ePTRI

European Paediatric Translational Research Infrastructure

Autism Spectrum Disorder: Linked-Read Sequencing Reveals New and Undetected Variants

Alessandra Mezzelani, PhD

National Research Council, Italy

Institute for Biomedical Technologies, Segrate (MI)

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 777554

Autism Spectrum Disorder (ASD): epidemiology and etiology

Epidemiology:

- It occurs within the first 3 years of life
- The prevalence is of 1:68 with male:female ratio=5:1

Etiology not completely understood

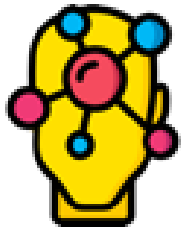
- Genetics play a key role
- Causative or predisposing genetics variants in 30% of patients
- Thousands of genes are involved
- Gene-environment interaction has been proposed

Objective and study design

- **Hypothesis**: variations in non coding and/or unaccessible regions could be involved in ASD

- **Study design**: linked-read WG-NGS of patients with ASD is used to access and discover phasing disease associated regions

- **Expectations:** new variants can:
  - be integrated in interactome studies
  - represent new biomarkers useful:
    - to design diagnostic protocols and kits
    - in patient stratification and personalized medicine



ePTRI
European Paediatric Translational Research Infrastructure

Methods

- Ten patients with ASD were involved: 7 males and 3 females, including 2 couples of male siblings and 1 couple of male-female siblings

- HMW DNA was submitted to 10Xgenomics microfluidics partitioning and barcoding technology and then to Illumina library preparation and WG-NGS

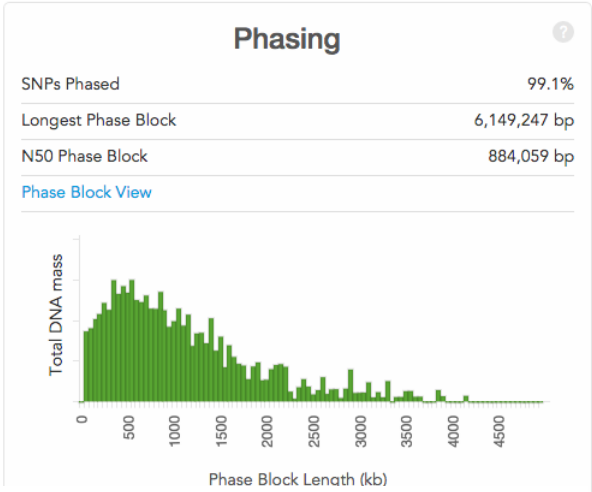
- 10X Long Ranger pipelines and Loupe Genome browser were used to find and phase variants across >10Mb haplotype blocks in comparison to reference genomes.

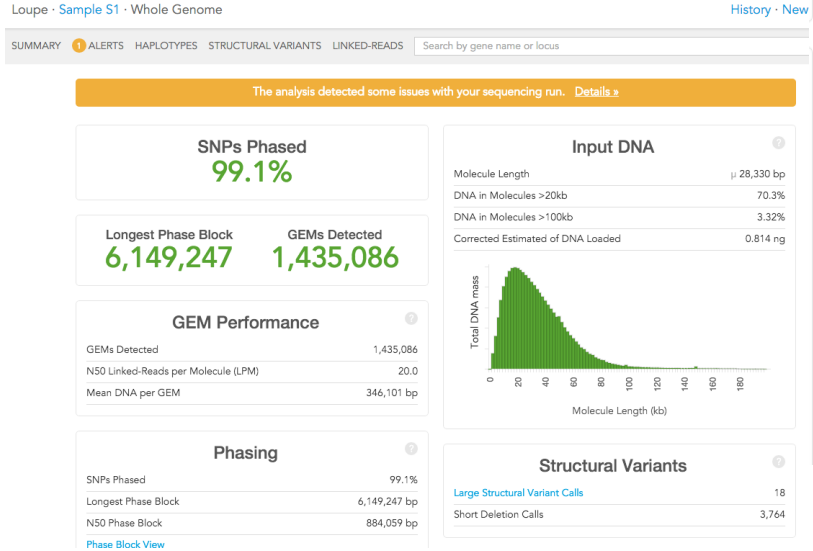
- Genes affected by variants were compared with those already associated with ASD

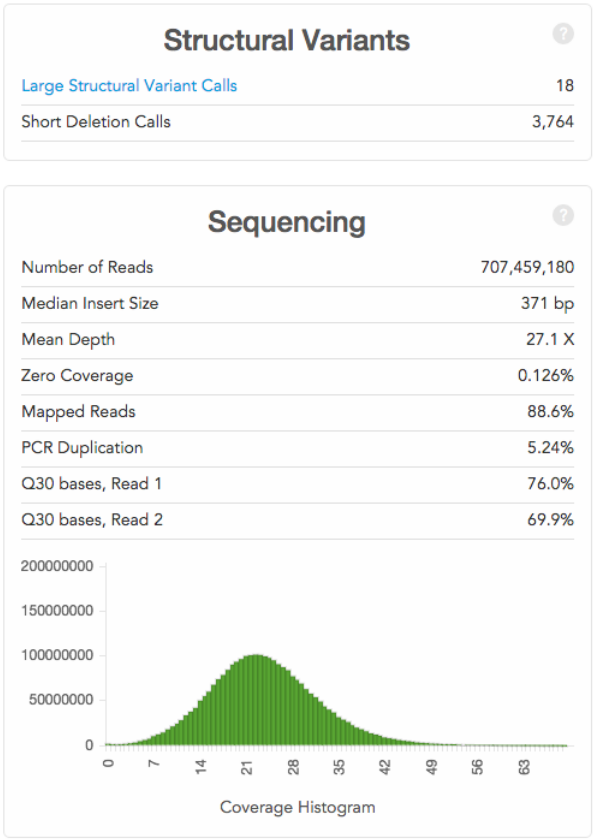
ePTRI
European Paediatric Translational Research Infrastructure

Results:
quality of linked read analysis

Sample 1:
input 1,25ng of HMW-DNA







EPTRI
European Paediatric Translational Research Infrastructure

Results

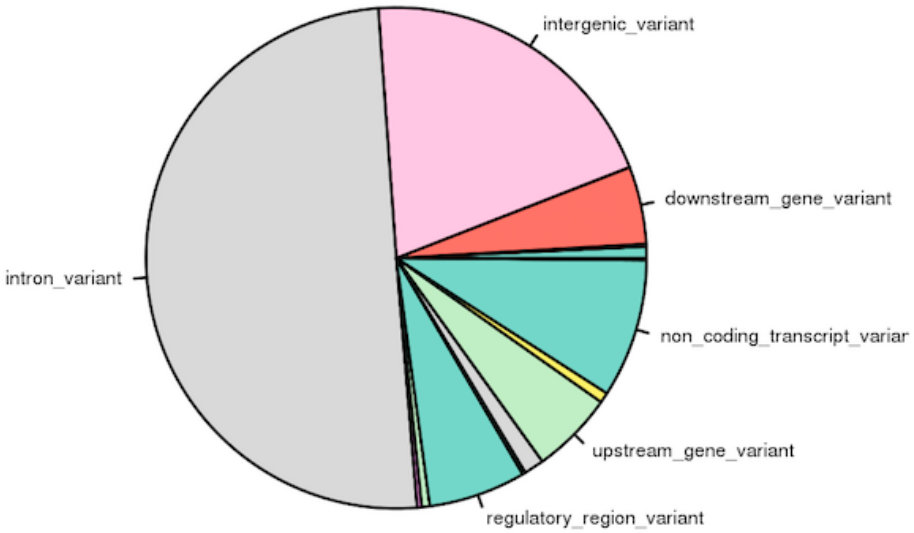
Number of variants	
Total genes	23'740
Total variants	13’297
Annotations within genes	759'163
Annotations out of genes	338'078
Existing variants	942'744
New variants	154'497

Type of variants	
3_prime_UTR_variant	splice_donor_variant
5_prime_UTR_variant	start_lost
coding_sequence_variant	stop_gained
downstream_gene_variant	stop_lost
frameshift_variant	stop_retained_variant
inframe_deletion	synonymous_variant
inframe_insertion	TF_binding_site_variant
intergenic_variant	transcript_ablation
intron_variant	upstream_gene_variant
missense_variant	feature_truncation
non_coding_transcript_exon_variant	non_coding_transcript_variant
protein_altering_variant	splice_region_variant
regulatory_region_variant	regulatory_region_ablation
splice_acceptor_variant	TFBS_ablation

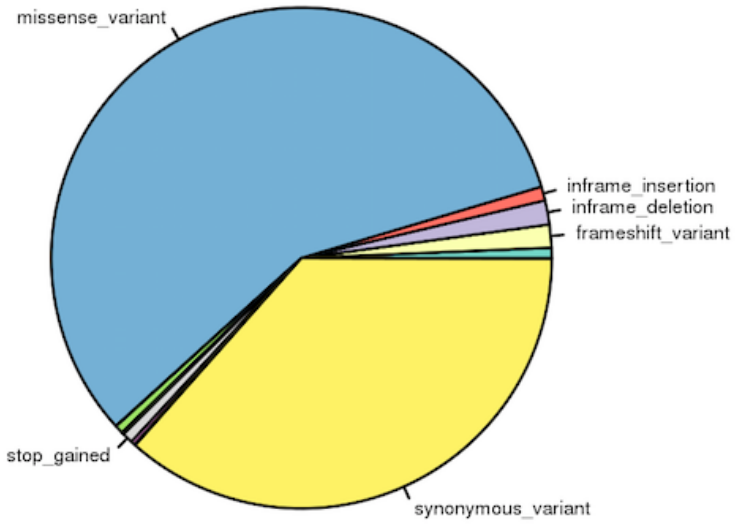
ePTRI
European Paediatric Translational Research Infrastructure

Results: variant consequences

Consequences (all)



Consequences (coding)

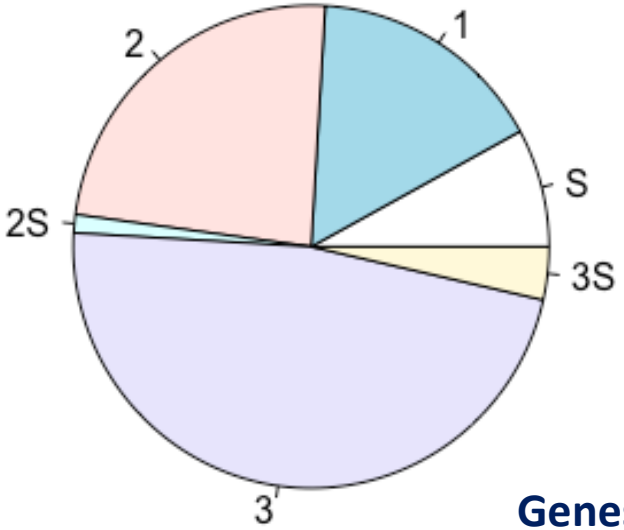


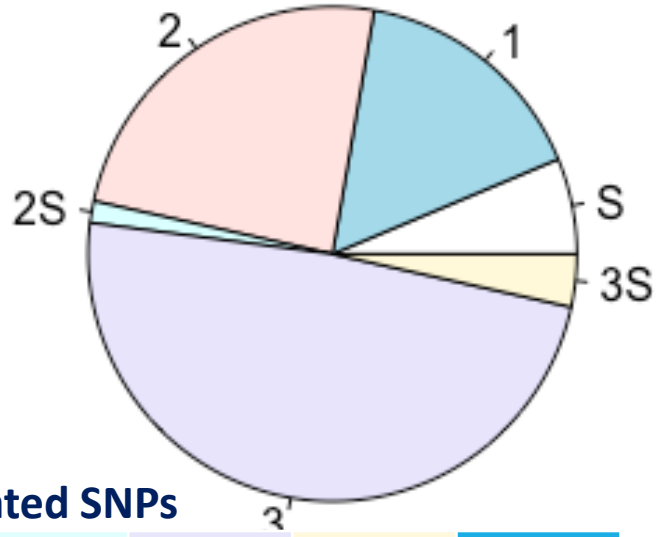
ePTRI
European Paediatric Translational Research Infrastructure

Results: genes containing SFARI annotated/not annotated SNPs distributed into SFARI categories

Genes containing SFARI annotated SNPs distributed into SFARI categories

Sfari modules	S	1	2	2S	3	3S	Tot
#genes	65	132	195	10	388	29	819





Genes containing SFARI not annotated SNPs

Sfari modules	S	1	2	2S	3	3S	Tot
#geni	31	79	120	7	240	17	494

SFARI database https://gene.sfari.org/

EPTRI
European Paediatric Translational Research Infrastructure

Conclusions

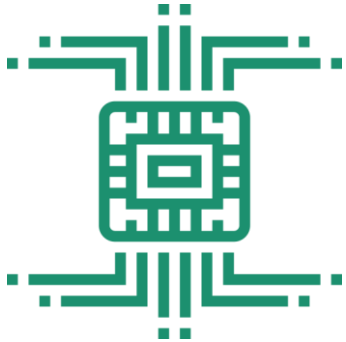
- Linked-read sequencing technology performed at CNR-ITB has allowed to decipherer the genomics heterogeneity in ADS
- Homozygous variants common to 3 couples of affected siblings was identified.
- Further studies will be performed to identify haploblocks carrying the risk alleles and to validate the new ASD variations
- The new biomarkers will be used:
    - to stratify ASD patients population
    - to identify pathways and gene networks involved in the disorder
    - design target-driven personalized treatment

ePTRI
EUROPEAN PAEDIATRIC TRANSLATIONAL RESEARCH INFRASTRUCTURE

Perspectives

- Further sophisticated bioinformatics analyses of the results can predict new ASD associated genes

- Possible development and application of dedicated AI, performed at CNR-ITB,
can accelerate these analyses



- These linked-read WG-NGS based approaches can be applied to many other pediatric diseases with unknown genetic causes

ePTRI
European Paediatric Translational Research Infrastructure

Acknowledgements

Coauthors:

Francesca Anna Cupaioli[1], Ettore Mosca[1], Noemi Di Nanni[1], Paride Pelucchi[1], Luciano Milanesi[1], Maria Elisabetta Raggi[2], Laura Villa[2]

1 National Research Council, Institute of Biomedical Technologies, Segrate, Italy

2 Scientific Institute, IRCCS Eugenio Medea, Bosisio Parini, Italy.

Projects:

EU project GEMMA (grant agreement No 825033) www.gemma-project.eu

EPTRI

CNRBiOmics

EPTRI
European Paediatric Translational Research Infrastructure

EPTRI

European Paediatric Translational Research Infrastructure

The views and opinions expressed in the following PowerPoint slides are those of the individual presenter and should not be attributed to EPTRI or the EC

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 777554