Reviewing the results of the first large-scale genomic study (Deciphering Developmental Disorders) in Wales

Dr Nayanatara Poobalan[1], Dr Andrew E. Fry[2]
1. School of Medicine, Cardiff University
2. All Wales Medical Genomics Service, University Hospital of Wales

Aim - Generate an overview, determine patient correspondence completion, and predict deleteriousness of variants of uncertain significance (VUS).

The Deciphering Developmental Disorders (DDD) project used new testing techniques including array comparative genomic hybridisation and trio-based whole exome sequencing on children with developmental disorders without a diagnosis.

In total, 12,600 patients from around the UK were recruited. 370 of those patients were recruited from Wales. Clinical interpretation, validation and feedback of the DDD results to Welsh patients were delegated to the NHS Medical Genetics Service in Wales.

1. Overview of DDD study

The DECIPHER website allowed access to DDD participient results and was used to create an inventory of all Welsh patients. This inventory comprised of 370 patients, 3124 phenotypes, 847 Human Phenotype Ontology (HPO) terms, and 393 variants. The 100 most common phenotype terms were grouped in categories and displayed in a graph (figure 1) to demonstrate the frequency of occurrence in Welsh DDD patients.

Dysmorphology is the largest recurring category followed by developmental delay or intellectual disability and abnormal behaviour. This demonstated common phenotypes suggesting genetic pathology. 356 variants were single nucleotide variants or small indels and the remaining 37 variants were copy number variants. The single nucleotide variants were categorised according to gene mutation (table 1) and missense variants were annotated for further analysis.

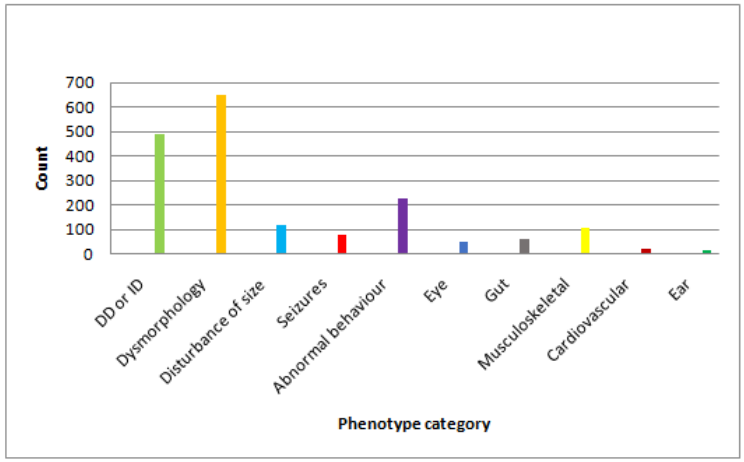

Figure 1. Frequency of usage of the top 100 most common phenotype terms grouped by category.

Consequence	Count
missense variant	263
frameshift variant	35
splice region variant	2
splice donor variant	7
splice acceptor variant	5
stop lost	1
stop gained	27
inframe deletion	5
inframe insertion	9
start lost	1
intron variant	1

Table 1. Frequency of different types of small-scale variation in DDD cohort.

Status	Count
Complete	25
Partially complete	7
Incomplete	3

Table 2. Table of actual status of patient correspondence marked incomplete on DECIPHER

2. Patient correspondence completion

Family correspondence status for the 370 patients were reviewed and 35 were listed as 'incomplete' on DECIPHER. Their medical notes were assessed to determine DDD result documentation and actual correspondence status, enabling the department to act on incomplete documentation and correspondence. A detailed account has been reported to the All Wales Medical Genomics Service and is summarised in the table above (table 2).

3. Predicting deleteriousness

Variants from DECIPHER were categorised according to gene mutation. Missense variants were identified as pathogenic, benign, and variant of uncertain significance (VUS). These missense variants were utilised to determine sensitivity and specificity and subsequent positive likelihood ratio of respective prediction programs. This ratio identified ReVe, MetaLR, and ReVEL as reliable programs and these prediction programs were used to predicted deleteriousness in 96 missense VUSs.

Prediction Program	Pathogenic (N=31)	Benign (N=25)	VUS (N=96)	Positive likelihood ratio
SIFT	0.94	0.6	0.66	2.35
Polyphen2 (HDIV)	0.9	0.16	0.64	1.07
Polyphen2 (HVAR)	0.84	0.36	0.49	1.31
LRT	0.87	0.36	0.58	1.36
MutationTaster	1	0.16	0.84	1.19
MutationAssessor	0.35	0.2	0.07	0.44
FATHMM	0.45	0.72	0.41	1.61
PROVEAN	0.94	0.4	0.52	1.57
VEST3	0.94	0.48	0.64	1.81
MetaSVM	0.65	0.72	0.39	2.32
MetaLR	0.65	0.76	0.44	2.71
M-CAP	0.97	0.08	0.8	1.05
CADD	0.97	0.16	0.83	1.15
DANN	0.87	0.24	0.65	1.14
FATHMM	0.97	0.12	0.86	1.1
Eigen	0.84	0.16	0.73	1
GenoCanyon	0.84	0.28	0.82	1.17
fitCons	0.65	0.4	0.5	1.08
GERP	0.94	0.08	0.9	1.02
phyloP	0.94	0.28	0.83	1.31
phastCons	0.94	0.32	0.79	1.38
SiPhy	0.9	0.36	0.72	1.41
REVEL	0.97	0.68	0.55	3.03
ReVe	0.87	0.64	0.5	2.42

Table 3. Table of sensitivity, specificity, proportion predicted deleterious and positive likelihood ratio.

30 variants were identified as deleterious by all 3 programs and 21 other variants were identified as deleterious by 2 out of the 3 programs. Of the 51 variants, gene variation in *SIN3A, KDM6A, SCN8A, KMT2D, HDAC8, and HCFC1* genes were noted as particularly interesting as they are as deleterious according to these programs and are known de novo mutations.

Conclusion - This review identified common phenotypes suggesting pathology and encouraged greater patient correspondence completion in Welsh DDD cases. Welsh DDD patients are generating interesting research findings e.g. *SIN3A* variant highlighted in the review.

Reference
1. Deciphering Developmental Disorders Study Large-scale discovery of novel genetic causes of developmental disorders. Nature. 2015;519(7542):223–228.doi: 10.1038/nature14135. - DOI - PMC - PubMed
2. Helen V Firth D. Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. The Lancet. 2014.
3. Heidi L. Rehm S. Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology [Internet]. PMC. 2015 [cited 26 June 2019]. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4544753/
4. Balasubramanian, M., Dingemans, A.J.M., Albaba, S., Richardson, R., Yates, T.M., Cox, H., Douzgou, S., Armstrong, R., Sansbury, F.H., Burke, K.B., Fry, A.E., Ragge, N., Sharif, S., Foster, A., De Sandre-Giovannoli, A., Elouej, S., Vasudevan, P., Mansour, S., Wilson, K. and Stewart, H. (2021). Comprehensive study of 28 individuals with SIN3A-related disorder underscoring the associated mild cognitive and distinctive facial phenotype. European Journal of Human Genetics, 29(4), pp.625–636.