding the protein phosphatase domain. This is expected to lead to a truncated protein with a dysfunctional phosphatase activity, drastically affecting UBLCP1 protein function.

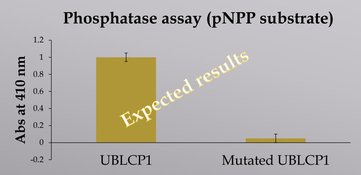
Generation of normal and mutated UBLCP1

In order to study the protein generated by this variation, we generated a recombinant human UBLCP1, fused to His-tag at N-terminus.



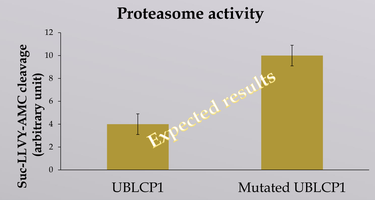
Phosphatase activity

UBLCP1 phosphatase activity will be determined using the calorimetric phosphatase substrate p-nitrophenyl phosphate (pNPP). The reaction produces a water soluble yellow end product that has a strong absorption at 405 nm. Purified normal and mutated UBLCP1 proteins will be added to pNPP in a controlled experiment, and their respective phosphatase activity will be measured.



Cellular activity

The direct effect of normal and mutated UBLCP1 on proteasome activity will be measured *in vitro* utilizing a fluorogenic synthetic peptide substrate, Suc-LLVY-MCA, which gets converted to a highly fluorescent degradation product, 7-amino-4-methylcoumarin (AMC), upon cleavage by proteasome.



To do list

- ✓ **Knockdown-replacement strategy: create a stable shRNA knockdown for the endogenous UBLCP1 gene, followed by co-tranfection with the mutated UBLCP1 vector. This strategy will be utilized in neuronal differentiated PC12 cells, that will be subjected to biochemical and histopathological studies:**
  - ✓ **Cell survival will be assessed using TUNNEL.**
  - ✓ **Immunofluorescence using antibodies against MAP2 to quantify neurite number/length and antibodies against tau to quantify axonal length.**
  - ✓ **Western blots will be used for the quantification of: β-III-tubulin (essential for neuronal dendritic formation), Synaptophysin (presynaptic protein), Drebrin and PSD-95 (postsynaptic proteins), ubiquitin (ubiquitinated cellular proteins).**
- ✓ **Establish knockout/knock-in cell lines using the CRISPR/Cas9 system.**
- ✓ **Use cultured fibroblasts established from a skin biopsy of the autistic patient to compare to control fibroblasts. These cells will be used to generate Human Induced Pluripotent Stem Cells (hiPSCs), genetically reprogrammed adult cells that exhibit a pluripotent stem cell-like state similar to embryonic stem cells. hiPSCs can be directionally differentiated into neurons, allowing study of the impact of the genetic variant on the iPSC-derived neuronal phenotype such as connectivity, synapses, spine density and expression of neurotransmitter receptors.**

Conclusion

Proposed studies will establish a powerful analysis protocol allowing identification of prioritized lists of potential deleterious variants in the Lebanese population enabling confirmation of ASD susceptibility genes, and, uncover novel ones. Extensive functional studies on *UBLCP1* and other candidate genes will link novel mutations to ASD.

References

1. Chaaya, M., Saab, D., Maalouf, F. T. & Boustany, R. M. Prevalence of Autism Spectrum Disorder in Nurseries in Lebanon: A Cross Sectional Study. J Autism Dev Disord (2015).
2. Soueid, J. et al. RYR2, PTDSS1 and AREG genes are implicated in a Lebanese population-based study of copy number variation in autism. Sci. Rep. 6, (2016).
3. Kourtian, S. and Soueid, J. et al. Candidate genes for inherited autism susceptibility in the Lebanese population. Sci. Rep. 7, (2017).

Acknowledgements

We thank families and staff from the Lebanese Autism Society (LAS), SESOBEL, and the AUBMC Special Kids Clinic, who participated and helped in this study. This research is supported by a generous grant from OpenMinds.