AWMGS
All Wales Medical Genomics Service

Test Information Sheet – Whole Genome Sequencing for Rare Diseases

Whole genome sequencing (WGS) services are currently available from AWMGS for specific clinical indications where samples from both parents of the proband are available.

Description

Using WGS, it is possible to look for changes across the genome to reduce the need for multiple different diagnostic tests and improve the patients' journey to finding a genetic cause of their clinical features.

Research studies have shown that the application of WGS in a routine clinical setting is highly effective in improving outcomes for patients by reducing the need for multiple diagnostic tests, facilitating earlier treatment decisions and preventing repeat inpatient admissions (Kingsmore et al., 2019, PMID: 31564432; French et al. 2019, PMID: 30847515; Mestek-Boukhibar et al. 2018, PMID: 30049826; Stark & Ellard 2021, PMID: 34744166; Lindstrand et al., 2019, PMID: 31694722; 100,000 Genomes Project, 2021, PMID: 34758253; Malinowski et al., 2020, PMID: 32203227).

Referral Criteria

These tests are only available on a trio basis where the proband and BOTH parents are available. Alternative genetic testing is available if only one parent is present. For full details, please see the referral criteria leaflet for these services on our website.

Reporting of Results and Possible Incidental Findings

It is possible that as part of this test, incidental or secondary findings that do not explain the current clinical features in the patient may be identified. The AWMGS policy is to analyse and report actionable (likely) pathogenic variants from a gene list. The list has been agreed by the AWMGS Clinical Genetics team and is based on the ACMG list for reporting of secondary findings v3.0 (PMID: 34012068). Note that variants in the list of genes will **not** be actively looked for. These will only be detected if the genes on the list are included on the gene panel applied or when performing agnostic analysis. Therefore, in cases where the genes on the list are not on the panel, our analysis has not excluded variants in these genes.

Trio WGS will reveal possible non-paternity (or non-maternity) and this result could be discussed with the referring clinician.

If a diagnosis is not found, the variant data from genome sequencing tests will be stored long-term in order that further analysis can be undertaken in the future. The results of re-analyses may be issued as a supplementary report to the referring clinician.

Test Methods

Genomic DNA extracted from the proband's and parents' blood samples is sequenced simultaneously using Next Generation Sequencing (NGS) of the whole genome using Illumina PCR-Free assay and sequenced on Illumina NovaSeq 6000 platform. Sequences are aligned to human genome assembly GRCh38 (hg38), and variants identified with the Illumina DRAGEN (Dynamic Read Analysis for GENomics) Bio-IT Platform.

Analysis includes evaluation of variants that are identified to be *de novo*, compound heterozygous, homozygous and X-linked using in-house bioinformatic pipelines. Variants are interpreted and reported following the latest ACGS/ACMG guidelines (Ellard et al., 2020; Richards et al., 2015).

The WGS analysis pathway utilises a two-phase approach for analysis of small variants (SNVs and Indels). In the first phase, a virtual panel of genes based on the PanelApp gene panels (such as the Paediatric Disorders panel - https://panelapp.genomicsengland.co.uk/panels/486/ or the Intellectual Disability panel - https://panelapp.genomicsengland.co.uk/panels/285/) is used to only look at variants in genes included on the gene panel. If no variant of interest is found in the first phase, analysis moves to the second phase. The second phase uses a "gene agnostic" approach to look for variants in genes not included in the first phase panel of genes and that are ranked highly by Exomiser based on the patient's phenotype. The top 8 Exomiser variants and relevant de novo, compound heterozygous and homozygous variants will be looked at. This means that nearly all genes in the human genome, related to the patient's phenotype, are inspected for variants that pass the bioinformatic filtering pipeline.

After receiving a normal result, inquiries often arise regarding the inclusion of specific genes in the analysis. If a normal result has been issued, it indicates that both phases of the analysis have been completed, and nearly all genes have been accounted for. However, if you wish to confirm whether a gene was included in the initial panel analysis, please contact the laboratory to obtain information about the version of the gene panel used for the analysis in your patient. Once the panel version is confirmed, you can then cross-reference the content of the relevant panel on PanelApp to verify if the gene of interest was included. It's important to note that we always aim to use the latest NHS Genomic Medicine Service (GMS) Signed Off versions of gene panels to ensure most up-to-date analysis.

Note for the Intellectual Disability and Congenital Malformations and/or Dysmorphisms services, only green genes on the PanelApp panel will be included.

As phenotypic based filtering methods are used in analysis, it is **vital** that clear and detailed phenotypic information for both the proband and parents is included on the request form in order to give accurate results. Phenotypic information is required in the Human Phenotype Ontology (HPO) form (e.g. HPO:XXXXX), which can be accessed here https://hpo.jax.org/app/

Incomplete Penetrance Analysis

In cases where it has been specifically indicated on the request form by ticking the parent affected box, an additional analysis step will be undertaken to ascertain if any variants inherited from the affected parent could be the cause of the clinical features in the proband. This will **only** be done for the genes on the panel used in the first analysis phase and will **only** be performed for the Intellectual Disability service.

Limitations

Currently this test is validated for the detection of small variants (SNV and Indels) and Copy Number Variants (CNVs) from whole genome sequencing data only. Other Structural Variants (SVs) may also be detected, but the sensitivity of detection is unknown. This analysis does not include complex structural genetic variants such as inversions or translocations, mitochondrial variants, triplet repeats or UPD. Small CNVs <5KB in size, whilst commonly detected, have reduced sensitivity and may not always be discovered. This assay may not always detect multi-nucleotide variants. SNP Array testing will be carried out concurrently for an interim period for CNV detection.

The WGS analysis is primarily looking for *de novo, recessive, compound heterozygous, and X linked* variation. Therefore, variants inherited from an affected parent, or from a parent with milder clinical features or causing disorders with a later age of onset may not be detected. These variants may be missed if they are present in genes with low or variable penetrance. Please see above for incomplete penetrance analysis, where applicable.

This test will not detect mosaicism for variants discovered through WGS.

CNVs detected via WGS will be included in analysis if they intersect a protein coding region or overlap significantly with an ISCA region on the virtual panel. Therefore, non-coding CNVs are unlikely to be detected.

Intronic variants: only conserved splice site variants (defined as variants located within 1–3 bases of an exon or 1–8 bases of an intron are analysed. Other non-coding variants are not currently analysed unless they were predicted to impact splicing by SpliceAI (Jaganathan et al., 2019) or previously reported in the ClinVar database as pathogenic.

On average 97% of coding region exons ± 5 base pairs will be covered at greater than 20x.

Given typical sequencing conditions, the lower confidence interval of the sensitivity of variant calling for Single Nucleotide Variant (SNV) detection is 99.9%. The lower confidence interval of the sensitivity of variant calling for Indel detection is 99.6%. The sensitivity was calculated over a set of high confidence genome regions which are accessible by NGS. The sensitivity of variant calling outside of these regions is unknown. The lower confidence interval of the sensitivity of variant calling for Copy Number Variant (CNV) on WGS is 95.5%.

It is well established that some regions of the genome are inherently difficult to sequence, e.g. pseudogene and GC-rich regions, and result in low coverage and other issues that may compromise the sensitivity of variant calling. In the event of a negative result, please contact the laboratory to discuss the coverage of any genes that may be considered a likely cause of the patient's phenotypic features. Genes within the Intellectual Disability and Congenital Malformations and/or Dysmorphisms panels which are known to have low coverage are described in the appendix. However, this list is not exhaustive.

It is possible that there is a pathogenic variant in the genome which has not currently been detected by this primary analysis. Variants may be identified but not recognised as causative due to insufficient scientific knowledge about the genomic region and its function at the time.

Re-analysis of the data to incorporate updated clinical information and/or newly emerging gene and variant information may be performed and reported at a later time.

Recessive carrier status will not be disclosed.

Samples Required

Please send samples in EDTA to the laboratory. We require at least 1ml from the proband and 3–4ml from the parents.

Reporting Times

The reporting time for rapid WGS is 14 calendar days – this is applicable to the R14 WINGS service only.

The reporting time for PRRS is 21 calendar days - this WGS service is applicable for PRRS referrals on blood samples only.

The reporting time for routine WGS is 84 calendar days.

Appendix

Table 1. Genes with known coverage issues within the Intellectual Disability and Congenital Malformations and/or Dysmorphisms panels when analysed on Whole Genome Sequencing. Genes where the median sample had less than 95% of exonic bases +- 5bp covered at 20x are shown.

Gene	Median Exonic Bases covered at greater than 20x
ACAN	0.84455
ANAPC1	0.6801
ARX	0.88915
CFC1	0.156
FBRSL1	0.8305
FMN2	0.8844
GBA	0.8785
HOXA13	0.8674
HYDIN	0.60265
IKBKG	0.33735
MNX1	0.9418
MSTO1	0.54555
NDUFB11	0.9076
NEB	0.8501
OCLN	0.663
OTOA	0.7423
POU3F3	0.8606
PRODH	0
SRY	0
STRC	0.2707
TUBB2A	0.6884
TUBB2B	0.83705
ZSWIM6	0.9367

References

1. Kingsmore, SF et al. A randomized, Controlled Trial of the Analytic and Diagnostic Performance of Singleton and Trio, Rapid Genome and Exome Sequencing in Ill Infants. Am J Hum Genet. 2019 Oct 3; 105(4): 719–33.
2. Mestek-Boukhibar, L et al. Rapid Paediatric Sequencing (RaPS): Comprehensive real-life workflow for the rapid diagnosis of critically ill children. J Med Genet. 2018.
3. French, CE et al. Whole genome sequencing reveals that genetic conditions are frequent in intensively ill children. Intensive Care Med. 2019 May 1; 45(5): 627–36.
4. Stark & Ellard. Rapid genomic testing for critically ill children: time to become standard of care? Eur J Hum Gen. 2021.
5. Lindstrand, A et al. From cytogenetics to cytogenomics: whole-genome sequencing as a first-line test comprehensively captures the diverse spectrum of disease-causing genetic variation underlying intellectual disability. Genome medicine. 2019 Dec 1; 11(1): 68.

6. Malinowski, J et al. Systematic evidence-based review: outcomes from exome and genome sequencing for pediatric patients with congenital anomalies or intellectual disability. Genet Med. 2020; 22(6): 986–1004.
7. 100,000 Genomes Project Pilot Investigators, et al. 100,000 Genomes Pilot on Rare-Disease Diagnosis in Health Care – Preliminary Report. N Engl J Med. 2021 Nov 11; 385(2): 1868–1880.
8. Miller DT - ACMG Secondary Findings Working Group. ACMG SF v3.0 list for reporting of secondary findings in clinical exome and genome sequencing: a policy statement of the American College of Medical Genetics and Genomics (ACMG). Genet Med. 2021 Aug;23(8):1381-1390 PMID: 34012068
9. Ellard, S et al. ACGS Best Practice Guidelines for Variant Classification in Rare Disease 2020. Association for Clinical Genomic Science.
10. Richards, S et al. ACMG Laboratory Quality Assurance Committee. 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. Genet Med. 2015 May; 17(5): 405–24.
11. Jaganathan et al. Predicting Splicing from Primary Sequence with Deep Learning. Cell. 2019 Jan; 176(3): 414–416.
12. Robinson PN et al. Improved exome prioritization of disease genes through cross-species phenotype comparison. Genome research 2014;24;2;340-8 PMID: 24162188