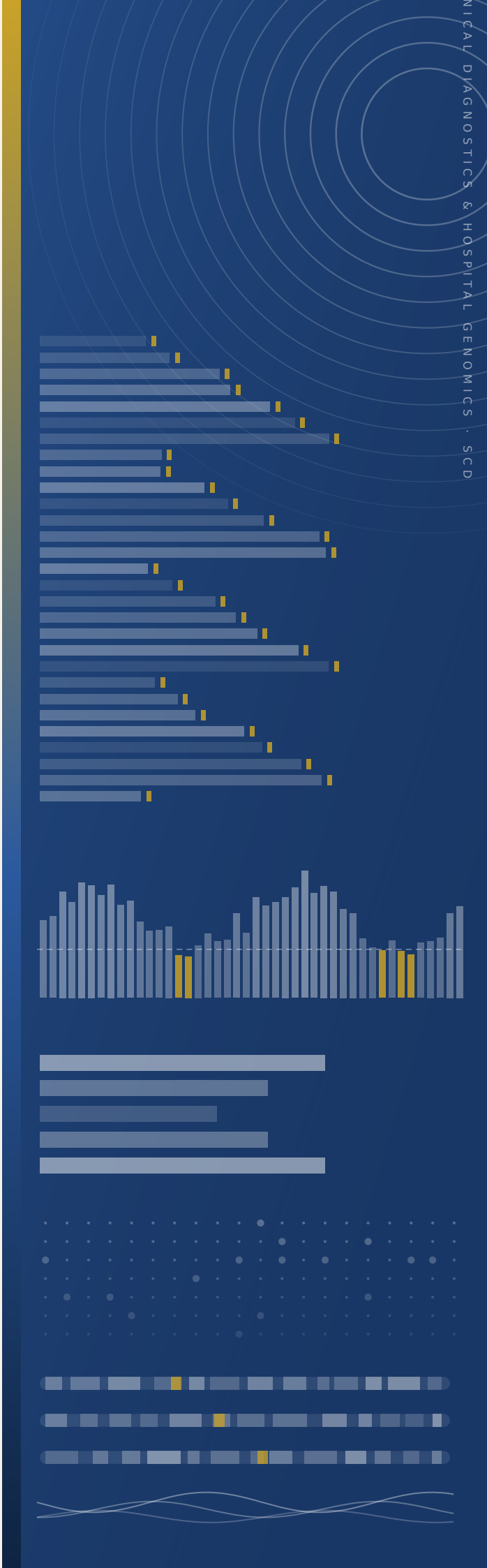
NTHRYS

OPC PVT LTD · HYDERABAD · INDIA

SECTOR CATALOGUE 01 · NGS DATA ANALYSIS

Clinical Diagnostics & *Hospital Genomics*

Twenty-nine industrial training programmes, one for each distinct type of NGS data analysis performed in accredited hospital and reference laboratories. Every programme is set out on a single page answering the questions a laboratory actually asks before it commits budget — why the analysis is done, when it is done, what it detects, what it does not detect, what makes it hard, who does this job, and what a trainee can do afterwards — with steps, duration, prerequisite and fee stated plainly.



29	502	2450	419
ANALYSIS TYPES	STEPS	HOURS	DAYS

CATALOGUE SCD · 2026 EDITION

INDUSTRIAL TRAINING · ACCREDITED LABORATORY PRACTICE

PIPELINE REGISTER

Clinical Diagnostics — *Analysis Types*

Twenty-nine distinct types of NGS data analysis performed in clinical diagnostic laboratories, each an independent analysis with its own input, its own pipeline and its own report set. Each is listed under its formal title with the industry name and the shorthand a laboratory uses in practice. Every programme carries its own step count, contact hours and fee, derived from the work that specific analysis involves and the accreditation obligations attached to it.

CODE	ANALYSIS TYPE · INDUSTRY NAME & SHORT NAME	STEPS	HOURS	DAYS	FEE (INR)
SCD-01	**Germline Whole Genome Analysis** Whole Genome Sequencing analysis · WGS analysis	24	120	20	2,85,000
SCD-02	**Germline Whole Exome Analysis** Whole Exome Sequencing analysis · WES analysis	21	104	18	2,45,000
SCD-03	**Clinical Exome Analysis** Clinical exome sequencing analysis · CES analysis · medical exome	18	88	15	2,05,000
SCD-04	**Mendeliome Analysis** Comprehensive Mendelian gene set analysis · Mendeliome	17	84	14	1,95,000
SCD-05	**Targeted Gene Panel Analysis** Capture panel sequencing analysis · Panel analysis · NGS panel	16	78	13	1,80,000
SCD-06	**Amplicon Panel Analysis** Amplicon sequencing analysis · Amplicon analysis	14	68	12	1,55,000
SCD-07	**Virtual Panel Analysis** In-silico panel analysis · Virtual panel · slice	12	58	10	1,30,000
SCD-08	**Trio Analysis** Proband-parent-parent family analysis · Trio · family analysis	22	110	19	2,60,000
SCD-09	**Duo Analysis** Proband-single parent analysis · Duo	17	82	14	1,90,000
SCD-10	**Family Segregation Analysis** Segregation and cascade testing analysis · Segregation study	13	62	11	1,40,000
SCD-11	**Cohort Joint Analysis** Joint genotyping and cohort-scale analysis · Joint genotyping	19	92	16	2,15,000
SCD-12	**Rapid Genome Analysis** Rapid whole genome sequencing analysis · rWGS analysis	23	114	19	2,70,000
SCD-13	**Ultra-Rapid Genome Analysis** Ultra-rapid whole genome sequencing analysis · urWGS analysis · 24-hour genome	24	122	21	2,95,000
SCD-14	**Rapid Exome Analysis** Rapid whole exome sequencing analysis · rWES analysis	20	98	17	2,30,000
SCD-15	**Copy Number Variant Analysis** CNV analysis from sequencing data · CNV analysis	18	86	15	2,00,000
SCD-16	**Structural Variant Analysis** SV analysis from whole genome data · SV analysis	21	102	17	2,40,000
SCD-17	**Digital Karyotyping** Chromosomal microarray-equivalent analysis · CMA-equivalent analysis	17	82	14	1,90,000
SCD-18	**Runs of Homozygosity Analysis** Homozygosity mapping and AOH analysis · ROH · AOH · homozygosity mapping	12	56	10	1,25,000
SCD-19	**Uniparental Disomy Analysis** UPD detection and imprinting analysis · UPD analysis	13	64	11	1,45,000
SCD-20	**Repeat Expansion Analysis** Short tandem repeat expansion detection · RED · repeat analysis	15	72	12	1,65,000
SCD-21	**Mitochondrial Genome Analysis** mtDNA sequencing analysis · mtDNA analysis	16	76	13	1,75,000
SCD-22	**Mosaicism Analysis** Low-level mosaic variant detection · Mosaicism analysis	19	94	16	2,20,000
SCD-23	**Pharmacogenomic Analysis** PGx star-allele and diplotype analysis · PGx analysis	16	78	13	1,80,000
SCD-24	**HLA Typing Analysis** HLA genotyping from sequencing data · HLA typing	13	62	11	1,40,000
SCD-25	**Carrier Screening Analysis** Expanded carrier screening analysis · ECS analysis	15	70	12	1,60,000
SCD-26	**Genomic Reanalysis** Periodic reanalysis and reclassification · Reanalysis	14	66	11	1,50,000
SCD-27	**Clinical Long-Read Analysis** Long-read sequencing analysis for diagnostics · Long-read analysis	20	100	17	2,35,000
SCD-28	**Diagnostic RNA-Seq Analysis** RNA studies for variant interpretation · RNA studies · splice confirmation	18	88	15	2,05,000
SCD-29	**Methylation Episignature Analysis** Episignature and imprinting methylation analysis · Episignature · EpiSign	15	74	13	1,70,000
Complete Sector Catalogue 01 — all twenty-nine analysis types		**502**	**2450**	**419**	**56,95,000**

SCD-01
SECTOR 01

Germline Whole Genome Analysis

(**Whole Genome Sequencing analysis** · WGS analysis)

WHY IT IS DONE

It is the broadest unbiased genomic test available, examining coding and non-coding sequence together in one assay. When a patient has an undiagnosed condition with a plausible genetic basis and narrower testing has failed, this is the analysis that resolves it or formally closes the question.

WHAT IT ANSWERS

Is there a genetic cause for this patient's presentation anywhere in the genome, and if so what is it, how was it inherited, and what is the recurrence risk for the family?

WHEN IT IS DONE

After a non-diagnostic exome or panel, or as a first-line test where the presentation is complex, multi-system or does not point to any specific gene. Increasingly used first-line in paediatric and neurological indications.

WHAT TRIGGERS IT

Clinician request for undiagnosed disease, a negative exome, a phenotype spanning multiple organ systems, or a suspected structural or non-coding cause that coding-only testing cannot address.

HOW IT IS DONE

Reads are aligned to a version-locked clinical reference, processed and recalibrated, then run through parallel calling arms for short variants, copy number, structural variation, repeats and mitochondrial sequence. Findings are annotated, filtered against population frequency, prioritised against the patient's phenotype and classified under ACMG-AMP before confirmation and sign-out.

WHAT IT PRODUCES

A signed clinical report, the classified variant set with full evidence records, the callable region file defining what the test could interrogate, copy number and structural findings, and the archived data package for future reanalysis.

WHAT IT DETECTS

Single nucleotide variants and small insertions and deletions genome-wide, copy number variants, structural rearrangements, repeat expansions, mitochondrial variants, and deep intronic and regulatory variation, with phasing across the genome.

WHAT IT DOES NOT DETECT

Methylation and imprinting defects, most balanced rearrangements in repeat-rich regions, very large repeat expansions beyond read length, low-level tissue-restricted mosaicism absent from blood, and variants in regions the reference genome does not represent. Interpretation of non-coding findings remains severely limited even where they are detected.

WHAT MAKES IT HARD

The volume of variation is the problem, not the sequencing. Five million variants must be reduced to a defensible shortlist without discarding the answer. Non-coding findings are abundant and almost uninterpretable, structural calls carry high false positive rates, and the callable region must be declared honestly rather than implied.

WHO DOES THIS JOB

Clinical bioinformatician or clinical scientist in genomics, typically at senior or lead level, working alongside a clinical geneticist who holds interpretive sign-out responsibility.

WHAT YOU CAN DO AFTERWARDS

Run and defend a complete germline genome pipeline end to end, declare its callable footprint and limitations accurately, classify variants to ACMG-AMP standard with documented evidence, and present findings at a multidisciplinary review.

WHO HIRES FOR IT

Hospital genomics laboratories, national genome programmes, diagnostic chains offering genome testing, genome-focused CROs, and pharmaceutical companies running rare disease discovery programmes.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
24	**120 hours · 20 days**	Linux command line, basic genetics	**2,85,000**

SCD-01
WORKFLOW

Germline Whole Genome Analysis

(**Whole Genome Sequencing analysis** · WGS analysis)

COMPLETE WORKFLOW — 24 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Case authorisation and consent verification**	Test scope, secondary findings election and storage consent confirmed before analysis begins	LIMS record, consent form	*Analysis authorisation*
02	**Data receipt and integrity check**	Checksums verified and file completeness confirmed against the sequencing run record	md5sum, seqkit stats	*Verified FASTQ*
03	**Raw read quality assessment**	Q30 fraction, adapter load, duplication estimate and GC profile measured against validated criteria	FastQC, MultiQC	*QC acceptance record*
04	**Read preprocessing**	Adapter read-through, low-quality tails and polyG artefacts removed with base retention accounting	fastp	*Trimmed FASTQ*
05	**Reference and resource version lock**	Reference build, analysis set and all known-sites resources confirmed and checksummed	Reference registry	*Version lock record*
06	**Alignment to clinical reference**	Reads aligned alt-aware with read groups assigned per sample and library	bwa-mem2, GRCh38 analysis set	*Aligned BAM*
07	**Duplicate marking**	PCR and optical duplicates flagged with flow-cell-appropriate optical pixel distance	Picard MarkDuplicates	*Marked BAM*
08	**Base quality recalibration**	Systematic base quality error patterns corrected against known variant sites	GATK BQSR	*Analysis-ready BAM*
09	**Alignment quality metrics**	Mapping rate, insert size, error rate and chimeric read fraction assessed	Picard, samtools stats	*Alignment QC record*
10	**Coverage and callable region analysis**	Genome-wide depth measured and the formal callable BED generated at diagnostic threshold	mosdepth, GATK CallableLoci	*Callable region BED*
11	**Sample identity fingerprinting**	Genotypes captured at a fixed SNP panel for identity and cross-assay comparison	somalier extract	*Identity fingerprint*
12	**Contamination estimation**	Within-sample contamination fraction estimated using ancestry-aware population allele frequencies	VerifyBamID2	*Contamination record*
13	**Sex and ancestry verification**	Inferred sex and ancestry compared against the clinical record with discrepancies flagged	X/Y coverage, PCA projection	*Verification record*
14	**Short variant calling**	SNVs and small indels called genome-wide with correct sex chromosome ploidy handling	HaplotypeCaller or DeepVariant	*Germline gVCF*
15	**Variant filtering and recalibration**	Call set filtered by recalibration or documented hard filters with metric review	GATK VQSR	*Filtered VCF*
16	**Copy number calling**	Read-depth copy number called genome-wide with GC and mappability correction	GATK gCNV, CNVnator	*CNV call set*
17	**Structural variant calling**	Deletions, duplications, inversions and translocations called by ensemble and merged	Manta, Delly, SURVIVOR	*SV call set*
18	**Repeat expansion genotyping**	Pathogenic repeat loci genotyped against the catalogue with threshold assessment	ExpansionHunter	*Repeat genotypes*
19	**Mitochondrial variant analysis**	mtDNA variants called with heteroplasmy quantification and nuclear mitochondrial sequence exclusion	Mutect2 mitochondrial mode	*mtDNA results*
20	**Annotation and population filtering**	Consequence, transcript and population frequency annotation applied on MANE Select transcripts	VEP, gnomAD, ClinVar	*Annotated variant set*
21	**Phenotype-driven prioritisation**	Candidates ranked against HPO terms with all inheritance models applied in parallel	Exomiser, HPO terms	*Ranked candidates*
22	**ACMG-AMP classification**	Evidence codes applied and documented individually for every candidate variant	ACMG-AMP with gene specifications	*Classified variants*
23	**Orthogonal confirmation**	Every reportable finding confirmed by an independent class-appropriate method	Sanger, MLPA, ddPCR	*Confirmation record*
24	**Report generation and sign-out**	Result, interpretation, limitations and callable footprint issued under dual signature	Validated report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
24	**120 h · 20 d**	**SCD-01**	**2,85,000**

SCD-02
SECTOR 01

Germline Whole Exome Analysis

(**Whole Exome Sequencing analysis** · WES analysis)

WHY IT IS DONE

Around eighty-five percent of known disease-causing variants sit in the two percent of the genome that codes for protein. Restricting analysis to that fraction delivers most of the diagnostic yield of a genome at a fraction of the cost, which is why it remains the highest-volume diagnostic genomic test worldwide.

WHAT IT ANSWERS

Is there a disease-causing variant in any protein-coding gene that explains this patient's presentation, and what are its inheritance and recurrence implications?

WHEN IT IS DONE

First-line for most suspected monogenic disease where the phenotype does not point to a single gene, and after targeted panel testing has returned negative. Standard of care in paediatric developmental disorders and many adult neurological and metabolic presentations.

WHAT TRIGGERS IT

A clinical presentation consistent with monogenic disease without a clear single-gene candidate, a negative panel result, or a family history suggesting Mendelian inheritance.

HOW IT IS DONE

Reads are aligned and processed within the capture target with clinical padding, capture performance is measured, and per-exon coverage is certified for every clinically curated gene. Short variants and exon-level copy number are called, annotated, filtered on population frequency, prioritised against phenotype, classified and confirmed orthogonally before sign-out.

WHAT IT PRODUCES

A signed clinical report, per-exon coverage and gap report, classified variant set with evidence records, exome copy number findings, and the orthogonal confirmation record for every reported variant.

WHAT IT DETECTS

Coding single nucleotide variants and small insertions and deletions, canonical splice site variants, exon-level copy number changes where a matched reference pool exists, and variants in the immediately flanking intronic sequence captured by the design.

WHAT IT DOES NOT DETECT

Deep intronic and regulatory variants, most structural rearrangements, repeat expansions, and any gene or exon that the capture kit covers poorly. GC-rich first exons drop out systematically, and pseudogene-bearing genes are frequently uninterpretable without a separate method. A negative exome is only as strong as its gap report.

WHAT MAKES IT HARD

Capture is never uniform. The gap report is the deliverable most often skipped and most often needed, because a negative result over an uncovered exon is not a negative result. Exome copy number calling depends entirely on reference pool quality and is routinely implemented without adequate validation.

WHO DOES THIS JOB

Clinical bioinformatician or clinical scientist in genomics, mid to senior level. The most common single job description in clinical genomics worldwide.

WHAT YOU CAN DO AFTERWARDS

Run a complete exome pipeline including copy number, produce and interpret an honest gap report, prioritise against phenotype, and classify variants to ACMG-AMP standard with full evidence documentation.

WHO HIRES FOR IT

Every hospital genomics laboratory and diagnostic chain offering genetic testing, reference laboratories, CROs, and public health genomics programmes. The largest single employment pool in this sector.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
21	**104 hours · 18 days**	Linux command line, basic genetics	**2,45,000**

SCD-02
WORKFLOW

Germline Whole Exome Analysis

(**Whole Exome Sequencing analysis** · WES analysis)

COMPLETE WORKFLOW — 21 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Case authorisation and consent verification**	Test scope and secondary findings election confirmed before analysis begins	LIMS record, consent form	*Analysis authorisation*
02	**Data receipt and integrity check**	Checksums and read counts verified against the sequencing run record	md5sum, seqkit stats	*Verified FASTQ*
03	**Raw read quality assessment**	Quality, adapter content and duplication measured against validated acceptance criteria	FastQC, MultiQC	*QC acceptance record*
04	**Read preprocessing**	Adapters and low-quality tails removed with base retention accounting per sample	fastp	*Trimmed FASTQ*
05	**Capture target preparation**	Bait and target intervals prepared with clinical padding and naming harmonised	Picard BedToIntervalList	*Interval lists*
06	**Alignment and duplicate marking**	Reads aligned to the clinical reference and duplicates flagged	bwa-mem2, Picard MarkDuplicates	*Marked BAM*
07	**Base quality recalibration**	Base quality error patterns corrected against version-matched known sites	GATK BQSR	*Analysis-ready BAM*
08	**Capture performance assessment**	On-target fraction, fold enrichment and fold-80 penalty measured against thresholds	Picard CollectHsMetrics	*Capture efficiency record*
09	**Per-exon coverage and gap analysis**	Depth measured for every exon of every curated gene with shortfalls listed	mosdepth, bedtools coverage	*Gap report and callable BED*
10	**Sample identity fingerprinting**	Fixed SNP panel genotypes captured for identity verification	somalier extract	*Identity fingerprint*
11	**Contamination estimation**	Contamination fraction estimated with ancestry-aware allele frequencies	VerifyBamID2	*Contamination record*
12	**Short variant calling**	SNVs and indels called across padded target intervals to recover splice-region variants	HaplotypeCaller or DeepVariant	*Exome VCF*
13	**Variant filtering**	Call set filtered with exome-appropriate thresholds and metric review	GATK hard filters or VQSR	*Filtered VCF*
14	**Reference pool selection**	Matched normals from the same capture kit and batch selected and compatibility verified	Internal reference pool	*Authorised reference pool*
15	**Exome copy number calling**	Exon-level copy number called with normalisation and denoising against the pool	gCNV, CNVkit or DECoN	*Exome CNV calls*
16	**Copy number quality filtering**	Calls quality-scored with stricter thresholds for single-exon events and non-informative exons declared	Caller quality metrics	*Filtered CNV set*
17	**Annotation and population filtering**	Consequence and frequency annotation applied on MANE Select with splice prediction	VEP, gnomAD, ClinVar, SpliceAI	*Annotated variant set*
18	**Phenotype-driven prioritisation**	Candidates ranked against HPO terms with inheritance models applied	Exomiser, HPO terms	*Ranked candidates*
19	**ACMG-AMP classification**	Evidence codes applied per variant with singleton-unavailable criteria documented	ACMG-AMP with gene specifications	*Classified variants*
20	**Orthogonal confirmation**	Reportable sequence and copy number findings confirmed by independent methods	Sanger, MLPA, ddPCR	*Confirmation record*
21	**Report generation and sign-out**	Result, interpretation and gap-derived limitations issued under dual signature	Validated report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
21	**104 h · 18 d**	**SCD-02**	**2,45,000**

SCD-03
SECTOR 01

Clinical Exome Analysis

(**Clinical exome sequencing analysis** · CES analysis · medical exome)

WHY IT IS DONE

A full exome generates a large burden of variants of uncertain significance in genes with no established disease link, which consumes interpretation time and creates clinical uncertainty. Restricting analysis to genes with proven disease association delivers most of the yield with a fraction of the interpretive noise and a far simpler consent conversation.

WHAT IT ANSWERS

Is there a disease-causing variant in a gene already known to cause human disease that explains this presentation?

WHEN IT IS DONE

Where a laboratory wants exome-level breadth without exome-level uncertainty, in settings with limited interpretive capacity, or where consent constraints make incidental discovery undesirable. Common as a first-line test in high-throughput diagnostic chains.

WHAT TRIGGERS IT

A suspected monogenic condition where the laboratory policy or the referring pathway specifies restricted-scope testing, or where the patient has declined broader analysis.

HOW IT IS DONE

The authorised gene list version is confirmed and recorded, coverage is certified per gene rather than globally, and variants and copy number changes within the list are called, annotated, prioritised and classified. The out-of-scope statement and reflex recommendation are generated as formal deliverables.

WHAT IT PRODUCES

A signed clinical report citing the gene list version, a per-gene coverage certificate, classified variants within scope, and an explicit scope limitation statement with reflex testing recommendations.

WHAT IT DETECTS

Coding and splice variants and exon-level copy number changes within the authorised disease gene list, at the depth and quality the assay is validated to deliver across that list.

WHAT IT DOES NOT DETECT

Anything outside the gene list, which includes every gene not yet associated with disease and every gene associated after the list version in force. Also misses deep intronic variants, structural rearrangements and repeat expansions. A negative result is bounded entirely by the list version and must be reported as such.

WHAT MAKES IT HARD

The gene list is the assay, and most laboratories govern it poorly. Without version control and a documented curation cycle, a negative result cannot be interpreted a year later because nobody can establish which genes were examined. Curation discipline matters more here than analytical skill.

WHO DOES THIS JOB

Clinical bioinformatician or variant scientist, mid level, working with a curation lead who owns the gene list under change control.

WHAT YOU CAN DO AFTERWARDS

Construct and govern a clinical gene list against gene-disease validity evidence, certify coverage per gene, and produce correctly bounded negative reports with defensible reflex recommendations.

WHO HIRES FOR IT

High-throughput diagnostic chains, hospital laboratories operating tiered testing pathways, and reference laboratories offering restricted-scope products alongside full exome.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
18	**88 hours · 15 days**	Linux command line, basic genetics	**2,05,000**

SCD-03
WORKFLOW

Clinical Exome Analysis

(**Clinical exome sequencing analysis** · CES analysis · medical exome)

COMPLETE WORKFLOW — 18 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Gene list version authorisation**	List version, curation date and gene-disease validity tier confirmed and recorded	Gene list register	*Authorised list version*
02	**Consent scope confirmation**	Restricted-scope consent confirmed with the patient informed of what lies outside scope	Consent record	*Scope confirmation*
03	**Data receipt and quality assessment**	Integrity verified and quality measured against validated acceptance criteria	md5sum, FastQC, MultiQC	*QC acceptance record*
04	**Read preprocessing**	Adapters and low-quality tails removed with retention accounting	fastp	*Trimmed FASTQ*
05	**Alignment and processing**	Reads aligned, duplicates marked and base qualities recalibrated within list intervals	bwa-mem2, Picard, GATK BQSR	*Analysis-ready BAM*
06	**Per-gene coverage certification**	Depth stated for every gene and exon on the list rather than as a global mean	mosdepth, bedtools coverage	*Per-gene coverage certificate*
07	**Identity and contamination verification**	Fingerprint captured and contamination estimated before interpretation	somalier, VerifyBamID2	*Identity clearance*
08	**Short variant calling**	Variants called across list intervals with splice-region and curated deep intronic padding	Validated caller	*Restricted VCF*
09	**Variant filtering**	Call set filtered with documented thresholds and metric review	GATK filters	*Filtered VCF*
10	**Targeted copy number calling**	Exon-level copy number called across list genes against a matched reference pool	gCNV, CNVkit or DECoN	*Gene-list CNV calls*
11	**Annotation and population filtering**	Consequence and frequency annotation with gene-specific thresholds where curated	VEP, gnomAD, ClinVar, SpliceAI	*Annotated variant set*
12	**Indication alignment check**	Confirmation that the applied list actually addresses the stated clinical indication	HPO terms against list content	*Alignment record*
13	**Phenotype-driven prioritisation**	Candidates ranked within the restricted gene set against phenotype	HPO scoring	*Ranked candidates*
14	**ACMG-AMP classification**	Evidence codes applied per variant with supporting data documented	ACMG-AMP with gene specifications	*Classified variants*
15	**Scope limitation documentation**	Explicit statement of what a restricted analysis cannot detect, with reflex options	Limitation template	*Limitation statement*
16	**Orthogonal confirmation**	Reportable findings confirmed by an independent method before release	Sanger, MLPA, ddPCR	*Confirmation record*
17	**Report generation**	Result, gene list version, per-gene coverage and reflex recommendation assembled	Validated report template	*Draft report*
18	**Clinical sign-out**	Report reviewed and authorised under dual signature with list version cited	Sign-out procedure	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
18	**88 h · 15 d**	**SCD-03**	**2,05,000**

SCD-04
SECTOR 01

Mendeliome Analysis

(**Comprehensive Mendelian gene set analysis** · Mendeliome)

WHY IT IS DONE

Sequencing several thousand established Mendelian genes to depths a full exome cannot economically reach converts coverage from a limitation into a strength. Reliable mosaicism sensitivity, dependable single-exon copy number resolution and near-zero gaps in the genes that matter are all consequences of that depth.

WHAT IT ANSWERS

Is there a disease-causing variant in any established Mendelian disease gene, including variants present at low mosaic fraction that shallower testing would miss?

WHEN IT IS DONE

Where a laboratory wants broad Mendelian coverage with depth sufficient for mosaicism and confident copy number calling, and where the additional discovery breadth of a research-grade exome is not required.

WHAT TRIGGERS IT

Suspected Mendelian disease across an uncertain differential, a presentation where mosaicism is plausible, or a laboratory pathway that positions the Mendeliome between panel and exome.

HOW IT IS DONE

The gene set version is authorised and recorded, coverage is certified per gene against the elevated threshold the assay is validated to, and variants are called at a depth that permits confident low-fraction calling. Mosaicism sensitivity is stated per gene, copy number is called at exon level, and findings are classified and confirmed.

WHAT IT PRODUCES

A signed clinical report, a high-depth coverage certificate stated per gene, a mosaicism sensitivity statement, exon-level copy number findings, and full classification records.

WHAT IT DETECTS

Coding and splice variants across the full Mendelian gene set, low-level mosaic variants at fractions standard exome depth cannot support, and single-exon copy number changes with genuinely usable resolution.

WHAT IT DOES NOT DETECT

Genes outside the curated Mendelian set, including newly associated genes and candidate genes under investigation. Deep intronic, regulatory and structural causes remain outside scope, as do repeat expansions and methylation defects.

WHAT MAKES IT HARD

The commercial claim is depth, so the depth must be demonstrated per gene rather than asserted as an average. Mosaicism sensitivity has to be translated from depth into a stated minimum detectable fraction, which most laboratories never do, leaving the assay's principal advantage undocumented.

WHO DOES THIS JOB

Clinical bioinformatician or clinical scientist, mid to senior level, in a laboratory operating a tiered testing menu.

WHAT YOU CAN DO AFTERWARDS

Operate a high-depth Mendelian assay, translate achieved depth into a defensible mosaicism sensitivity claim per gene, and call single-exon copy number reliably.

WHO HIRES FOR IT

Diagnostic chains with tiered product menus, hospital laboratories, and reference laboratories differentiating on depth rather than breadth.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
17	**84 hours · 14 days**	Linux command line, basic genetics	**1,95,000**

SCD-04
WORKFLOW

Mendeliome Analysis

(**Comprehensive Mendelian gene set analysis** · Mendeliome)

COMPLETE WORKFLOW — 17 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Gene set version authorisation**	Mendelian gene set version, curation date and evidence tiers confirmed	Gene set register	*Authorised gene set*
02	**Consent and scope verification**	Scope and secondary findings election confirmed before analysis	Consent record	*Analysis authorisation*
03	**Data receipt and quality assessment**	Integrity verified and quality measured against validated criteria	md5sum, FastQC, MultiQC	*QC acceptance record*
04	**Read preprocessing**	Adapters and low-quality tails removed with retention accounting	fastp	*Trimmed FASTQ*
05	**Alignment and processing**	Reads aligned, duplicates marked and base qualities recalibrated across the gene set	bwa-mem2, Picard, GATK BQSR	*Analysis-ready BAM*
06	**High-depth coverage certification**	Per-gene and per-exon depth certified against the elevated validated threshold	mosdepth, bedtools coverage	*Coverage certificate*
07	**Identity and contamination verification**	Fingerprint captured and contamination estimated before interpretation	somalier, VerifyBamID2	*Identity clearance*
08	**Short variant calling at depth**	Variants called at depth permitting confident reduced-allele-fraction calling	Validated caller	*Mendeliome VCF*
09	**Mosaicism sensitivity screening**	Allele fractions modelled against depth to surface candidate mosaic variants	Allele fraction analysis	*Mosaic candidates*
10	**Variant filtering**	Call set filtered with documented thresholds and metric review	GATK filters	*Filtered VCF*
11	**Exon-level copy number calling**	Single-exon copy number called using the uniform high depth of the assay	gCNV, CNVkit or DECoN	*Exon-level CNV calls*
12	**Annotation and population filtering**	Consequence and frequency annotation on MANE Select with splice prediction	VEP, gnomAD, ClinVar, SpliceAI	*Annotated variant set*
13	**Phenotype-driven prioritisation**	Candidates ranked against HPO terms with gene-disease validity weighting	Exomiser, HPO terms	*Ranked candidates*
14	**ACMG-AMP classification**	Evidence codes applied per variant with full reasoning documented	ACMG-AMP with gene specifications	*Classified variants*
15	**Mosaicism sensitivity statement**	Achieved depth translated into minimum detectable mosaic fraction per gene	Depth-based sensitivity modelling	*Sensitivity statement*
16	**Orthogonal confirmation**	Reportable findings confirmed including mosaics at matched sensitivity	Sanger, ddPCR	*Confirmation record*
17	**Report generation and sign-out**	Result, gene set version, coverage and sensitivity issued under dual signature	Validated report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
17	**84 h · 14 d**	**SCD-04**	**1,95,000**

SCD-05
SECTOR 01

Targeted Gene Panel Analysis

(**Capture panel sequencing analysis** · Panel analysis · NGS panel)

WHY IT IS DONE

Where the clinical question points to a defined set of genes, a focused panel delivers the answer faster, cheaper and at higher depth than any broader test, with a fraction of the interpretive burden. It remains the highest-volume clinical NGS assay in most laboratories by a wide margin.

WHAT IT ANSWERS

Is there a disease-causing variant in the specific genes associated with this defined clinical indication?

WHEN IT IS DONE

First-line where the presentation clearly indicates a gene group — cardiomyopathy, hereditary cancer, epilepsy, hearing loss, immunodeficiency — and where a rapid, high-depth, well-validated answer is preferred to breadth.

WHAT TRIGGERS IT

A clinical presentation matching an established indication-specific panel, a family history in a known disease group, or a screening pathway that specifies panel testing.

HOW IT IS DONE

The panel design version and its validated copy number claim scope are confirmed, coverage is certified per amplicon or exon, and variants and copy number changes are called with parameters validated for that specific design. Regions the design cannot resolve are routed to documented alternative methods before classification and sign-out.

WHAT IT PRODUCES

A signed clinical report citing the design version, a per-amplicon coverage certificate, panel copy number findings with per-gene sensitivity, and the fill-in results for any region requiring an alternative method.

WHAT IT DETECTS

Coding and splice variants across the panel genes at high depth, exon-level copy number changes where the panel carries a validated copy number claim, and low-fraction variants supported by the elevated depth.

WHAT IT DOES NOT DETECT

Any gene outside the panel design, which is the principal risk when the clinical suspicion turns out to be wrong. Also misses deep intronic and regulatory variation, structural rearrangements, and any region the design covers poorly, particularly pseudogene-bearing and repeat-rich genes.

WHAT MAKES IT HARD

Every panel has blind spots, and the discipline is declaring them rather than ignoring them. Panel-wide copy number claims are almost never true, and pseudogene-bearing genes on clinical panels are precisely where disease occurs, so routing them to a valid alternative method is essential rather than optional.

WHO DOES THIS JOB

Clinical bioinformatician or variant scientist, entry to mid level. The most accessible route into clinical genomics and the highest-volume role in the sector.

WHAT YOU CAN DO AFTERWARDS

Operate and validate an indication-specific panel, certify coverage per target, make a defensible per-gene copy number claim, and manage difficult regions with documented alternative methods.

WHO HIRES FOR IT

Every diagnostic laboratory offering genetic testing, hospital laboratories, oncology and cardiology services, and private diagnostic chains at scale.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
16	**78 hours · 13 days**	Linux command line, basic genetics	**1,80,000**

SCD-05
WORKFLOW

Targeted Gene Panel Analysis

(**Capture panel sequencing analysis** · Panel analysis · NGS panel)

COMPLETE WORKFLOW — 16 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Panel design version verification**	Design version, gene content, coverage intent and documented blind spots confirmed	Panel design register	*Authorised panel version*
02	**Consent and indication confirmation**	Indication-scoped consent confirmed and the referral question recorded	Consent record, requisition	*Analysis authorisation*
03	**Data receipt and quality assessment**	Integrity verified and quality measured against panel-specific acceptance criteria	md5sum, FastQC, MultiQC	*QC acceptance record*
04	**Chemistry-appropriate preprocessing**	Adapters removed and primers trimmed where the design is amplicon-based	fastp, primer trimming	*Preprocessed FASTQ*
05	**Alignment and duplicate handling**	Reads aligned with duplicate marking applied or suppressed according to chemistry	bwa-mem2, Picard	*Analysis-ready BAM*
06	**Per-target coverage certification**	Depth certified per amplicon or exon against the panel's validated threshold	mosdepth, bedtools coverage	*Coverage certificate*
07	**Identity and contamination verification**	Fingerprint captured and contamination estimated before interpretation	somalier, VerifyBamID2	*Identity clearance*
08	**High-depth variant calling**	Variants called with parameters tuned to the panel's depth and chemistry	Validated caller	*Panel VCF*
09	**Variant filtering**	Call set filtered with panel-validated thresholds and metric review	GATK filters	*Filtered VCF*
10	**Reference pool verification**	Matched normals from the same design and batch confirmed compatible	Internal reference pool	*Authorised pool*
11	**Panel copy number calling**	Exon-level copy number called across genes carrying a validated CNV claim	DECoN, CNVkit	*Panel CNV calls*
12	**Difficult region handling**	Pseudogene-bearing and repeat-rich genes routed to documented alternative methods	MLPA, ddPCR, long-read	*Fill-in results*
13	**Annotation and population filtering**	Consequence and frequency annotation with indication-specific thresholds	VEP, gnomAD, ClinVar, SpliceAI	*Annotated variant set*
14	**ACMG-AMP classification**	Evidence codes applied per variant with supporting data documented	ACMG-AMP with gene specifications	*Classified variants*
15	**Orthogonal confirmation**	Reportable sequence and copy number findings confirmed independently	Sanger, MLPA, ddPCR	*Confirmation record*
16	**Report generation and sign-out**	Result, design version, coverage and declared blind spots issued under dual signature	Validated report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
16	**78 h · 13 d**	**SCD-05**	**1,80,000**

SCD-06
SECTOR 01

Amplicon Panel Analysis

(**Amplicon sequencing analysis** · Amplicon analysis)

WHY IT IS DONE

Amplicon designs deliver very high depth over small target regions with minimal input DNA and short turnaround, which suits focused clinical questions, degraded samples and point-of-need testing. The analysis differs materially from capture and applying capture logic to amplicon data produces systematically wrong results.

WHAT IT ANSWERS

What variants are present in a small, precisely defined set of targets, at what allele fraction, and at what depth?

WHEN IT IS DONE

Where the target region is small and fixed, input material is limited or degraded, turnaround must be short, or very high depth is required for low-fraction detection. Common in focused hereditary panels, confirmatory testing and resource-limited settings.

WHAT TRIGGERS IT

A focused clinical question over a small target set, limited or degraded sample material, or a pathway requiring rapid high-depth confirmation.

HOW IT IS DONE

Primer sequences are trimmed before analysis and duplicate marking is suppressed because amplicon reads share start positions by design. Per-amplicon coverage is certified, variants are called with amplicon-appropriate parameters, and primer-site variants are actively screened for as an allele dropout risk before classification.

WHAT IT PRODUCES

A signed clinical report, per-amplicon coverage and uniformity certificate, variant calls with allele fractions, and the allele dropout risk assessment for primer-overlapping sites.

WHAT IT DETECTS

Variants within the amplified regions at very high depth, including low allele fraction variants that capture-based depth would not support reliably.

WHAT IT DOES NOT DETECT

Anything outside the amplified regions, copy number changes in most designs because amplification destroys the depth-copy number relationship, and large insertions or deletions spanning primer sites. Allele dropout from a variant under a primer produces false homozygous calls and false negatives, which is the characteristic failure of the method.

WHAT MAKES IT HARD

Two errors dominate. Applying duplicate marking to amplicon data removes almost all the reads, and failing to screen for primer-site variants produces confident false homozygous calls. Amplicon uniformity is also poor by nature, so per-amplicon certification is mandatory rather than good practice.

WHO DOES THIS JOB

Clinical bioinformatician or laboratory scientist with analysis responsibility, entry to mid level.

WHAT YOU CAN DO AFTERWARDS

Process amplicon data correctly with primer-aware handling, certify per-amplicon coverage, and identify and manage allele dropout risk.

WHO HIRES FOR IT

Diagnostic laboratories running focused panels, hospital laboratories, point-of-care and decentralised testing services, and public health laboratories.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
14	**68 hours · 12 days**	Linux command line, basic genetics	**1,55,000**

SCD-06
WORKFLOW

Amplicon Panel Analysis

(**Amplicon sequencing analysis** · Amplicon analysis)

COMPLETE WORKFLOW — 14 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Amplicon design verification**	Design version, amplicon coordinates and primer sequences confirmed and recorded	Design register, primer file	*Authorised design version*
02	**Consent and indication confirmation**	Scope confirmed and the referral question recorded	Consent record, requisition	*Analysis authorisation*
03	**Data receipt and quality assessment**	Integrity verified and read quality measured against acceptance criteria	md5sum, FastQC	*QC acceptance record*
04	**Primer trimming**	Primer sequences removed from read ends before any variant analysis	cutadapt, iVar, fgbio	*Primer-trimmed FASTQ*
05	**Alignment without duplicate marking**	Reads aligned with duplicate marking suppressed since amplicons share start positions	bwa-mem2	*Analysis-ready BAM*
06	**Per-amplicon coverage certification**	Depth and uniformity certified for every amplicon against the validated threshold	mosdepth, bedtools coverage	*Coverage certificate*
07	**Amplicon uniformity assessment**	Inter-amplicon depth variation measured with under-performing amplicons named	Coverage ratio analysis	*Uniformity record*
08	**Identity and contamination verification**	Fingerprint captured where the design permits and contamination assessed	somalier, allele balance review	*Identity clearance*
09	**High-depth variant calling**	Variants called with amplicon-appropriate parameters at very high depth	Validated caller, amplicon mode	*Amplicon VCF*
10	**Allele dropout screening**	Primer-overlapping variants identified as a false homozygous call risk	Primer coordinate intersection	*Dropout risk assessment*
11	**Variant filtering**	Call set filtered with strand and position bias review appropriate to amplicon data	GATK filters, bias metrics	*Filtered VCF*
12	**Annotation and population filtering**	Consequence and frequency annotation applied on validated transcripts	VEP, gnomAD, ClinVar	*Annotated variant set*
13	**ACMG-AMP classification**	Evidence codes applied per variant with supporting data documented	ACMG-AMP with gene specifications	*Classified variants*
14	**Confirmation and report sign-out**	Findings confirmed independently and the report issued under dual signature	Sanger, validated template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
14	**68 h · 12 d**	**SCD-06**	**1,55,000**

SCD-07
SECTOR 01

Virtual Panel Analysis

(**In-silico panel analysis** · Virtual panel · slice)

WHY IT IS DONE

Sequencing broadly and analysing narrowly has become the dominant clinical model. It contains the burden of uncertain findings and simplifies consent while leaving the underlying data available for later expansion without re-sequencing the patient, which is both faster and cheaper than repeat testing.

WHAT IT ANSWERS

Is there a disease-causing variant in this specific gene set, within data that were already generated under a broader assay?

WHEN IT IS DONE

Applied to existing exome or genome data, either at the time of first analysis or months and years later when a new clinical question arises. Increasingly the default operating model for laboratories running genome-first pathways.

WHAT TRIGGERS IT

A new or refined clinical question against existing data, a tiered testing policy that restricts first-pass analysis, or a request to apply a newly published gene panel to an archived case.

HOW IT IS DONE

Both the panel version and the backbone assay version are recorded, sample identity is re-verified against the current requisition, and coverage is recomputed specifically for the panel genes rather than inherited from backbone metrics. Variants are extracted, classified and reported with an explicit out-of-panel handling statement.

WHAT IT PRODUCES

A signed clinical report citing both panel and backbone versions, a panel-specific coverage certificate, classified variants within scope, and the documented out-of-panel handling record.

WHAT IT DETECTS

Variants and copy number changes within the authorised virtual panel genes, at whatever quality the underlying backbone assay achieved in those specific genes.

WHAT IT DOES NOT DETECT

Anything outside the applied panel, and anything the backbone assay itself could not detect. The critical limitation is inherited: a virtual panel applied to exome data carries every exome limitation, and backbone coverage in the panel genes may be far worse than the backbone's global average suggests.

WHAT MAKES IT HARD

Two versions govern every result and both must be recorded or the report cannot be reproduced. Recomputing coverage for the panel genes is routinely skipped, so laboratories inherit a global backbone claim that is simply untrue for the specific genes being reported.

WHO DOES THIS JOB

Variant scientist or clinical bioinformatician, entry to mid level, often the highest-throughput analytical role in a genome-first laboratory.

WHAT YOU CAN DO AFTERWARDS

Apply and govern virtual panels against archived data, recompute and certify panel-specific coverage, and produce correctly bounded reports citing both governing versions.

WHO HIRES FOR IT

Laboratories operating genome-first or exome-first pathways, national genomic medicine programmes, and diagnostic chains running tiered analytical products off a single sequencing backbone.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
12	**58 hours · 10 days**	Familiarity with VCF and clinical annotation	**1,30,000**

SCD-07
WORKFLOW

Virtual Panel Analysis

(**In-silico panel analysis** · Virtual panel · slice)

COMPLETE WORKFLOW — 12 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Virtual panel version authorisation**	Panel version, curation source and date confirmed and recorded	Panel list register	*Authorised panel version*
02	**Backbone provenance verification**	Original assay, pipeline version and sequencing date established for the source data	Original case record	*Provenance record*
03	**Consent scope confirmation**	Confirmation that the applied panel falls within the originally consented scope	Consent record	*Scope confirmation*
04	**Archive retrieval and integrity check**	Archived BAM or CRAM retrieved and integrity and reference decodability verified	Checksum verification	*Verified archived data*
05	**Identity re-verification**	Fingerprint compared to confirm the data belong to the patient on the current requisition	somalier fingerprint comparison	*Identity clearance*
06	**Panel-specific coverage recomputation**	Depth recomputed for the panel genes specifically rather than inherited from backbone metrics	mosdepth, bedtools coverage	*Panel coverage certificate*
07	**Variant extraction and restriction**	Variants extracted and restricted to authorised panel intervals with clinical padding	bcftools, interval restriction	*Restricted variant set*
08	**Copy number restriction and reassessment**	Backbone copy number calls restricted to panel genes with sensitivity reassessed at gene level	Interval intersection	*Panel-restricted CNV*
09	**Annotation and population filtering**	Consequence and frequency annotation applied with panel-appropriate transcripts	VEP, gnomAD, ClinVar	*Annotated variant set*
10	**ACMG-AMP classification**	Evidence codes applied per variant within the restricted gene set	ACMG-AMP with gene specifications	*Classified variants*
11	**Out-of-panel handling record**	Documented policy on findings outside the applied panel recorded as applied	Laboratory policy	*Policy application record*
12	**Report generation and sign-out**	Result citing both panel and backbone versions issued under dual signature	Validated report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
12	**58 h · 10 d**	**SCD-07**	**1,30,000**

SCD-08
SECTOR 01

Trio Analysis

(**Proband-parent-parent family analysis** · Trio · family analysis)

WHY IT IS DONE

Adding both parents transforms interpretation. De novo variants become directly identifiable rather than inferred, compound heterozygous pairs can be proven to sit on opposite alleles, and the candidate list typically contracts by an order of magnitude. Diagnostic yield rises substantially for the same sequencing chemistry.

WHAT IT ANSWERS

Which candidate variants arose de novo, which were inherited and from which parent, and do any inheritance patterns fit the presentation and give a defensible recurrence risk?

WHEN IT IS DONE

Wherever both parents are available and the presentation suggests a dominant de novo or autosomal recessive mechanism. Standard of care in paediatric developmental disorders, and increasingly the default rather than the escalation.

WHAT TRIGGERS IT

A sporadic severe presentation in a child with unaffected parents, a suspected recessive condition, consanguinity, or a singleton result that identified candidates requiring segregation evidence.

HOW IT IS DONE

Pedigree is verified from the data before any inheritance logic is applied, all three samples are processed identically, and joint genotyping recovers low-confidence parental calls. De novo candidates are filtered stringently and reviewed at read level in all three individuals, inheritance models are applied in parallel, and compound heterozygotes are phased.

WHAT IT PRODUCES

A signed clinical report with recurrence risk, the pedigree verification record, the de novo call set with read-level evidence, inheritance model filtering results, compound heterozygote phasing evidence, and any relationship discrepancy record.

WHAT IT DETECTS

De novo single nucleotide and structural variants, confirmed compound heterozygous pairs in trans, correctly assigned inheritance for every candidate, and parent-of-origin for imprinting-relevant findings.

WHAT IT DOES NOT DETECT

Nothing additional to what the underlying assay detects — trio analysis improves interpretation, not detection. It cannot resolve variants in regions where any of the three samples is inadequately covered, and it does not compensate for germline mosaicism in a parent, which produces apparent de novo variants that carry real recurrence risk.

WHAT MAKES IT HARD

De novo calls carry the strongest interpretive weight and the highest false positive rate simultaneously, so read-level review of every candidate is unavoidable. Poor coverage in one parent silently converts inherited variants into apparent de novo events. Non-paternity is discovered as a byproduct and must be handled under a policy agreed before analysis, not improvised at sign-out.

WHO DOES THIS JOB

Clinical bioinformatician or clinical scientist, mid to senior level, working closely with a genetic counsellor on the family and disclosure implications.

WHAT YOU CAN DO AFTERWARDS

Verify pedigrees from sequence data, produce a defensible de novo call set with read-level evidence, apply inheritance models correctly, phase compound heterozygotes, and derive recurrence risk.

WHO HIRES FOR IT

Paediatric and developmental genomics services, national rare disease programmes, hospital genomics laboratories, and reference laboratories offering family-based testing.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
22	**110 hours · 19 days**	Prior exome or genome analysis experience	**2,60,000**

SCD-08
WORKFLOW

Trio Analysis

(**Proband-parent-parent family analysis** · Trio · family analysis)

COMPLETE WORKFLOW — 22 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Pedigree and consent verification**	Consent obtained from all three individuals including the relationship-discrepancy policy	PED file, consent records	*Analysis authorisation*
02	**Data receipt and integrity check**	Integrity verified for all three samples against their run records	md5sum, seqkit stats	*Verified trio FASTQ*
03	**Per-sample quality assessment**	Each of the three samples independently measured against validated acceptance criteria	FastQC, MultiQC	*Per-sample QC records*
04	**Read preprocessing**	Adapters and low-quality tails removed identically across all three samples	fastp	*Trimmed FASTQ*
05	**Uniform alignment and processing**	All three samples aligned and processed with one identical pipeline version	bwa-mem2, Picard, GATK BQSR	*Analysis-ready BAMs*
06	**Identity fingerprinting**	Fixed panel genotypes captured for each family member individually	somalier extract	*Family fingerprints*
07	**Relatedness and pedigree confirmation**	Kinship coefficients verified against the declared pedigree before inheritance logic	somalier relate, KING	*Pedigree verification*
08	**Contamination estimation**	Contamination fraction estimated separately for each family member	VerifyBamID2	*Contamination records*
09	**Trio callable region analysis**	Regions where all three samples reached diagnostic depth defined and declared	mosdepth, interval intersection	*Trio callable BED*
10	**Per-sample variant calling**	Per-sample gVCFs generated with consistent caller version and parameters	HaplotypeCaller or DeepVariant	*Per-sample gVCFs*
11	**Joint genotyping**	Family-aware joint genotyping applied to recover low-confidence parental genotypes	GATK GenotypeGVCFs	*Joint trio VCF*
12	**Variant filtering**	Joint call set filtered by recalibration or documented hard filters	GATK VQSR	*Filtered trio VCF*
13	**De novo candidate identification**	Trio genotype logic applied with stringent depth and allele balance filters	Trio genotype logic	*De novo candidates*
14	**De novo read-level review**	Every de novo candidate inspected at read level in proband and both parents	IGV review	*Reviewed de novo set*
15	**Inheritance model filtering**	Recessive, compound heterozygous, X-linked and de novo models applied in parallel	PED-aware filtering	*Model-filtered candidates*
16	**Compound heterozygote phasing**	Variant pairs confirmed to lie on opposite parental alleles rather than in cis	Trio-based phasing	*Phased compound pairs*
17	**Trio copy number and structural analysis**	Copy number and structural variants called with inheritance established across the trio	gCNV, Manta	*Trio CNV and SV calls*
18	**Annotation and population filtering**	Consequence and frequency annotation applied with model-specific thresholds	VEP, gnomAD, ClinVar, SpliceAI	*Annotated candidates*
19	**Phenotype-driven prioritisation**	Candidates ranked combining phenotype scoring with segregation evidence	Exomiser, HPO terms	*Ranked candidates*
20	**ACMG-AMP classification**	Evidence codes applied including de novo criteria where parentage is confirmed	ACMG-AMP with gene specifications	*Classified variants*
21	**Orthogonal confirmation**	Findings confirmed in the proband and demonstrated absent in both parents	Sanger across the trio	*Confirmed de novo status*
22	**Report generation and sign-out**	Result, inheritance evidence and recurrence risk issued under dual signature	Validated trio report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
22	**110 h · 19 d**	**SCD-08**	**2,60,000**

SCD-09
SECTOR 01

Duo Analysis

(**Proband-single parent analysis** · Duo)

WHY IT IS DONE

One parent delivers a meaningful share of trio benefit at lower cost, and is frequently the only option available where a parent is deceased, unavailable or declines testing. Handled properly it narrows the candidate list substantially; handled carelessly it produces overstated de novo claims.

WHAT IT ANSWERS

Which candidates were inherited from the tested parent, and which were not — recognising that the second category remains unresolved between de novo and inheritance from the untested parent?

WHEN IT IS DONE

Where only one parent can be sampled, in single-parent families, or where a trio is planned but one sample is delayed and the clinical timeline cannot wait.

WHAT TRIGGERS IT

A monogenic presentation with only one parent available, or an urgent case where waiting for the second parent would delay a clinically necessary result.

HOW IT IS DONE

The relationship is verified from the data, both samples are processed identically and jointly genotyped, and candidates are partitioned into inherited-from-tested-parent and not-present-in-tested-parent. The second category is carried forward with an explicit unresolved flag, and the extension recommendation for testing the second parent is generated as a formal deliverable.

WHAT IT PRODUCES

A signed clinical report stating unresolved elements plainly, the relationship verification record, inheritance partitioning results, the flagged candidate de novo set, partial phasing evidence, and the second-parent extension recommendation.

WHAT IT DETECTS

Inheritance from the tested parent, candidate de novo events flagged as unresolved, and partial phasing of compound heterozygous pairs where the tested parent carries one of the two variants.

WHAT IT DOES NOT DETECT

Confirmed de novo status, which requires both parents and cannot be inferred from one. It also cannot exclude inheritance from the untested parent, cannot fully phase compound heterozygotes where the tested parent carries neither variant, and gives no recurrence risk with trio-level confidence.

WHAT MAKES IT HARD

The temptation is to present candidate de novo variants as de novo, which is a serious overstatement and the characteristic failure of duo reporting. The interpretive framework is the deliverable as much as the variant list, and recurrence risk cannot be stated with trio confidence.

WHO DOES THIS JOB

Clinical bioinformatician or variant scientist, mid level, with genetic counselling input on how unresolved inheritance is communicated.

WHAT YOU CAN DO AFTERWARDS

Run duo analysis correctly, partition inheritance defensibly, recognise and state the limits of duo-derived de novo claims, and construct an extension recommendation that secures the completing sample.

WHO HIRES FOR IT

Hospital genomics laboratories, rare disease services, and any laboratory offering family-based testing where complete trios are not always achievable.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
17	**82 hours · 14 days**	Prior exome or genome analysis experience	**1,90,000**

SCD-09
WORKFLOW

Duo Analysis

(**Proband-single parent analysis** · Duo)

COMPLETE WORKFLOW — 17 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Relationship and consent verification**	Declared relationship recorded and consent obtained from both individuals	PED file, consent records	*Analysis authorisation*
02	**Data receipt and integrity check**	Integrity verified for both samples against their run records	md5sum, seqkit stats	*Verified duo FASTQ*
03	**Per-sample quality assessment**	Both samples independently measured against validated acceptance criteria	FastQC, MultiQC	*Per-sample QC records*
04	**Read preprocessing**	Adapters and low-quality tails removed identically across both samples	fastp	*Trimmed FASTQ*
05	**Uniform alignment and processing**	Both samples aligned and processed with one identical pipeline version	bwa-mem2, Picard, GATK BQSR	*Analysis-ready BAMs*
06	**Identity fingerprinting**	Fixed panel genotypes captured for both individuals	somalier extract	*Fingerprints*
07	**Relationship confirmation**	Kinship coefficient verified before any inheritance partitioning is trusted	somalier relate, KING	*Relationship verification*
08	**Contamination estimation**	Contamination fraction estimated separately for each individual	VerifyBamID2	*Contamination records*
09	**Duo-informative region analysis**	Regions where both samples reached diagnostic depth defined and declared	mosdepth, interval intersection	*Duo callable BED*
10	**Joint genotyping**	Pairwise joint genotyping applied to improve parental genotype confidence	GATK GenotypeGVCFs	*Joint duo VCF*
11	**Inheritance partitioning**	Candidates partitioned into inherited from the tested parent and not present in that parent	Duo genotype logic	*Partitioned candidates*
12	**Candidate de novo flagging**	Variants absent in the tested parent flagged as unresolved between de novo and other-parent inheritance	Explicit uncertainty flag	*Flagged candidate set*
13	**Partial compound heterozygote resolution**	Trans configuration inferred where the tested parent carries one of the two variants	Duo phasing logic	*Partially phased pairs*
14	**Annotation and prioritisation**	Consequence and frequency annotation applied with duo-appropriate phenotype weighting	VEP, gnomAD, Exomiser	*Ranked candidates*
15	**ACMG-AMP classification**	Evidence codes applied with de novo criteria explicitly recorded as unavailable	ACMG-AMP with gene specifications	*Classified variants*
16	**Extension recommendation**	Diagnostic gain from testing the second parent quantified and recommended	Documented criteria	*Extension recommendation*
17	**Report generation and sign-out**	Result with unresolved elements stated plainly, issued under dual signature	Validated duo report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
17	**82 h · 14 d**	**SCD-09**	**1,90,000**

SCD-10
SECTOR 01

Family Segregation Analysis

(**Segregation and cascade testing analysis** · Segregation study)

WHY IT IS DONE

A variant of uncertain significance can often be resolved by testing whether it tracks with disease across a family. Segregation is one of the few evidence lines that can move a variant decisively toward pathogenic or benign, and it uses samples the family can readily provide.

WHAT IT ANSWERS

Does this variant co-segregate with the phenotype across affected and unaffected relatives, and does that pattern strengthen or weaken the case for pathogenicity?

WHEN IT IS DONE

After a variant of uncertain significance has been identified and the family structure offers informative meioses. Also used for cascade testing once a pathogenic variant is confirmed in a proband.

WHAT TRIGGERS IT

A VUS in a gene consistent with the family phenotype, a request to test at-risk relatives of a confirmed proband, or a reclassification review requiring additional evidence.

HOW IT IS DONE

The pedigree is verified from data before segregation is trusted, informative meioses are counted, and each relative is genotyped for the specific variant by a targeted method. The segregation pattern is scored against ACMG-AMP criteria and the resulting evidence strength is documented for reclassification.

WHAT IT PRODUCES

A segregation report with the annotated pedigree, per-relative genotype results, the informative meiosis count, the ACMG-AMP evidence weight derived, and cascade testing reports for each relative tested.

WHAT IT DETECTS

Presence or absence of the specific familial variant in each relative tested, the segregation pattern across the pedigree, and the statistical weight that pattern contributes to classification.

WHAT IT DOES NOT DETECT

Any variant other than the one being tracked, so a relative can be negative for the familial variant and still have a genetic condition from another cause. Segregation is also confounded by incomplete penetrance, phenocopies, variable expressivity and small pedigrees that generate too few informative meioses to matter.

WHAT MAKES IT HARD

Small families rarely generate enough informative meioses to shift a classification, and the analysis is frequently performed without first assessing whether it can possibly produce useful evidence. Incomplete penetrance is routinely misread as failed segregation, and phenotyping of relatives is often too weak to support the conclusion drawn.

WHO DOES THIS JOB

Variant scientist or clinical scientist, mid level, working directly with genetic counselling on family recruitment and phenotype ascertainment.

WHAT YOU CAN DO AFTERWARDS

Assess in advance whether a family can generate decisive segregation evidence, execute cascade testing, and score segregation correctly under ACMG-AMP for reclassification.

WHO HIRES FOR IT

Clinical genetics services, hereditary cancer and cardiac genetics units, hospital laboratories, and any service running variant reclassification programmes.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
13	**62 hours · 11 days**	Familiarity with VCF and variant classification	**1,40,000**

SCD-10
WORKFLOW

Family Segregation Analysis

(**Segregation and cascade testing analysis** · Segregation study)

COMPLETE WORKFLOW — 13 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Referral and variant scoping**	The specific familial variant and the clinical question defined and recorded	Requisition, proband report	*Scoped referral*
02	**Pedigree construction and review**	Family structure documented with affected and unaffected status recorded per relative	Pedigree drawing standard	*Annotated pedigree*
03	**Informative meiosis assessment**	Whether the family can generate decisive evidence assessed before testing proceeds	Segregation power assessment	*Feasibility record*
04	**Consent for relatives**	Individual consent obtained from each relative including incidental relationship findings	Consent records	*Analysis authorisation*
05	**Sample receipt and identity capture**	Each relative's sample received and an identity fingerprint captured	somalier extract	*Relative fingerprints*
06	**Relatedness verification**	Declared relationships verified from genotype data before segregation is trusted	somalier relate, KING	*Relationship verification*
07	**Targeted genotyping of relatives**	Each relative genotyped specifically for the familial variant by a targeted method	Sanger, ddPCR, targeted NGS	*Per-relative genotypes*
08	**Genotype quality review**	Each genotype reviewed for depth, allele balance and technical adequacy	Read-level review	*Verified genotypes*
09	**Segregation pattern analysis**	Co-segregation with phenotype assessed across affected and unaffected relatives	Segregation analysis	*Segregation pattern*
10	**Penetrance and phenocopy assessment**	Incomplete penetrance, phenocopy and variable expressivity considered before conclusion	Clinical correlation review	*Confounder assessment*
11	**ACMG-AMP evidence scoring**	Segregation scored under the applicable evidence criteria at the correct strength	ACMG-AMP PP1 or BS4	*Evidence weight*
12	**Variant reclassification review**	The variant reassessed with segregation evidence incorporated into the full evidence set	ACMG-AMP reassessment	*Reclassification outcome*
13	**Reporting and family communication**	Segregation report and individual cascade reports issued with counselling support	Validated report templates	*Signed reports*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
13	**62 h · 11 d**	**SCD-10**	**1,40,000**

SCD-11
SECTOR 01

Cohort Joint Analysis

(**Joint genotyping and cohort-scale analysis** · Joint genotyping)

WHY IT IS DONE

Calling samples jointly rather than individually recovers genotypes at sites where a single sample had marginal evidence, produces a consistent variant representation across the cohort, and enables internal allele frequencies that are far more useful for filtering than any public database.

WHAT IT ANSWERS

What is the consistent genotype for every sample at every variant site across the cohort, and what is the internal frequency of each variant in this population?

WHEN IT IS DONE

Wherever a laboratory accumulates samples over time and wants comparable calls across them, in family and population studies, and in any programme building an internal frequency reference to support diagnostic filtering.

WHAT TRIGGERS IT

Accumulation of a defined sample set, a research or programme cohort requiring uniform calls, or a laboratory initiative to build an internal allele frequency resource.

HOW IT IS DONE

Per-sample gVCFs are consolidated into a scalable store, jointly genotyped, and filtered by variant quality recalibration or documented hard filters. Sample-level quality control covering relatedness, ancestry, contamination and sex is applied across the cohort, and an internal frequency database is generated and versioned.

WHAT IT PRODUCES

A joint cohort VCF, the sample-level quality control matrix, the relatedness and ancestry report, an internal allele frequency database with version, and the documented sample exclusion register.

WHAT IT DETECTS

Consistent genotypes across all samples including reference and no-call states, variants recovered at marginal sites through joint evidence, internal allele frequencies, and cohort-level quality outliers.

WHAT IT DOES NOT DETECT

Anything the individual assays could not detect. Joint calling improves genotype consistency and marginal recovery; it does not extend the callable region. It also cannot correct batch effects, which it will faithfully propagate across the whole cohort if they are not identified first.

WHAT MAKES IT HARD

Scale is the engineering problem and batch effect is the scientific one. A cohort assembled across chemistry or pipeline versions will carry systematic differences that joint calling propagates rather than resolves, and internal frequency databases built on unrecognised batch structure produce filtering that removes real findings.

WHO DOES THIS JOB

Clinical bioinformatician or computational genomics specialist, mid to senior level, often with responsibility for the laboratory's internal frequency resource.

WHAT YOU CAN DO AFTERWARDS

Operate joint genotyping at scale, apply cohort-level quality control, detect batch structure before it propagates, and build and govern an internal allele frequency database.

WHO HIRES FOR IT

National genomics programmes, biobanks, large diagnostic chains, hospital laboratories building internal frequency resources, and population genomics initiatives.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
19	**92 hours · 16 days**	Linux, scripting, prior variant calling experience	**2,15,000**

SCD-11
WORKFLOW

Cohort Joint Analysis

(**Joint genotyping and cohort-scale analysis** · Joint genotyping)

COMPLETE WORKFLOW — 19 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Cohort definition and scope**	Sample set, inclusion criteria and analytical objectives defined and recorded	Cohort register	*Cohort definition*
02	**Metadata assembly**	Batch, run, lane, extraction date, kit lot and instrument attached to every sample	LIMS export	*Annotated sample manifest*
03	**Pipeline version harmonisation**	Confirmation that all samples were processed under compatible pipeline versions	Version audit	*Harmonisation record*
04	**gVCF ingestion and validation**	Per-sample gVCFs validated for completeness, reference build and interval consistency	bcftools, GATK ValidateVariants	*Validated gVCF set*
05	**Genomic datastore construction**	gVCFs consolidated into a scalable store designed for cohort-wide genotyping	GenomicsDBImport or Hail	*Genomic datastore*
06	**Joint genotyping**	All samples genotyped together to recover marginal sites and standardise representation	GATK GenotypeGVCFs	*Joint cohort VCF*
07	**Variant quality recalibration**	Call set filtered by recalibration models or documented hard filters	GATK VQSR	*Filtered cohort VCF*
08	**Variant representation normalisation**	Multi-allelic sites split and indels left-aligned for consistent downstream comparison	bcftools norm	*Normalised VCF*
09	**Per-sample quality metrics**	Call rate, transition-transversion ratio, heterozygosity and singleton burden computed per sample	bcftools stats	*Sample QC matrix*
10	**Contamination screening**	Contamination estimated across the cohort with outliers identified	VerifyBamID2	*Contamination results*
11	**Relatedness estimation**	Kinship computed across all sample pairs with unexpected relationships flagged	KING, somalier	*Relatedness matrix*
12	**Ancestry inference**	Samples projected onto a reference panel to assign ancestry for correct frequency use	PCA projection	*Ancestry assignment*
13	**Sex verification**	Inferred sex compared against records with aneuploidy and discrepancy flagged	X/Y coverage, X heterozygosity	*Sex check results*
14	**Batch effect assessment**	Technical variables tested for association with cohort structure and quality metrics	Variance partitioning	*Batch effect report*
15	**Sample exclusion decisions**	Exclusions applied against pre-registered criteria with justification recorded	Documented criteria	*Exclusion register*
16	**Internal allele frequency generation**	Cohort frequencies computed and versioned for use in diagnostic filtering	bcftools, frequency pipeline	*Internal frequency database*
17	**Frequency database validation**	Internal frequencies compared against public references to detect systematic bias	gnomAD comparison	*Validation record*
18	**Cohort release packaging**	Final VCF, QC matrix, frequency database and documentation packaged with versions	Release procedure	*Cohort release package*
19	**Cohort readiness sign-off**	Analysable sample set and residual technical structure declared for downstream analysis	Sign-off procedure	*Cohort readiness statement*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
19	**92 h · 16 d**	**SCD-11**	**2,15,000**

SCD-12
SECTOR 01

Rapid Genome Analysis

(**Rapid whole genome sequencing analysis** · rWGS analysis)

WHY IT IS DONE

In acutely unwell patients a diagnosis that arrives in weeks arrives too late to change management. Rapid genome analysis compresses the same work into days by pre-provisioning compute, capturing phenotype before data arrive, and interpreting concurrently with sequencing rather than after it.

WHAT IT ANSWERS

Is there a genetic diagnosis that should change management for this acutely unwell patient, and can it be established within the clinical window?

WHEN IT IS DONE

In acute paediatric and neonatal presentations, intensive care admissions with suspected genetic cause, and any situation where the genetic result could alter treatment within days.

WHAT TRIGGERS IT

An acutely unwell patient with a presentation suggesting a genetic cause, escalation from an intensive care or acute paediatric team, and acceptance onto a validated rapid pathway.

HOW IT IS DONE

The pathway is activated with escalation contacts and a committed turnaround, phenotype is captured before data arrive, compute is reserved rather than queued, and quality is assessed during sequencing so a failing run is aborted early. Accelerated alignment and calling run against a demonstrated concordance to the standard pipeline, and actionable findings are communicated verbally the moment they are defensible.

WHAT IT PRODUCES

A provisional verbal communication record, the signed final report, the accelerated pipeline concordance statement, acute gene set coverage, and a stage-by-stage turnaround audit against the commitment.

WHAT IT DETECTS

The same variant classes as standard genome analysis — short variants, copy number, structural, repeat and mitochondrial — prioritised so that acutely actionable findings surface first.

WHAT IT DOES NOT DETECT

The same limitations as standard genome analysis apply, and compression adds its own: broader genome-wide review is necessarily curtailed, non-coding assessment is limited, and a rapid negative is a weaker negative than an unhurried one. That distinction must reach the clinician rather than being lost in an urgent handover.

WHAT MAKES IT HARD

The analysis is not the difficulty; the operation is. A queued compute job, an absent reviewer or a delayed confirmation destroys the turnaround. Speed must also be shown not to have cost accuracy, which requires a documented concordance to the standard assay rather than an assurance.

WHO DOES THIS JOB

Senior clinical bioinformatician on a rapid rota, working alongside on-call clinical genetics. Requires both analytical depth and operational discipline under time pressure.

WHAT YOU CAN DO AFTERWARDS

Design and operate a rapid analytical pathway, demonstrate concordance of an accelerated pipeline, manage verbal provisional reporting safely, and audit turnaround performance.

WHO HIRES FOR IT

Tertiary paediatric hospitals, neonatal and paediatric intensive care genomics services, national rapid genome programmes, and reference laboratories offering acute pathways.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
23	**114 hours · 19 days**	Prior genome analysis experience	**2,70,000**

SCD-12
WORKFLOW

Rapid Genome Analysis

(**Rapid whole genome sequencing analysis** · rWGS analysis)

COMPLETE WORKFLOW — 23 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Rapid pathway activation**	Case accepted onto the rapid pathway with clock started and turnaround committed	Escalation protocol	*Activated rapid case*
02	**Escalation contact confirmation**	On-call analytical, clinical and confirmation contacts confirmed available	On-call roster	*Escalation record*
03	**Pre-arrival phenotype capture**	Structured phenotype captured with the referring team before data arrive	HPO structured capture	*Structured phenotype*
04	**Compute pre-provisioning**	Dedicated capacity reserved with no queueing permitted on the rapid pathway	Reserved cluster or cloud capacity	*Provisioned compute*
05	**Streaming quality assessment**	Quality assessed during sequencing so a failing run is aborted rather than completed	Real-time run metrics	*Early acceptance decision*
06	**Accelerated alignment**	Reads aligned using a validated accelerated implementation with concordance demonstrated	GPU-accelerated aligner	*Aligned BAM*
07	**Accelerated variant calling**	Short variants called using the validated accelerated caller configuration	Accelerated caller	*Rapid VCF*
08	**Concurrent identity verification**	Identity and contamination verified in parallel rather than sequentially	somalier, VerifyBamID2	*Identity clearance*
09	**Acute gene coverage assessment**	Coverage sufficiency assessed across curated acute-presentation gene sets	mosdepth on acute gene sets	*Acute coverage report*
10	**Concurrent copy number calling**	Copy number called in parallel using a pre-built reference pool	gCNV with maintained pool	*Rapid CNV calls*
11	**Concurrent structural variant calling**	Structural variants called in parallel rather than deferred to a later pass	Manta	*Rapid SV calls*
12	**Repeat and mitochondrial screening**	Pathogenic repeat loci and mitochondrial variants screened concurrently	ExpansionHunter, Mutect2	*Repeat and mtDNA results*
13	**Annotation and acute-set filtering**	Annotation applied with acute-presentation gene sets examined first	VEP, gnomAD, ClinVar	*Annotated candidates*
14	**Phenotype-driven rapid prioritisation**	Candidates ranked against the captured phenotype with acute sets weighted first	Exomiser, HPO terms	*Priority candidates*
15	**Genome-wide expansion**	Analysis widened beyond acute sets while priority review proceeds in parallel	Genome-wide filtering	*Expanded candidate set*
16	**Rapid multidisciplinary review**	Candidates reviewed with the clinical team on a committed timeline	Scheduled review	*Reviewed candidates*
17	**ACMG-AMP classification**	Evidence codes applied to candidates under review with reasoning documented	ACMG-AMP with gene specifications	*Classified variants*
18	**Provisional verbal communication**	Actionable findings communicated verbally the moment they become defensible	Verbal report protocol	*Provisional communication record*
19	**Accelerated orthogonal confirmation**	Confirmation initiated on an accelerated pathway without routine batching	Rapid Sanger or ddPCR	*Confirmed findings*
20	**Concordance evidence attachment**	Accelerated pipeline concordance to the standard assay attached to the case file	Validation record	*Concordance statement*
21	**Final report generation**	Full written report assembled with limitations from compressed analysis stated	Validated rapid template	*Draft report*
22	**Clinical sign-out**	Report authorised under dual signature superseding the provisional communication	Sign-out procedure	*Signed clinical report*
23	**Turnaround performance audit**	Every stage timestamped against the committed turnaround with delays attributed	Timestamped stage log	*Turnaround audit*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
23	**114 h · 19 d**	**SCD-12**	**2,70,000**

SCD-13
SECTOR 01

Ultra-Rapid Genome Analysis

(**Ultra-rapid whole genome sequencing analysis** · urWGS analysis · 24-hour genome)

WHY IT IS DONE

Where management decisions must be taken within a single day, the entire diagnostic path has to be compressed below twenty-four hours. At that compression the work stops being a pipeline and becomes an orchestrated operation with a standing team and reserved infrastructure.

WHAT IT ANSWERS

Can a genetic diagnosis be established and communicated within the same day the sample was taken, in time to change an imminent clinical decision?

WHEN IT IS DONE

In the most acute intensive care presentations where a decision on escalation, withdrawal or specific therapy is imminent, and where a result the following week would have no effect on the outcome.

WHAT TRIGGERS IT

An imminent management decision in a critically unwell patient, activation of the ultra-rapid pathway by a joint clinical and laboratory decision, and confirmation that the standing team is available.

HOW IT IS DONE

The team is placed on standby and infrastructure reserved before the sample arrives. Basecalling, alignment and calling overlap with sequencing, interim call sets are issued at defined coverage checkpoints, and hypothesis genes are examined at first sufficient depth. Review is continuous rather than scheduled, and confirmation is initiated in parallel rather than after interpretation.

WHAT IT PRODUCES

Checkpoint coverage and interim call records, the first-pass hypothesis result, the progressive expansion record, every provisional communication with timing, the signed final report, and a full operational timing audit.

WHAT IT DETECTS

The same variant classes as genome analysis, examined progressively — pre-specified hypothesis genes first at the earliest sufficient coverage, then acute panels, then genome-wide as data accumulate.

WHAT IT DOES NOT DETECT

Everything standard genome analysis misses, plus the consequences of interpreting on partial data. Coverage at the point of first-pass interpretation is incomplete, structural and non-coding review is curtailed, and a negative issued at twenty-four hours is provisional rather than definitive.

WHAT MAKES IT HARD

Handoffs, not compute, are what fail. A hypothesis-directed first pass also risks stopping once a plausible candidate appears, so the progressive expansion record exists specifically to prove the analysis was completed systematically rather than abandoned early.

WHO DOES THIS JOB

Senior or lead clinical bioinformatician on a dedicated rota, in a service with an established rapid programme. Not an entry point into clinical genomics.

WHAT YOU CAN DO AFTERWARDS

Operate a sub-day genomic pathway, interpret safely on partial data, run progressive analytical expansion, and audit an operation where handoffs determine success.

WHO HIRES FOR IT

A small number of tertiary centres and national programmes operating ultra-rapid services. Scarce roles, high visibility, and strong signalling value for a career in acute genomics.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
24	**122 hours · 21 days**	Prior rapid genome analysis experience	**2,95,000**

SCD-13
WORKFLOW

Ultra-Rapid Genome Analysis

(**Ultra-rapid whole genome sequencing analysis** · urWGS analysis · 24-hour genome)

COMPLETE WORKFLOW — 24 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Ultra-rapid activation and standby**	Clinical, laboratory and analytical staff placed on standby before the sample arrives	Dedicated on-call rota	*Standing team*
02	**Joint clinical framing**	Clinical question framed as the imminent management decision requiring genomic input	Joint clinical protocol	*Framed question*
03	**Hypothesis capture**	Differential diagnoses and named candidate genes recorded before data are available	Structured hypothesis capture	*Recorded hypotheses*
04	**Dedicated infrastructure reservation**	GPU and storage capacity held exclusively with no shared queueing permitted	Reserved infrastructure	*Provisioned infrastructure*
05	**Streaming basecalling**	Basecalling run in real time concurrently with the sequencing run	Real-time basecaller	*Progressive basecalls*
06	**Progressive alignment**	Alignment proceeds on partial data as sequencing continues rather than after completion	Accelerated aligner	*Progressive alignment*
07	**Checkpoint interim calling**	Interim call sets issued at defined partial-coverage milestones	Checkpoint calling	*Interim call sets*
08	**Earliest identity verification**	Identity confirmed at the first point sufficient depth is available	somalier	*Identity clearance*
09	**Hypothesis gene first pass**	Pre-specified candidate genes examined at first sufficient coverage	Targeted interval analysis	*First-pass result*
10	**Acute panel expansion**	Analysis widened from hypothesis genes to curated acute presentation panels	Acute gene sets	*Acute panel result*
11	**Genome-wide expansion**	Analysis widened to the full genome as coverage accumulates	Genome-wide filtering	*Expanded candidate set*
12	**Concurrent structural and copy number analysis**	Structural and copy number calling run in parallel throughout	Manta, gCNV	*SV and CNV candidates*
13	**Repeat and mitochondrial screening**	Repeat loci and mitochondrial variants screened concurrently	ExpansionHunter, Mutect2	*Repeat and mtDNA results*
14	**Annotation and prioritisation**	Annotation and phenotype scoring applied continuously to emerging candidates	VEP, Exomiser	*Ranked candidates*
15	**Rolling clinical review**	Candidates assessed continuously by the standing team as they emerge	Continuous review protocol	*Rolling review record*
16	**ACMG-AMP classification**	Evidence codes applied with reasoning documented under time pressure	ACMG-AMP with gene specifications	*Classified variants*
17	**Immediate provisional communication**	Actionable findings communicated verbally at the moment they become defensible	Verbal report protocol	*Communication record*
18	**Parallel orthogonal confirmation**	Confirmation initiated before interpretation completes to protect the clock	Rapid Sanger or ddPCR	*Confirmed findings*
19	**Progressive expansion documentation**	Record demonstrating the analysis was completed systematically rather than stopped early	Expansion log	*Expansion record*
20	**Concordance evidence attachment**	Accelerated pipeline concordance to the standard assay attached to the case file	Validation record	*Concordance statement*
21	**Final report generation**	Written report assembled stating what compression cost in analytical breadth	Validated ultra-rapid template	*Draft report*
22	**Clinical sign-out**	Report authorised under dual signature superseding all provisional communications	Sign-out procedure	*Signed clinical report*
23	**Operational timing audit**	Every stage and handoff timestamped with delays attributed and remediation identified	Operational log	*Timing audit*
24	**Pathway review and remediation**	Audit reviewed jointly to correct the specific handoff that limited turnaround	Pathway review meeting	*Remediation actions*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
24	**122 h · 21 d**	**SCD-13**	**2,95,000**

SCD-14
SECTOR 01

Rapid Exome Analysis

(**Rapid whole exome sequencing analysis** · rWES analysis)

WHY IT IS DONE

Rapid exome analysis delivers most of the acute diagnostic yield of a rapid genome at lower cost and lower compute burden, which makes it the realistic rapid pathway for the majority of laboratories that cannot resource a genome service.

WHAT IT ANSWERS

Is there a coding-region genetic diagnosis that should change management for this acutely unwell patient, within the clinical window?

WHEN IT IS DONE

In acute presentations where a rapid answer is required and genome capacity is unavailable or unjustified, and as the first rapid pathway a laboratory establishes before progressing to genome.

WHAT TRIGGERS IT

An acutely unwell patient with a suspected genetic cause, escalation from an acute clinical team, and acceptance onto the validated rapid exome pathway.

HOW IT IS DONE

The pathway is activated with a committed turnaround, phenotype is captured before data arrive, and compute is reserved. Quality is assessed during sequencing, accelerated alignment and calling run against a demonstrated concordance to the routine exome, copy number is called from a pre-built reference pool so it does not delay the pathway, and findings are communicated verbally when defensible.

WHAT IT PRODUCES

The provisional communication record, signed final report, accelerated concordance statement, acute gene coverage report, rapid copy number findings, the escalation recommendation, and the turnaround audit.

WHAT IT DETECTS

Coding and splice variants and exon-level copy number changes, prioritised so that acute-presentation genes are examined first.

WHAT IT DOES NOT DETECT

Everything exome analysis misses — deep intronic, regulatory, structural and repeat causes — compounded by compressed review. A rapid exome negative is not a genome negative, and the escalation recommendation exists specifically to ensure that distinction reaches the clinician.

WHAT MAKES IT HARD

Copy number analysis is the element most often dropped under time pressure, which removes a real share of acute diagnostic yield. Maintaining a current reference pool in advance is what makes it possible to include it without cost to turnaround.

WHO DOES THIS JOB

Clinical bioinformatician on a rapid rota, mid to senior level. The most accessible entry point into acute genomics work.

WHAT YOU CAN DO AFTERWARDS

Operate a rapid exome pathway including copy number, demonstrate accelerated pipeline concordance, manage provisional verbal reporting, and construct defensible escalation recommendations.

WHO HIRES FOR IT

Hospital genomics laboratories establishing acute services, paediatric hospitals, regional genomics networks, and diagnostic chains offering urgent testing.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
20	**98 hours · 17 days**	Prior exome analysis experience	**2,30,000**

SCD-14
WORKFLOW

Rapid Exome Analysis

(**Rapid whole exome sequencing analysis** · rWES analysis)

COMPLETE WORKFLOW — 20 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Rapid pathway activation**	Case accepted with clock started, turnaround committed and escalation contacts confirmed	Escalation protocol	*Activated rapid case*
02	**Pre-arrival phenotype capture**	Structured phenotype captured with the referring team before data arrive	HPO structured capture	*Structured phenotype*
03	**Compute pre-provisioning**	Dedicated capacity reserved with no queueing permitted	Reserved capacity	*Provisioned compute*
04	**Streaming quality assessment**	Quality assessed during sequencing with an accept or abort decision taken early	Real-time run metrics	*Early acceptance decision*
05	**Accelerated capture-aware alignment**	Reads aligned within the capture target using a validated accelerated implementation	Accelerated aligner	*Aligned BAM*
06	**Accelerated variant calling**	Short variants called across padded target intervals with the validated configuration	Accelerated caller	*Rapid exome VCF*
07	**Concurrent identity verification**	Identity and contamination verified in parallel to protect the clock	somalier, VerifyBamID2	*Identity clearance*
08	**Acute gene coverage assessment**	Coverage sufficiency assessed across curated acute-presentation genes	mosdepth on acute gene sets	*Acute coverage report*
09	**Rapid copy number calling**	Copy number called using a reference pool maintained in advance	DECoN or gCNV	*Rapid CNV calls*
10	**Annotation and acute-set filtering**	Annotation applied with acute gene sets examined ahead of broader analysis	VEP, gnomAD, ClinVar	*Annotated candidates*
11	**Phenotype-driven rapid prioritisation**	Candidates ranked against the captured phenotype with acute weighting	Exomiser, HPO terms	*Priority candidates*
12	**Exome-wide expansion**	Analysis widened beyond acute sets while priority review proceeds	Exome-wide filtering	*Expanded candidate set*
13	**Rapid multidisciplinary review**	Candidates reviewed with the clinical team on a committed timeline	Scheduled review	*Reviewed candidates*
14	**ACMG-AMP classification**	Evidence codes applied with reasoning documented for candidates under review	ACMG-AMP with gene specifications	*Classified variants*
15	**Provisional verbal communication**	Actionable findings communicated verbally when defensible with timing recorded	Verbal report protocol	*Communication record*
16	**Accelerated orthogonal confirmation**	Confirmation performed on an accelerated pathway without routine batching	Rapid Sanger or ddPCR	*Confirmed findings*
17	**Escalation recommendation**	Explicit genome escalation recommendation generated for every rapid negative	Documented escalation criteria	*Escalation recommendation*
18	**Concordance evidence attachment**	Accelerated pipeline concordance to the routine exome attached to the case file	Validation record	*Concordance statement*
19	**Final report generation and sign-out**	Report with coding-restriction limitations issued under dual signature	Validated rapid template	*Signed clinical report*
20	**Turnaround performance audit**	Stage-by-stage timestamps recorded against the committed turnaround	Timestamped stage log	*Turnaround audit*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
20	**98 h · 17 d**	**SCD-14**	**2,30,000**

SCD-15
SECTOR 01

Copy Number Variant Analysis

(**CNV analysis from sequencing data** · CNV analysis)

WHY IT IS DONE

Copy number changes account for a substantial share of monogenic diagnostic yield and are entirely invisible to sequence variant calling. Detecting them from the same data rather than by separate array or MLPA removes a second test, a second cost and a second delay from every diagnostic pathway.

WHAT IT ANSWERS

Are there deletions or duplications affecting clinically relevant genes, at what copy state, and what is their classification?

WHEN IT IS DONE

Alongside every sequence variant analysis as standard practice, and specifically where the phenotype suggests a dosage-sensitive mechanism or where sequence analysis has returned negative.

WHAT TRIGGERS IT

Routine inclusion in exome, panel and genome analysis, a negative sequence variant result, or a clinical suspicion of a known microdeletion or microduplication syndrome.

HOW IT IS DONE

A matched reference pool is constructed and governed, batch compatibility is verified, and coverage is normalised for GC and mappability bias before segmentation. Calls are quality-scored with stricter thresholds for single-exon events, non-informative regions are declared explicitly, and findings are classified under the ACMG-ClinGen copy number framework before orthogonal confirmation.

WHAT IT PRODUCES

An exon-level copy number call report with quality evidence, the reference pool governance record, a sensitivity and non-informative region statement, ACMG-ClinGen classification records, and orthogonal confirmation results.

WHAT IT DETECTS

Exon-level, gene-level and multi-gene deletions and duplications, with copy state assignment and quality evidence, at a resolution determined by the underlying assay and the reference pool.

WHAT IT DOES NOT DETECT

Balanced rearrangements, which carry normal copy number and are entirely invisible to this method. Also misses copy number changes in regions of poor mappability or capture, single-exon events below the validated sensitivity, and the precise breakpoints of any event, which require a separate analysis.

WHAT MAKES IT HARD

The method is only as good as its reference pool, and a stale or batch-mismatched pool produces both false positives and silent false negatives. Sensitivity differs sharply between single-exon and multi-exon events, so a single sensitivity claim is misleading and the non-informative regions must be declared rather than quietly omitted.

WHO DOES THIS JOB

Clinical bioinformatician or variant scientist, mid level. A distinct and consistently under-supplied skill relative to sequence variant analysis.

WHAT YOU CAN DO AFTERWARDS

Build and govern a reference pool, call and quality-score copy number reliably, state sensitivity honestly by event size, and classify under the ACMG-ClinGen framework.

WHO HIRES FOR IT

Every laboratory offering exome, panel or genome testing. Copy number competence is one of the most frequently cited gaps in clinical bioinformatics recruitment.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
18	**86 hours · 15 days**	Prior alignment and coverage analysis experience	**2,00,000**

SCD-15
WORKFLOW

Copy Number Variant Analysis

(**CNV analysis from sequencing data** · CNV analysis)

COMPLETE WORKFLOW — 18 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Reference pool construction**	Matched normals from the same capture kit, chemistry and batch assembled into a pool	Internal normal sample set	*Reference pool*
02	**Reference pool governance**	Pool size, kit version, construction date and refresh schedule documented	Pool register	*Pool governance record*
03	**Batch compatibility verification**	Test sample coverage profile correlated against the pool with drift flagged	Correlation analysis	*Compatibility record*
04	**Target interval preparation**	Target intervals prepared with consistent naming and padding across sample and pool	bedtools, Picard	*Interval definitions*
05	**Coverage matrix construction**	Per-target read depth extracted identically across the test sample and every pool member	mosdepth, bedtools coverage	*Coverage matrix*
06	**GC bias correction**	Coverage normalised across GC content bins to remove systematic capture bias	GC correction	*GC-corrected coverage*
07	**Mappability correction**	Coverage adjusted for mappability so unmappable regions are not called as deletions	Umap tracks	*Corrected coverage*
08	**Principal component denoising**	Systematic technical variation removed by denoising against the reference pool	PCA denoising	*Denoised coverage*
09	**Segmentation**	Adjacent targets with consistent copy ratio merged into candidate copy number segments	CBS or HMM segmentation	*Segments*
10	**Copy number state assignment**	Integer copy state assigned to each segment with quality scoring	gCNV, CNVkit, DECoN	*Raw CNV calls*
11	**Event-size-specific filtering**	Single-exon calls held to stricter thresholds than multi-exon events	Caller quality metrics	*Filtered CNV calls*
12	**Non-informative region declaration**	Targets where the method cannot call reliably identified and declared explicitly	Per-target informativeness	*Non-informative region list*
13	**Population frequency filtering**	Common copy number polymorphism removed using public and internal frequency data	gnomAD-SV, DGV, internal database	*Rare CNV set*
14	**Dosage sensitivity annotation**	Gene overlap and dosage sensitivity evidence annotated for each remaining call	ClinGen dosage sensitivity	*Annotated CNV set*
15	**Sensitivity determination**	Detectable event size established empirically by event type and genomic context	Reference material benchmarking	*Sensitivity statement*
16	**ACMG-ClinGen classification**	Points-based copy number classification applied with dosage and phenotype evidence	ACMG-ClinGen CNV framework	*Classified CNV set*
17	**Orthogonal confirmation**	Every reportable copy number finding confirmed by an independent method	MLPA, qPCR, ddPCR	*Confirmed CNVs*
18	**Report generation and sign-out**	Calls, sensitivity statement and non-informative regions issued under dual signature	Validated CNV report template	*Signed CNV report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
18	**86 h · 15 d**	**SCD-15**	**2,00,000**

SCD-16
SECTOR 01

Structural Variant Analysis

(**SV analysis from whole genome data** · SV analysis)

WHY IT IS DONE

Structural variants cause disease through gene disruption, dosage change and position effect, and they are missed by both sequence variant calling and array. In cohorts where exome and array have failed, structural analysis of genome data is one of the highest-yield remaining options.

WHAT IT ANSWERS

Are there deletions, duplications, inversions, insertions or translocations affecting clinically relevant genes or regulatory architecture, and where exactly are their breakpoints?

WHEN IT IS DONE

In genome analysis as standard, after negative exome and array testing, where the phenotype suggests a contiguous gene or position effect mechanism, or to characterise a rearrangement already seen on karyotype.

WHAT TRIGGERS IT

Routine inclusion in genome analysis, a negative exome and array workup, or an abnormal cytogenetic finding requiring molecular resolution.

HOW IT IS DONE

Multiple callers with complementary class strengths run in parallel, calls are merged and uniformly re-genotyped, and class-specific filters are applied. Every reportable call is reviewed at read level at both breakpoints, repeat context is flagged, and classification distinguishes dosage mechanisms from disruption mechanisms before class-appropriate confirmation.

WHAT IT PRODUCES

A merged and genotyped structural variant call set with caller provenance, per-class sensitivity statement, breakpoint evidence review records, gene disruption and fusion assessment, and classification records.

WHAT IT DETECTS

Deletions, duplications, insertions, inversions and translocations with breakpoint coordinates, gene disruption assessment, predicted fusions, and the parental origin of events where family data are available.

WHAT IT DOES NOT DETECT

Sensitivity varies by an order of magnitude across variant classes, so a single figure is meaningless. Events in repeat-rich and segmentally duplicated regions are unreliable, very large insertions exceeding read length are missed, and short-read data resolve inversions and balanced events far more poorly than deletions.

WHAT MAKES IT HARD

Short-read structural calling has the highest false positive rate in clinical genomics. No single caller is adequate, so ensemble calling is mandatory, and read-level review of every reportable breakpoint is the only reliable control. Repeat-driven artefacts dominate unfiltered call sets.

WHO DOES THIS JOB

Senior clinical bioinformatician. One of the most technically demanding roles in clinical genomics and among the hardest positions for laboratories to fill.

WHAT YOU CAN DO AFTERWARDS

Operate ensemble structural variant calling, review breakpoint evidence competently, state per-class sensitivity defensibly, and classify balanced and unbalanced events under the correct frameworks.

WHO HIRES FOR IT

Genome-capable hospital laboratories, national genomics programmes, cytogenetics services transitioning to sequencing, and reference laboratories offering structural analysis.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
21	**102 hours · 17 days**	Prior genome analysis and alignment experience	**2,40,000**

SCD-16
WORKFLOW

Structural Variant Analysis

(**SV analysis from whole genome data** · SV analysis)

COMPLETE WORKFLOW — 21 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Alignment preparation**	Coordinate-sorted, duplicate-marked alignments verified for read group and pairing integrity	samtools, Picard	*Analysis-ready BAM*
02	**Caller selection and version lock**	Callers selected for complementary variant class strengths with versions locked	Validation record	*Caller configuration*
03	**Ensemble structural calling**	Multiple callers run in parallel across the full genome	Manta, Delly, GRIDSS	*Per-caller SV sets*
04	**Call merging and harmonisation**	Calls merged with defined breakpoint tolerance and caller agreement recorded	SURVIVOR or Jasmine	*Merged SV set*
05	**Uniform re-genotyping**	Every merged call re-genotyped uniformly so calls are comparable across callers	SVTyper, graph genotyping	*Genotyped SV set*
06	**Class-specific quality filtering**	Deletions, duplications, inversions and translocations filtered on separate criteria	Per-class thresholds	*Filtered SV set*
07	**Repeat and blacklist annotation**	Calls in repeats, segmental duplications and blacklisted regions flagged by confidence tier	RepeatMasker, ENCODE blacklist	*Context-annotated calls*
08	**Copy number corroboration**	Depth evidence assessed at each breakpoint to distinguish balanced from unbalanced events	mosdepth, read depth analysis	*Balance assessment*
09	**Population frequency filtering**	Common structural polymorphism removed using public and internal frequency data	gnomAD-SV, DGV, internal database	*Rare SV set*
10	**Breakpoint read-level review**	Both breakpoints of every reportable call inspected at read level	IGV review	*Reviewed calls*
11	**Junction assembly**	Local assembly performed at breakpoints to reconstruct junction sequence	Local assembly	*Junction sequences*
12	**Gene disruption assessment**	Genes interrupted at each breakpoint identified with exon and transcript consequence	Gene and transcript overlap	*Disruption assessment*
13	**Regulatory and domain assessment**	Enhancer separation and topological domain disruption assessed for position effect	Regulatory and TAD annotation	*Regulatory assessment*
14	**Fusion prediction**	Predicted gene fusions assessed for reading frame and product plausibility	Frame and orientation analysis	*Fusion predictions*
15	**Per-class sensitivity determination**	Sensitivity established separately for each variant class by benchmarking	Reference material benchmarking	*Per-class sensitivity*
16	**ACMG-ClinGen classification**	Dosage evidence applied to unbalanced events and disruption evidence to balanced events	ACMG-ClinGen CNV framework	*Classified SV set*
17	**Class-appropriate confirmation**	Confirmation method matched to variant class including breakpoint-spanning PCR	PCR, MLPA, karyotype, long-read	*Confirmed SVs*
18	**Report generation**	Calls, per-class sensitivity and breakpoint evidence assembled into the report	Validated SV report template	*Draft report*
19	**Clinical sign-out**	Report reviewed and authorised under dual signature	Sign-out procedure	*Signed SV report*
20	**Internal artefact database update**	Recurrent false positives added to the internal artefact database for future filtering	Internal artefact register	*Updated artefact database*
21	**Validation evidence maintenance**	Per-class sensitivity claims refreshed against current reference materials	Benchmarking cycle	*Maintained validation record*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
21	**102 h · 17 d**	**SCD-16**	**2,40,000**

SCD-17
SECTOR 01

Digital Karyotyping

(**Chromosomal microarray-equivalent analysis** · CMA-equivalent analysis)

WHY IT IS DONE

Sequencing-based copy number analysis now exceeds chromosomal microarray in resolution and is not constrained by probe density, which allows laboratories to consolidate two services into one. Replacing an established array service requires demonstrated equivalence rather than an assertion.

WHAT IT ANSWERS

Does sequencing-based analysis detect everything the array would have detected, at equal or better resolution, and can it therefore replace the array service?

WHEN IT IS DONE

When a laboratory is consolidating array and sequencing services, validating a sequencing replacement for an existing array menu, or offering array-equivalent reporting to clinicians who expect it.

WHAT TRIGGERS IT

A service consolidation decision, a business case to retire array infrastructure, or a clinical governance requirement to demonstrate equivalence before withdrawing an established test.

HOW IT IS DONE

The incumbent array's resolution and reporting thresholds are documented as the benchmark, sequencing depth and bin size are designed to meet or exceed them, and calling includes mosaic-aware analysis and absence of heterozygosity. Retrospective concordance is established across a defined case series with every discordance investigated, and findings are reported in ISCN-style nomenclature.

WHAT IT PRODUCES

The equivalence specification, retrospective concordance validation report with discordance analysis, resolution comparison against array probe spacing, genome-wide copy number findings, absence of heterozygosity results, and ISCN-style descriptions.

WHAT IT DETECTS

Genome-wide copy number changes at resolution matching or exceeding array, mosaic aneuploidy, absence of heterozygosity, and the same reportable categories clinicians currently receive from cytogenetics.

WHAT IT DOES NOT DETECT

Balanced rearrangements, exactly as array cannot detect them. Also constrained by mappability rather than probe density, so the blind spots differ from the array's blind spots — which is precisely why discordance analysis during validation matters more than the headline concordance figure.

WHAT MAKES IT HARD

Equivalence is a validation exercise, not an analytical one. The discordances are more informative than the concordances, and clinical acceptance depends on reporting in nomenclature that cytogeneticists and clinicians can read without retraining.

WHO DOES THIS JOB

Senior clinical bioinformatician or clinical scientist, working jointly with cytogenetics. A bridging role between two disciplines and correspondingly scarce.

WHAT YOU CAN DO AFTERWARDS

Design and execute a service equivalence validation, investigate discordance rigorously, and produce sequencing reports that integrate cleanly with existing cytogenetic practice.

WHO HIRES FOR IT

Hospital laboratories consolidating cytogenetics and molecular services, national programmes retiring array platforms, and diagnostic chains rationalising test menus.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
17	**82 hours · 14 days**	Prior copy number analysis experience	**1,90,000**

SCD-17
WORKFLOW

Digital Karyotyping

(**Chromosomal microarray-equivalent analysis** · CMA-equivalent analysis)

COMPLETE WORKFLOW — 17 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Equivalence scope definition**	The array platform being replaced documented with probe density and reporting thresholds	Incumbent platform specification	*Equivalence specification*
02	**Sequencing design specification**	Depth and bin size designed to meet or exceed the array's effective resolution	Design calculation	*Design specification*
03	**Validation case series selection**	A defined series of cases with known array results assembled for concordance testing	Case selection criteria	*Validation series*
04	**Coverage extraction**	Binned read depth extracted genome-wide at the designed bin size	mosdepth	*Binned coverage*
05	**Bias correction**	Coverage corrected for GC content and mappability before segmentation	GC and mappability correction	*Corrected coverage*
06	**Blacklist masking**	Centromeric, telomeric and blacklisted regions masked and declared rather than called	ENCODE blacklist	*Masked coverage*
07	**Segmentation and state assignment**	Segments identified and integer copy states assigned with ploidy awareness	CBS or HMM, gCNV	*CNV call set*
08	**Mosaic aneuploidy analysis**	Low-level mosaic aneuploidy detected through allele fraction and depth modelling	Mosaic-aware calling	*Mosaic call set*
09	**Absence of heterozygosity detection**	Homozygous regions detected from SNP genotypes to match the array's capability	ROH detection	*AOH regions*
10	**Retrospective concordance analysis**	Sensitivity, specificity and concordance computed against known array results	Concordance analysis	*Concordance metrics*
11	**Discordance investigation**	Every discordant case investigated and the cause of difference established	Case-by-case review	*Discordance analysis*
12	**Resolution comparison**	Effective resolution compared directly against array probe spacing by genomic region	Benchmarking analysis	*Resolution comparison*
13	**Population frequency filtering**	Common copy number polymorphism removed in line with established array conventions	gnomAD-SV, DGV, ISCA sets	*Rare CNV set*
14	**ACMG-ClinGen classification**	Points-based classification applied consistently with established array reporting practice	ACMG-ClinGen CNV framework	*Classified CNV set*
15	**ISCN-style description generation**	Findings expressed in nomenclature clinicians already read from cytogenetics	ISCN conventions	*ISCN descriptions*
16	**Report generation**	Findings, resolution statement and AOH results assembled in array-continuous format	Validated report template	*Draft report*
17	**Clinical sign-out**	Report reviewed and authorised under dual signature	Sign-out procedure	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
17	**82 h · 14 d**	**SCD-17**	**1,90,000**

SCD-18
SECTOR 01

Runs of Homozygosity Analysis

(**Homozygosity mapping and AOH analysis** · ROH · AOH · homozygosity mapping)

WHY IT IS DONE

Homozygous regions localise recessive disease genes in consanguineous families, flag uniparental disomy, and quantify inbreeding — all from data already generated, at negligible additional cost. In populations with high consanguinity it is one of the highest-yield analyses available.

WHAT IT ANSWERS

Which genomic regions are homozygous, what does their distribution say about parental relatedness, and which recessive disease genes fall within them?

WHEN IT IS DONE

Routinely alongside exome and genome analysis, and specifically where consanguinity is known or suspected, where a recessive mechanism is likely, or where uniparental disomy needs to be excluded.

WHAT TRIGGERS IT

Known or suspected consanguinity, a recessive presentation, an unexplained homozygous variant, or a phenotype consistent with an imprinting disorder.

HOW IT IS DONE

Reporting thresholds and the disclosure pathway are confirmed before analysis begins. Genotypes are filtered and annotated with ancestry-matched frequencies, homozygous runs are detected and filtered on length and density, inbreeding is quantified, and the distribution is classified against expected consanguinity patterns. Chromosome-restricted homozygosity is separated from genome-wide patterns.

WHAT IT PRODUCES

The runs of homozygosity map with genes contained, the inbreeding coefficient with its relationship band, consanguinity pattern classification, uniparental disomy discrimination, a candidate recessive gene shortlist, and the disclosure handling record.

WHAT IT DETECTS

Runs of homozygosity with coordinates and length, the genomic inbreeding coefficient, consanguinity patterns consistent with defined degrees of relationship, chromosome-restricted homozygosity indicating uniparental disomy, and recessive candidate genes within homozygous regions.

WHAT IT DOES NOT DETECT

It localises rather than diagnoses — a homozygous region contains many genes and the analysis does not identify which one is responsible. Heterodisomic uniparental disomy produces no homozygosity and is missed entirely. Short regions arising from linkage disequilibrium are not clinically meaningful and must be excluded.

WHAT MAKES IT HARD

The analysis is technically simple and ethically complex. It reveals consanguinity that families have not disclosed and occasionally patterns that raise safeguarding concerns, so the disclosure policy must be agreed before analysis rather than negotiated after a finding appears.

WHO DOES THIS JOB

Variant scientist or clinical bioinformatician, entry to mid level, with clinical genetics and safeguarding governance available for disclosure decisions.

WHAT YOU CAN DO AFTERWARDS

Detect and filter homozygous regions correctly, quantify inbreeding, distinguish uniparental disomy from consanguinity, and apply a disclosure policy to sensitive incidental findings.

WHO HIRES FOR IT

Laboratories serving populations with high consanguinity, rare disease services, hospital genomics laboratories, and clinical genetics units across South Asia and the Middle East particularly.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
12	**56 hours · 10 days**	Familiarity with VCF and genotype data	**1,25,000**

SCD-18
WORKFLOW

Runs of Homozygosity Analysis

(**Homozygosity mapping and AOH analysis** · ROH · AOH · homozygosity mapping)

COMPLETE WORKFLOW — 12 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Reporting and disclosure policy confirmation**	Reporting thresholds and the consanguinity disclosure pathway confirmed before analysis	Laboratory policy	*Policy confirmation*
02	**Genotype extraction**	Genotypes extracted at informative common SNP sites with depth thresholds applied	bcftools, genotype extraction	*Genotype set*
03	**Genotype quality filtering**	Sites with inadequate depth, call rate or quality removed before detection	Quality filters	*Filtered genotypes*
04	**Ancestry-matched frequency annotation**	Population allele frequencies applied from an ancestry-appropriate reference	gnomAD by ancestry	*Frequency-annotated genotypes*
05	**Homozygous run detection**	Extended homozygous regions detected with validated length and density parameters	HMM or sliding window detection	*Raw ROH regions*
06	**Linkage disequilibrium filtering**	Short regions arising from linkage disequilibrium excluded from clinical consideration	Length and density thresholds	*Filtered ROH regions*
07	**Inbreeding coefficient calculation**	Genomic inbreeding coefficient computed from total homozygous burden	Total ROH burden analysis	*Inbreeding coefficient*
08	**Consanguinity pattern classification**	Region size distribution classified against patterns expected for defined relationships	Distribution analysis	*Consanguinity classification*
09	**Uniparental disomy discrimination**	Chromosome-restricted homozygosity separated from genome-wide consanguinity	Per-chromosome distribution analysis	*UPD flag*
10	**Candidate gene localisation**	Recessive disease genes within homozygous regions identified and ranked against phenotype	Gene overlap, HPO scoring	*Candidate gene shortlist*
11	**Homozygous variant assessment**	Rare homozygous variants within the regions classified and correlated with phenotype	VEP, ACMG-AMP	*Classified candidates*
12	**Disclosure handling and reporting**	Consanguinity findings handled under the agreed pathway and the report issued	Documented policy, report template	*Signed report and disclosure record*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
12	**56 h · 10 d**	**SCD-18**	**1,25,000**

SCD-19
SECTOR 01

Uniparental Disomy Analysis

(**UPD detection and imprinting analysis** · UPD analysis)

WHY IT IS DONE

Uniparental disomy causes disease by disrupting imprinting or by unmasking a recessive variant, and it is entirely invisible to copy number analysis because the copy number is normal. It is a genuine diagnosis that is routinely missed because nobody looks for it.

WHAT IT ANSWERS

Were both copies of a chromosome or chromosomal segment inherited from a single parent, from which parent, and does that explain the presentation through imprinting or recessive unmasking?

WHEN IT IS DONE

Where the phenotype suggests an imprinting disorder, where a homozygous recessive variant appears with only one carrier parent, where growth restriction or developmental delay is unexplained, and as a standard check within family-based analysis.

WHAT TRIGGERS IT

A phenotype consistent with a known imprinting syndrome, an unexpected homozygous variant, chromosome-restricted homozygosity found during other analysis, or a discordant inheritance pattern in trio data.

HOW IT IS DONE

The available sample configuration is assessed because sensitivity differs sharply with and without parents. Mendelian inconsistency patterns are analysed where parents are available, homozygosity blocks are detected and distinguished from consanguinity, mechanism and extent are classified, parental origin is determined, and imprinted locus overlap and recessive unmasking are both assessed.

WHAT IT PRODUCES

The sample configuration and sensitivity record, Mendelian inconsistency report, homozygosity and consanguinity discrimination, mechanism and extent classification, parental origin report, imprinted locus assessment, and recessive unmasking findings.

WHAT IT DETECTS

Whole-chromosome and segmental uniparental disomy, the distinction between isodisomy and heterodisomy, parental origin where family samples are available, overlap with imprinted loci, and recessive variants unmasked by isodisomy.

WHAT IT DOES NOT DETECT

Heterodisomy is essentially undetectable without parental samples because it produces no homozygosity signature. The analysis also cannot detect imprinting defects arising from epimutation rather than disomy, which require methylation analysis and are a common alternative cause of the same clinical syndromes.

WHAT MAKES IT HARD

Distinguishing chromosome-restricted homozygosity from consanguinity is the crux, and the two have entirely different clinical consequences. Parental origin determines which syndrome results, so the same chromosome gives different disease depending on which parent contributed it.

WHO DOES THIS JOB

Variant scientist or clinical bioinformatician, mid level, working with clinical genetics on imprinting syndrome correlation.

WHAT YOU CAN DO AFTERWARDS

Detect and classify uniparental disomy with and without parental samples, determine parental origin, and link findings to imprinting syndromes and recessive unmasking.

WHO HIRES FOR IT

Clinical genetics services, paediatric genomics laboratories, imprinting disorder specialist centres, and hospital laboratories offering family-based testing.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
13	**64 hours · 11 days**	Familiarity with genotype data and inheritance logic	**1,45,000**

SCD-19
WORKFLOW

Uniparental Disomy Analysis

(**UPD detection and imprinting analysis** · UPD analysis)

COMPLETE WORKFLOW — 13 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Sample configuration assessment**	Trio, duo or proband-only configuration determined with its sensitivity limits recorded	Sample inventory review	*Configuration record*
02	**Genotype extraction**	High-quality genotypes extracted at informative sites across all chromosomes	bcftools, genotype extraction	*Genotype set*
03	**Relatedness verification**	Declared parentage verified before any inheritance-based inference is trusted	somalier relate, KING	*Relationship verification*
04	**Mendelian inconsistency analysis**	Patterns of Mendelian error analysed by chromosome where parental data exist	Trio genotype logic	*Inconsistency patterns*
05	**Homozygosity block detection**	Extended homozygous regions detected as candidate isodisomic segments	ROH detection	*Homozygosity blocks*
06	**Consanguinity discrimination**	Chromosome-restricted homozygosity distinguished from a genome-wide consanguinity pattern	Genome-wide distribution analysis	*Consanguinity assessment*
07	**Mechanism classification**	Isodisomy, heterodisomy and mixed patterns distinguished with mechanistic implications	Genotype pattern analysis	*Mechanism classification*
08	**Extent determination**	Segmental events delineated with boundaries and whole-chromosome events confirmed	Boundary determination	*Event extent*
09	**Parental origin determination**	Maternal or paternal origin established, determining which imprinting consequence follows	Trio genotype analysis	*Parental origin*
10	**Imprinted locus assessment**	Overlap with catalogued imprinted regions assessed and the syndrome implicated identified	Imprinted region catalogue	*Imprinting assessment*
11	**Recessive unmasking assessment**	Homozygous recessive variants within the isodisomic region identified and classified	Variant analysis, ACMG-AMP	*Unmasked variant assessment*
12	**Confirmation testing**	Findings confirmed by an independent method appropriate to the mechanism identified	Methylation testing, microsatellites	*Confirmed findings*
13	**Report generation and sign-out**	Mechanism, extent, origin, clinical implication and recurrence risk issued under dual signature	Validated UPD report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
13	**64 h · 11 d**	**SCD-19**	**1,45,000**

SCD-20
SECTOR 01

Repeat Expansion Analysis

(**Short tandem repeat expansion detection** · RED · repeat analysis)

WHY IT IS DONE

Repeat expansions cause a substantial group of neurological and neuromuscular diseases and are invisible to standard variant calling, which cannot represent them. Detecting them from existing sequencing data replaces a panel of separate single-gene tests with one analysis.

WHAT IT ANSWERS

Is any pathogenic repeat locus expanded beyond its normal range, what is the approximate size, and does it fall in the premutation or full mutation range?

WHEN IT IS DONE

Alongside genome and exome analysis where the phenotype includes ataxia, myopathy, dystonia, cognitive decline or a movement disorder, and as a replacement for individual repeat-specific tests in laboratories consolidating their menus.

WHAT TRIGGERS IT

A neurological or neuromuscular phenotype consistent with repeat expansion disease, a family history suggesting anticipation, or a negative standard variant analysis in a compatible presentation.

HOW IT IS DONE

A catalogue of pathogenic loci is applied with a validated genotyping method, repeat sizes are estimated with confidence bounds, and calls are assessed against locus-specific normal, premutation and pathogenic thresholds. Sizing limitations are stated explicitly and any expanded finding is routed to an orthogonal sizing method for confirmation.

WHAT IT PRODUCES

A repeat genotype report per locus with size estimates and confidence bounds, threshold assessment against locus-specific ranges, an explicit sizing limitation statement, and orthogonal confirmation results for expanded findings.

WHAT IT DETECTS

Expansions at catalogued pathogenic repeat loci with approximate sizing, genotype relative to normal, premutation and pathogenic thresholds, and novel repeat motifs where a discovery-mode method is applied.

WHAT IT DOES NOT DETECT

Expansions substantially longer than the sequencing read length cannot be sized accurately and are typically reported only as expanded. Somatic instability and interruption patterns within the repeat are not resolved by short reads, and any locus absent from the catalogue is not examined at all.

WHAT MAKES IT HARD

Short reads cannot size large expansions, and reporting an estimate as if it were a measurement is the characteristic error. The correct output for a large expansion is a categorical expanded call with an explicit statement that accurate sizing requires an orthogonal method.

WHO DOES THIS JOB

Clinical bioinformatician, mid level, typically in a laboratory serving neurology or neuromuscular services.

WHAT YOU CAN DO AFTERWARDS

Genotype repeat loci from short-read data, apply locus-specific thresholds correctly, state sizing limitations honestly, and route findings to appropriate confirmatory sizing.

WHO HIRES FOR IT

Neurogenetics laboratories, hospital genomics services, neuromuscular specialist centres, and laboratories consolidating single-gene repeat tests into sequencing pathways.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
15	**72 hours · 12 days**	Prior alignment analysis experience	**1,65,000**

SCD-20
WORKFLOW

Repeat Expansion Analysis

(**Short tandem repeat expansion detection** · RED · repeat analysis)

COMPLETE WORKFLOW — 15 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Locus catalogue authorisation**	The pathogenic repeat locus catalogue and its version confirmed and recorded	Repeat locus register	*Authorised catalogue version*
02	**Alignment verification**	Alignments verified as suitable for repeat genotyping including read length and depth	samtools stats	*Alignment suitability record*
03	**Repeat region read extraction**	Reads spanning and flanking each catalogued repeat locus extracted	Targeted extraction	*Locus read sets*
04	**Repeat genotyping**	Repeat unit counts genotyped at each catalogued locus with the validated method	ExpansionHunter, GangSTR	*Repeat genotypes*
05	**Confidence interval estimation**	Size estimates produced with confidence bounds reflecting read length constraints	Genotyper confidence intervals	*Sized genotypes*
06	**Read-level evidence review**	Spanning, flanking and in-repeat read evidence inspected for each expanded call	IGV, read visualisation	*Reviewed calls*
07	**Locus-specific threshold assessment**	Genotypes assessed against locus-specific normal, premutation and pathogenic ranges	Curated threshold tables	*Threshold assessment*
08	**Motif and interruption assessment**	Repeat motif composition assessed where the method and data support it	Motif analysis	*Motif assessment*
09	**Novel expansion screening**	Genome-wide screening for expansions outside the catalogue where indicated	ExpansionHunter Denovo, STRling	*Novel expansion candidates*
10	**Sizing limitation documentation**	Explicit statement of where read length prevents accurate sizing of large expansions	Limitation template	*Sizing limitation statement*
11	**Population comparison**	Genotypes compared against population repeat length distributions for context	Population repeat references	*Population context*
12	**Phenotype correlation**	Findings correlated against the clinical presentation and family history	Clinical correlation review	*Correlation record*
13	**Orthogonal sizing confirmation**	Expanded findings routed to an orthogonal method capable of accurate sizing	Repeat-primed PCR, Southern blot, long-read	*Confirmed sizing*
14	**Report generation**	Genotypes, thresholds, sizing limitations and confirmation assembled into the report	Validated report template	*Draft report*
15	**Clinical sign-out**	Report reviewed and authorised under dual signature with anticipation counselling noted	Sign-out procedure	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
15	**72 h · 12 d**	**SCD-20**	**1,65,000**

SCD-21
SECTOR 01

Mitochondrial Genome Analysis

(**mtDNA sequencing analysis** · mtDNA analysis)

WHY IT IS DONE

Mitochondrial disease is caused by variants in either the mitochondrial or the nuclear genome, and mitochondrial variants behave unlike nuclear ones: they are present at variable fraction rather than as clean genotypes, and that fraction determines whether disease occurs at all.

WHAT IT ANSWERS

Is there a pathogenic mitochondrial variant or deletion, at what heteroplasmy level, in this tissue, and is that level consistent with the clinical presentation?

WHEN IT IS DONE

Where the presentation suggests mitochondrial disease — multi-system involvement, lactic acidaemia, myopathy, encephalopathy, maternal inheritance pattern — and as a routine arm of genome analysis.

WHAT TRIGGERS IT

A clinical suspicion of mitochondrial disease, a maternal inheritance pattern, characteristic biochemical findings, or routine inclusion within genome analysis.

HOW IT IS DONE

Mitochondrial reads are extracted with attention to nuclear mitochondrial sequence contamination, variants are called in a mode designed for variable allele fraction rather than genotype assignment, and heteroplasmy is quantified with confidence bounds. Large deletions are detected from coverage discontinuity, copy number is assessed for depletion, and haplogroup is assigned to inform interpretation.

WHAT IT PRODUCES

A mitochondrial variant report with heteroplasmy levels and bounds, large deletion findings with breakpoints, mitochondrial copy number assessment, haplogroup assignment, and a tissue-specific limitation statement.

WHAT IT DETECTS

Mitochondrial point variants with heteroplasmy quantification, large-scale single deletions, mitochondrial copy number for depletion syndromes, and haplogroup assignment for interpretive context.

WHAT IT DOES NOT DETECT

Nuclear-encoded mitochondrial disease, which accounts for the majority of paediatric mitochondrial disorders and requires a separate nuclear analysis. Tissue-restricted heteroplasmy is also missed where the tested tissue is unaffected, so a blood-negative result does not exclude muscle-restricted disease.

WHAT MAKES IT HARD

Nuclear mitochondrial sequences are near-identical to mitochondrial sequence and generate convincing false variants at low heteroplasmy if not actively excluded. Heteroplasmy is also tissue-dependent, so the report must state which tissue was tested and what that does and does not exclude.

WHO DOES THIS JOB

Clinical bioinformatician or clinical scientist, mid level, in a laboratory serving metabolic or mitochondrial disease services.

WHAT YOU CAN DO AFTERWARDS

Call mitochondrial variants with correct nuclear mitochondrial sequence handling, quantify heteroplasmy defensibly, detect deletions and depletion, and bound results by tissue.

WHO HIRES FOR IT

Mitochondrial disease specialist centres, metabolic services, hospital genomics laboratories, and reference laboratories offering mitochondrial testing.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
16	**76 hours · 13 days**	Prior alignment and variant calling experience	**1,75,000**

SCD-21
WORKFLOW

Mitochondrial Genome Analysis

(**mtDNA sequencing analysis** · mtDNA analysis)

COMPLETE WORKFLOW — 16 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Tissue and indication recording**	Tissue type recorded because heteroplasmy is tissue-dependent and bounds the result	Sample provenance record	*Tissue provenance record*
02	**Mitochondrial read extraction**	Reads mapping to the mitochondrial genome extracted with circularity handled correctly	samtools, chrM extraction	*Mitochondrial read set*
03	**Nuclear mitochondrial sequence exclusion**	Reads originating from nuclear mitochondrial sequences identified and excluded	NUMT-aware realignment	*NUMT-filtered reads*
04	**Mitochondrial alignment**	Reads realigned to the revised Cambridge reference with shifted-origin handling	bwa-mem2, shifted reference	*Mitochondrial BAM*
05	**Coverage assessment**	Per-base mitochondrial coverage measured with uniformity and depth thresholds applied	mosdepth	*Coverage record*
06	**Variant calling in mitochondrial mode**	Variants called on allele fraction rather than genotype assignment	Mutect2 mitochondrial mode	*Mitochondrial VCF*
07	**Heteroplasmy quantification**	Allele fraction estimated per variant with confidence bounds	Allele fraction analysis	*Heteroplasmy estimates*
08	**Detection limit determination**	Depth translated into the minimum reliably detectable heteroplasmy level	Depth-based modelling	*Detection limit statement*
09	**Artefact filtering**	Strand bias, alignment artefacts and residual NUMT signal removed	Bias metrics, artefact filters	*Filtered variant set*
10	**Large deletion detection**	Single large-scale deletions detected from coverage discontinuity with breakpoints defined	Coverage breakpoint analysis	*Deletion findings*
11	**Copy number assessment**	Mitochondrial copy number estimated relative to nuclear coverage for depletion assessment	Coverage ratio analysis	*mtDNA copy number*
12	**Haplogroup assignment**	Haplogroup assigned to provide interpretive context for observed variants	HaploGrep	*Haplogroup assignment*
13	**Annotation and population filtering**	Variants annotated against mitochondrial-specific databases and frequency references	MITOMAP, gnomAD mitochondrial	*Annotated variants*
14	**Classification**	Variants classified under mitochondrial-specific interpretation criteria	Mitochondrial classification framework	*Classified variants*
15	**Orthogonal confirmation**	Findings confirmed by a method sensitive at the heteroplasmy level claimed	ddPCR, deep amplicon sequencing	*Confirmed findings*
16	**Report generation and sign-out**	Variants, heteroplasmy levels and tissue limitations issued under dual signature	Validated report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
16	**76 h · 13 d**	**SCD-21**	**1,75,000**

SCD-22
SECTOR 01

Mosaicism Analysis

(**Low-level mosaic variant detection** · Mosaicism analysis)

WHY IT IS DONE

Mosaic variants sit below the allele fraction at which standard germline callers operate and are therefore missed by routine analysis. They cause a recognised group of overgrowth, vascular, neurocutaneous and epilepsy syndromes, and identifying them changes both diagnosis and, increasingly, treatment.

WHAT IT ANSWERS

Is there a variant present in only a fraction of cells, at what fraction, in which tissue, and what does the distribution imply about when in development the event occurred?

WHEN IT IS DONE

Where the phenotype is asymmetric, segmental or patterned, where a mosaic syndrome is suspected clinically, where a parent of an apparently de novo case may be gonadal mosaic, and in tumour-predisposition contexts.

WHAT TRIGGERS IT

A segmental or asymmetric clinical presentation, a suspected mosaic overgrowth or vascular syndrome, recurrence of an apparently de novo variant in a sibling, or an unexplained low allele fraction observed in routine analysis.

HOW IT IS DONE

Tissue provenance is documented, minimum depth for the target fraction is confirmed, and an empirical background error model is built from known-negative positions in the same assay and tissue. Variants are called on allele fraction significance rather than genotype, strand and position bias artefacts are removed, and fractions are estimated with bounds and confirmed by a method itself sensitive at that fraction.

WHAT IT PRODUCES

Mosaic variant calls with fractions and confidence bounds, the empirical background error model, a depth-derived detection limit per region, tissue distribution and timing inference, artefact filtering records, and matched-sensitivity confirmation results.

WHAT IT DETECTS

Variants present at fractions well below heterozygous levels, with fraction estimates and confidence bounds, tissue distribution where matched samples are available, and developmental timing inference from that distribution.

WHAT IT DOES NOT DETECT

A variant absent from the tested tissue is not detected regardless of depth, so a blood-negative result says little about affected tissue in a segmental condition. Sensitivity is bounded by depth and by the assay's background error rate, and below that floor a negative is uninformative rather than negative.

WHAT MAKES IT HARD

Everything depends on distinguishing signal from the assay's own error floor, which must be measured rather than assumed. Confirmation is the trap: a confirmatory method that is not itself sensitive at the claimed fraction proves nothing, and this is routinely overlooked.

WHO DOES THIS JOB

Senior clinical bioinformatician with statistical competence. A specialist skill with few practitioners and rising demand as targeted therapies for mosaic conditions become available.

WHAT YOU CAN DO AFTERWARDS

Build empirical error models, detect and quantify low-fraction variants defensibly, state tissue-specific detection limits, and design confirmation at matched sensitivity.

WHO HIRES FOR IT

Specialist mosaic and overgrowth syndrome centres, paediatric genomics services, dermatology and vascular anomaly genetics, and laboratories supporting targeted therapy pathways.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
19	**94 hours · 16 days**	Prior variant calling experience, statistics basics	**2,20,000**

SCD-22
WORKFLOW

Mosaicism Analysis

(**Low-level mosaic variant detection** · Mosaicism analysis)

COMPLETE WORKFLOW — 19 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Tissue provenance documentation**	Tissue type and collection method recorded because mosaic distribution is tissue-dependent	Sample provenance record	*Provenance record*
02	**Target fraction specification**	The mosaic fraction the assay must detect specified from the clinical question	Assay specification	*Target sensitivity*
03	**Depth adequacy verification**	Achieved depth confirmed sufficient for the target fraction before analysis proceeds	mosdepth, depth calculation	*Depth acceptance*
04	**UMI-aware processing**	Molecular consensus applied where the chemistry provides unique molecular identifiers	fgbio, UMI-tools	*Consensus BAM*
05	**Duplicate and artefact handling**	Amplification artefacts suppressed without collapsing genuine independent molecules	Duplicate handling policy	*Processed BAM*
06	**Known-negative site selection**	Positions expected to carry no variant selected for error rate measurement	Negative site panel	*Error model sites*
07	**Background error model construction**	Per-substitution and per-context error rates measured empirically for this assay and tissue	Error profiling	*Empirical error model*
08	**Allele fraction modelling**	Variants called on statistical significance against the error model rather than genotype	Statistical calling	*Candidate mosaic variants*
09	**Strand bias filtering**	Calls with strand imbalance inconsistent with genuine variants removed	Strand bias metrics	*Bias-filtered candidates*
10	**Read position bias filtering**	Calls clustered at read ends or in poor alignment context removed	Position bias metrics	*Filtered candidates*
11	**Read-level review**	Every surviving candidate inspected at read level before it is carried forward	IGV review	*Reviewed candidates*
12	**Mosaic fraction estimation**	Allele fraction estimated with confidence intervals and tissue composition considered	Fraction estimation	*Fraction estimates*
13	**Detection limit statement**	Minimum detectable fraction stated per region from achieved depth and error floor	Sensitivity modelling	*Detection limit statement*
14	**Multi-tissue comparison**	Variant fraction compared across matched tissues where available	Matched tissue analysis	*Tissue distribution*
15	**Developmental timing inference**	Timing of the mosaic event inferred from the distribution across tissues	Distribution analysis	*Timing inference*
16	**Annotation and classification**	Variants annotated and classified accounting for mosaic rather than constitutional presence	VEP, ACMG-AMP with mosaic considerations	*Classified variants*
17	**Matched-sensitivity confirmation**	Confirmation performed by a method demonstrably sensitive at the fraction claimed	ddPCR, deep amplicon sequencing	*Confirmed findings*
18	**Recurrence implication assessment**	Germline and gonadal mosaicism implications for recurrence risk assessed	Clinical correlation	*Recurrence assessment*
19	**Report generation and sign-out**	Variants, fractions, bounds and tissue-specific limits issued under dual signature	Validated report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
19	**94 h · 16 d**	**SCD-22**	**2,20,000**

SCD-23
SECTOR 01

Pharmacogenomic Analysis

(**PGx star-allele and diplotype analysis** · PGx analysis)

WHY IT IS DONE

Pharmacogenomic variants determine whether a standard drug dose is therapeutic, ineffective or toxic. Unlike diagnostic genomics the result applies for life across many prescriptions, and for a small number of drugs pre-emptive testing prevents severe and predictable harm.

WHAT IT ANSWERS

What is this patient's diplotype at each pharmacogene, what metabolic phenotype does it translate to, and what prescribing change does the relevant guideline recommend?

WHEN IT IS DONE

Pre-emptively before prescribing drugs with established pharmacogenomic guidance, reactively after an adverse reaction or treatment failure, and increasingly as a standing panel result held in the record for future prescribing.

WHAT TRIGGERS IT

A planned prescription of a drug with pharmacogenomic guidance, an unexplained adverse drug reaction, treatment failure at standard dosing, or enrolment in a pre-emptive testing programme.

HOW IT IS DONE

Star-allele definitions are applied from a versioned nomenclature resource, diplotypes are called with phasing where achievable and ambiguity declared where not, and complex loci with structural and hybrid alleles are analysed with locus-specific methods. Diplotypes are translated to metabolic phenotype and mapped to current prescribing guidelines.

WHAT IT PRODUCES

A diplotype report per pharmacogene with the allele definition version, phenotype translation, guideline-mapped prescribing recommendations, an ambiguity and limitation statement, and a record suitable for durable storage in the patient record.

WHAT IT DETECTS

Star-allele diplotypes across the pharmacogenes, structural and hybrid alleles at complex loci, translated metabolic phenotypes, and guideline-mapped prescribing recommendations.

WHAT IT DOES NOT DETECT

Novel or uncatalogued alleles are typically defaulted to normal function, which can be wrong. Phasing ambiguity means some diplotypes cannot be resolved from short reads alone, and the analysis explains only the genetic component of drug response, not interactions, adherence, renal or hepatic function.

WHAT MAKES IT HARD

Star-allele calling is not variant calling. It requires haplotype-level reasoning that short reads support imperfectly, and complex loci with hybrid structures defeat naive approaches entirely. Translation from diplotype to phenotype to recommendation must follow current guidelines exactly, since these are updated regularly.

WHO DOES THIS JOB

Clinical bioinformatician or pharmacogenomics scientist, mid level, working alongside clinical pharmacy. A distinct discipline that sits between genomics and pharmacy.

WHAT YOU CAN DO AFTERWARDS

Call star alleles and diplotypes including at complex loci, translate to phenotype correctly, apply current prescribing guidelines, and produce durable records for lifelong reuse.

WHO HIRES FOR IT

Hospital pharmacy and therapeutics services, pre-emptive pharmacogenomics programmes, diagnostic chains offering PGx panels, and pharmaceutical companies running clinical pharmacogenomics.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
16	**78 hours · 13 days**	Familiarity with variant calling and haplotype concepts	**1,80,000**

SCD-23
WORKFLOW

Pharmacogenomic Analysis

(**PGx star-allele and diplotype analysis** · PGx analysis)

COMPLETE WORKFLOW — 16 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Pharmacogene panel authorisation**	The gene set, drug scope and guideline source confirmed and recorded	PGx panel register	*Authorised panel version*
02	**Allele definition version lock**	Star-allele nomenclature and definition table version confirmed and checksummed	PharmVar allele definitions	*Definition version record*
03	**Coverage verification at pharmacogenes**	Depth confirmed adequate across every defining position for every star allele	mosdepth, bedtools coverage	*Coverage certificate*
04	**Variant calling at defining positions**	Variants called at all allele-defining positions including small indels	Validated caller	*Pharmacogene VCF*
05	**Structural variant detection at complex loci**	Gene deletions, duplications and hybrid structures detected at complex pharmacogenes	Copy number analysis, Cyrius	*Structural findings*
06	**Phasing**	Variants phased into haplotypes where read or population data permit	WhatsHap, statistical phasing	*Phased haplotypes*
07	**Star-allele assignment**	Haplotypes matched against the versioned allele definition table	PharmCAT, StellarPGx	*Star-allele calls*
08	**CYP2D6 specialist analysis**	Copy number, hybrid alleles and tandem arrangements resolved with a locus-specific method	Cyrius, StellarPGx	*CYP2D6 diplotype*
09	**Diplotype determination**	Diplotype assigned per gene with ambiguity declared where phasing is unresolved	Diplotype caller	*Diplotype calls*
10	**Ambiguity documentation**	Unresolvable diplotype pairs and uncatalogued alleles documented explicitly	Ambiguity register	*Ambiguity statement*
11	**Phenotype translation**	Diplotypes translated to metabolic phenotype using the current activity score system	Activity score tables	*Metabolic phenotypes*
12	**Guideline mapping**	Phenotypes mapped to current prescribing recommendations for each drug in scope	CPIC, DPWG guidelines	*Prescribing recommendations*
13	**Guideline version recording**	Guideline version and date recorded since recommendations are revised regularly	Guideline register	*Guideline version record*
14	**Orthogonal confirmation**	High-impact findings confirmed by an independent genotyping method	Targeted genotyping, ddPCR	*Confirmed findings*
15	**Durable record preparation**	Results formatted for lifelong storage and reuse in the patient record	Structured PGx record format	*Durable PGx record*
16	**Report generation and sign-out**	Diplotypes, phenotypes and prescribing recommendations issued under dual signature	Validated PGx report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
16	**78 h · 13 d**	**SCD-23**	**1,80,000**

SCD-24
SECTOR 01

HLA Typing Analysis

(**HLA genotyping from sequencing data** · HLA typing)

WHY IT IS DONE

HLA typing underpins transplant matching, drug hypersensitivity risk prediction and disease association testing. Deriving it from existing sequencing data rather than by dedicated typing removes a separate test, though the extreme polymorphism of the locus makes it analytically demanding.

WHAT IT ANSWERS

What are this individual's HLA alleles at each locus, to what field resolution, and with what confidence given the ambiguity inherent in the region?

WHEN IT IS DONE

For transplant donor and recipient matching, before prescribing drugs with HLA-linked hypersensitivity risk, in disease association studies, and as an additional output extracted from existing genome or exome data.

WHAT TRIGGERS IT

A transplant workup, a planned prescription of a drug with an HLA risk association, a research or association study requirement, or a request to extract typing from existing sequencing data.

HOW IT IS DONE

Reads mapping to the HLA region and to unplaced contigs are extracted, and a specialist typing method that handles the region's polymorphism explicitly is applied against a versioned allele database. Resolution and ambiguity are declared per locus, and results are compared against orthogonal typing where a clinical decision depends on them.

WHAT IT PRODUCES

A per-locus allele call report with field resolution and the database version, an ambiguity statement per locus, a resolution comparison against dedicated typing where available, and risk allele findings where drug hypersensitivity screening was the purpose.

WHAT IT DETECTS

HLA alleles at class I and class II loci to two-field or higher resolution depending on method and data type, with ambiguity where the data cannot discriminate between allele pairs.

WHAT IT DOES NOT DETECT

Resolution from short-read exome data is materially lower than from dedicated typing, and some allele pairs cannot be discriminated at all. Novel alleles absent from the reference database are typically forced to the nearest known allele, and rare alleles in under-represented populations are systematically less reliable.

WHAT MAKES IT HARD

The HLA region is the most polymorphic in the human genome and standard alignment fails there, so a specialist method is not optional. Resolution and ambiguity must be reported honestly, because a transplant decision taken on an over-stated resolution is a serious error.

WHO DOES THIS JOB

Clinical bioinformatician, mid level, in a transplant, immunogenetics or pharmacogenomics setting.

WHAT YOU CAN DO AFTERWARDS

Extract and type HLA from sequencing data with an appropriate specialist method, declare resolution and ambiguity honestly, and identify drug hypersensitivity risk alleles.

WHO HIRES FOR IT

Transplant and immunogenetics laboratories, tissue typing services, pharmacogenomics programmes, and hospital laboratories consolidating typing into sequencing workflows.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
13	**62 hours · 11 days**	Prior alignment analysis experience	**1,40,000**

SCD-24
WORKFLOW

HLA Typing Analysis

(**HLA genotyping from sequencing data** · HLA typing)

COMPLETE WORKFLOW — 13 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Typing scope and purpose definition**	Loci required and the resolution needed established from the clinical purpose	Requisition review	*Typing scope*
02	**Allele database version lock**	The IMGT/HLA database release used for typing confirmed and recorded	IMGT/HLA release register	*Database version record*
03	**HLA region read extraction**	Reads mapping to the HLA region, alternate contigs and unplaced contigs extracted	samtools region extraction	*HLA read set*
04	**Unmapped read recovery**	Unmapped reads recovered since divergent alleles frequently fail standard alignment	Unmapped read extraction	*Recovered read set*
05	**Coverage assessment at HLA loci**	Depth confirmed adequate at each locus for the resolution being claimed	mosdepth	*Coverage record*
06	**Specialist HLA typing**	Reads typed against the allele database with a method built for extreme polymorphism	HLA-LA, OptiType, xHLA	*Raw allele calls*
07	**Resolution assessment**	Field resolution achieved determined per locus from the supporting evidence	Resolution analysis	*Resolution record*
08	**Ambiguity determination**	Allele pairs the data cannot discriminate identified and declared per locus	Ambiguity analysis	*Ambiguity statement*
09	**Novel allele flagging**	Sequences not matching any catalogued allele flagged rather than forced to the nearest match	Mismatch analysis	*Novel allele flags*
10	**Risk allele screening**	Drug hypersensitivity risk alleles specifically screened where that is the purpose	Risk allele panel	*Risk allele findings*
11	**Orthogonal comparison**	Results compared against dedicated typing where a clinical decision depends on them	Sequence-based typing, SSO	*Comparison record*
12	**Interpretation for purpose**	Results interpreted for matching, risk prediction or association as the referral requires	Clinical interpretation	*Interpretation record*
13	**Report generation and sign-out**	Alleles, resolution, ambiguity and database version issued under dual signature	Validated report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
13	**62 h · 11 d**	**SCD-24**	**1,40,000**

SCD-25
SECTOR 01

Carrier Screening Analysis

(**Expanded carrier screening analysis** · ECS analysis)

WHY IT IS DONE

Carrier screening identifies reproductive risk in people who are themselves healthy, allowing informed reproductive choices before or during pregnancy. Because the population tested is asymptomatic, the tolerance for false positives and for over-called variants is far lower than in diagnostic testing.

WHAT IT ANSWERS

Is this individual or couple a carrier for any recessive or X-linked condition on the screened gene set, and what is the resulting reproductive risk?

WHEN IT IS DONE

Pre-conception, in early pregnancy, before assisted reproduction, and in populations with recognised founder variants. Increasingly offered routinely rather than only on family history.

WHAT TRIGGERS IT

Pre-conception counselling, a fertility or assisted reproduction pathway, known consanguinity, ethnicity-associated founder risk, or a family history of a recessive condition.

HOW IT IS DONE

The screened gene set and its version are authorised, per-gene coverage is certified, and both sequence and copy number analysis are applied with particular attention to pseudogene-bearing loci requiring alternative methods. Classification is conservative given the asymptomatic population, and residual risk is calculated per gene and per ancestry.

WHAT IT PRODUCES

A carrier status report per gene, per-gene coverage certificate, residual risk calculation by gene and ancestry, combined couple risk where applicable, and reproductive options counselling support material.

WHAT IT DETECTS

Carrier status for sequence variants and copy number changes across the screened gene set, founder variants at population-appropriate sensitivity, and combined couple risk where both partners are tested.

WHAT IT DOES NOT DETECT

Carrier status is not exhaustive — residual risk always remains from variants outside the panel, variants of uncertain significance, and technically difficult genes. Notably weak in pseudogene-bearing genes central to carrier screening, and it does not predict severity or age of onset in an affected pregnancy.

WHAT MAKES IT HARD

Residual risk calculation is the deliverable that matters most and is most often omitted. Carrier screening also concentrates on genes that are analytically difficult, so a panel-wide sensitivity claim is unsupportable and gene-specific residual risk must be stated instead.

WHO DOES THIS JOB

Variant scientist or clinical bioinformatician, mid level, working closely with genetic counselling on reproductive risk communication.

WHAT YOU CAN DO AFTERWARDS

Operate a carrier screening pathway, manage pseudogene-bearing loci correctly, calculate and communicate residual risk per gene, and derive combined couple risk.

WHO HIRES FOR IT

Reproductive genetics and IVF chains, prenatal diagnostic laboratories, hospital genetics services, and consumer and clinical carrier screening providers.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
15	**70 hours · 12 days**	Familiarity with variant classification	**1,60,000**

SCD-25
WORKFLOW

Carrier Screening Analysis

(**Expanded carrier screening analysis** · ECS analysis)

COMPLETE WORKFLOW — 15 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Screened gene set authorisation**	Gene set version, condition list and severity criteria confirmed and recorded	Carrier panel register	*Authorised gene set*
02	**Ancestry and family history capture**	Reported ancestry and family history recorded for residual risk calculation	Requisition, counselling record	*Ancestry record*
03	**Consent and scope confirmation**	Reproductive scope, partner testing and incidental finding policy confirmed	Consent record	*Analysis authorisation*
04	**Coverage certification per gene**	Depth certified for every gene on the panel with shortfalls named individually	mosdepth, bedtools coverage	*Coverage certificate*
05	**Identity verification**	Fingerprint captured to confirm attribution, critical where couples are tested together	somalier extract	*Identity clearance*
06	**Variant calling**	Variants called across the screened gene set with validated parameters	Validated caller	*Carrier VCF*
07	**Copy number calling**	Exon-level copy number called since deletions are a major carrier mechanism	gCNV, DECoN	*Carrier CNV calls*
08	**Pseudogene-aware locus handling**	Genes with pseudogene interference routed to documented alternative methods	MLPA, long-read, specialist assay	*Specialist locus results*
09	**Founder variant screening**	Population-specific founder variants specifically interrogated at required sensitivity	Founder variant panel	*Founder variant results*
10	**Annotation and population filtering**	Variants annotated with carrier-appropriate frequency thresholds by ancestry	VEP, gnomAD by ancestry	*Annotated variants*
11	**Conservative classification**	Variants classified conservatively given testing of an asymptomatic population	ACMG-AMP with gene specifications	*Classified variants*
12	**Residual risk calculation**	Post-test residual carrier risk calculated per gene and per reported ancestry	Detection rate and frequency tables	*Residual risk figures*
13	**Combined couple risk calculation**	Joint reproductive risk calculated where both partners have been tested	Couple risk calculation	*Couple risk assessment*
14	**Orthogonal confirmation**	Reportable carrier findings confirmed by an independent method	Sanger, MLPA, ddPCR	*Confirmed findings*
15	**Report generation and sign-out**	Carrier status, residual risk and reproductive options issued under dual signature	Validated report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
15	**70 h · 12 d**	**SCD-25**	**1,60,000**

SCD-26
SECTOR 01

Genomic Reanalysis

(**Periodic reanalysis and reclassification** · Reanalysis)

WHY IT IS DONE

A negative genomic test is provisional, not final. Gene-disease associations are published continuously, classification criteria are refined, pipelines improve and phenotypes evolve. Systematic reanalysis converts a meaningful share of previously negative cases into diagnoses at a fraction of the cost of new sequencing.

WHAT IT ANSWERS

Given everything that has changed since the original analysis, does this previously negative or uncertain case now have a diagnosis, and does any previously reported variant require reclassification?

WHEN IT IS DONE

At a defined interval after a negative result, when the phenotype changes materially, when a specific new gene-disease association is published that fits the case, or when a batch reclassification review is triggered.

WHAT TRIGGERS IT

Elapsed reanalysis interval under laboratory policy, a clinician request following phenotype evolution, publication of a relevant new gene-disease association, or a knowledge base update flagging a previously reported variant.

HOW IT IS DONE

The reanalysis trigger is authorised and archived data are retrieved with integrity and identity verified. The knowledge base delta is established, a documented decision is taken on whether re-calling is warranted, phenotype updates are captured, and re-annotation and reclassification proceed against current resources with active referrer notification on any amendment.

WHAT IT PRODUCES

The trigger and authorisation record, archive integrity and identity verification, knowledge base delta report, phenotype update record, reclassification outcomes, newly reportable findings, and the amended report with documented referrer notification.

WHAT IT DETECTS

Newly reportable variants arising from new gene-disease associations, variants requiring upgrade or downgrade under current criteria, and findings recoverable through improved pipeline sensitivity.

WHAT IT DOES NOT DETECT

Anything the original data could not capture. Reanalysis cannot recover a region that was never covered, cannot detect variant classes the original assay was blind to, and cannot compensate for a fundamentally inadequate original test. Where the limitation is the data rather than the interpretation, re-sequencing is required.

WHAT MAKES IT HARD

Bounding the exercise is the discipline. Without a defined knowledge delta, reanalysis becomes an unbounded repetition of the original work. Downgrading a previously reported pathogenic variant is also harder to communicate than reporting a new finding, and an amended report that reaches nobody has no clinical effect.

WHO DOES THIS JOB

Variant scientist or clinical bioinformatician, mid level, often as a dedicated function in laboratories operating formal reanalysis programmes.

WHAT YOU CAN DO AFTERWARDS

Run a governed reanalysis programme, establish and bound the knowledge delta, execute reclassification in both directions, and close the communication loop on amended reports.

WHO HIRES FOR IT

Hospital genomics laboratories, national rare disease programmes, diagnostic chains offering reanalysis as a product, and any laboratory with an accumulating negative case archive.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
14	**66 hours · 11 days**	Prior exome or genome analysis experience	**1,50,000**

SCD-26
WORKFLOW

Genomic Reanalysis

(**Periodic reanalysis and reclassification** · Reanalysis)

COMPLETE WORKFLOW — 14 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Reanalysis trigger authorisation**	The trigger criterion, elapsed interval and authorising clinician recorded	Reanalysis policy	*Reanalysis authorisation*
02	**Archive retrieval**	Archived alignment and variant files retrieved from the storage tier holding them	Archive retrieval procedure	*Retrieved data*
03	**Integrity and decodability verification**	Checksums verified and reference decodability confirmed for compressed archives	Checksum verification, reference registry	*Verified archived data*
04	**Identity re-verification**	Fingerprint compared to confirm the data belong to the patient on the current request	somalier fingerprint comparison	*Identity clearance*
05	**Original analysis reconstruction**	The pipeline, resources and knowledge base versions used originally established	Original case record	*Baseline record*
06	**Knowledge base delta assessment**	Genes newly associated with disease and variants reclassified since sign-out identified	ClinVar, OMIM, gene validity resources	*Knowledge delta*
07	**Pipeline delta assessment**	Current pipeline compared against the original to decide whether re-calling is warranted	Version comparison	*Re-calling decision*
08	**Phenotype update capture**	New or evolved clinical features captured in current HPO terms	Updated clinical summary	*Updated phenotype*
09	**Selective re-calling**	Re-calling performed only where pipeline improvement materially affects sensitivity	Current validated caller	*Refreshed call set*
10	**Re-annotation**	Full re-annotation performed against current resources and transcript sets	VEP, gnomAD, ClinVar	*Re-annotated variants*
11	**Re-prioritisation**	Candidates re-ranked against the updated phenotype and new gene-disease associations	Exomiser, HPO terms	*Re-prioritised candidates*
12	**Reclassification review**	Previously reported variants reassessed under current criteria in both directions	ACMG-AMP current specifications	*Reclassification outcomes*
13	**Newly reportable finding assessment**	Candidates arising from new associations classified in full and confirmed	ACMG-AMP, Sanger confirmation	*Newly classified variants*
14	**Amended report issue and notification**	Amended report issued, original superseded and the referrer actively notified	Amended report template, notification log	*Amended report and notification record*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
14	**66 h · 11 d**	**SCD-26**	**1,50,000**

SCD-27
SECTOR 01

Clinical Long-Read Analysis

(**Long-read sequencing analysis for diagnostics** · Long-read analysis)

WHY IT IS DONE

Long reads resolve the regions short reads cannot: repeat expansions, segmental duplications, pseudogene-bearing loci, structural rearrangement breakpoints and phasing across long distances. They also read methylation directly from the same data, without a separate assay.

WHAT IT ANSWERS

What is present in the regions short-read sequencing cannot resolve, and can a case that short-read analysis left unexplained be solved by reading through the difficult sequence?

WHEN IT IS DONE

After a negative short-read workup in a phenotype suggesting a structural, repeat or pseudogene-mediated cause, to resolve a rearrangement breakpoint, or as a targeted assay for specific difficult loci.

WHAT TRIGGERS IT

A negative short-read genome or exome in a compatible phenotype, a known difficult locus requiring resolution, an unresolved structural finding, or a suspected repeat expansion beyond short-read sizing capacity.

HOW IT IS DONE

Basecalling model and version are recorded because they materially affect accuracy, reads are aligned with a long-read aware method, and small variants, structural variants and repeats are called with long-read specific tools. Phasing is applied across long distances and methylation is extracted from the same data where the chemistry supports it.

WHAT IT PRODUCES

A phased variant call set, structural variant findings with breakpoint resolution, accurately sized repeat genotypes, difficult locus resolution results, native methylation output where applicable, and the basecalling model and version record.

WHAT IT DETECTS

Repeat expansions with accurate sizing, structural variants with breakpoint resolution, variants in segmental duplications and pseudogene-bearing genes, long-range phasing, and native methylation from the same reads.

WHAT IT DOES NOT DETECT

Per-base accuracy is lower than short reads for some chemistries, so small variant calling in easy regions is not necessarily better. Throughput and cost per sample remain higher, the clinical evidence base and reference resources are less mature, and validation for accredited use is substantially more work.

WHAT MAKES IT HARD

It is a genuinely different analytical stack rather than a variation on short-read analysis, and the tooling changes rapidly. Validating it to accredited clinical standard is significantly harder than validating a short-read assay, and the reference and population resources are far less developed.

WHO DOES THIS JOB

Senior clinical bioinformatician. Currently one of the scarcest skill sets in clinical genomics and among the strongest differentiators on a CV in this sector.

WHAT YOU CAN DO AFTERWARDS

Operate a clinical long-read analysis stack, resolve loci short reads cannot address, phase across long distances, extract native methylation, and approach validation of a long-read clinical assay.

WHO HIRES FOR IT

Advanced hospital genomics laboratories, national programmes adopting long-read platforms, specialist neurogenetics and structural variant services, and technology providers building clinical products.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
20	**100 hours · 17 days**	Prior short-read genome analysis experience	**2,35,000**

SCD-27
WORKFLOW

Clinical Long-Read Analysis

(**Long-read sequencing analysis for diagnostics** · Long-read analysis)

COMPLETE WORKFLOW — 20 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Platform and chemistry recording**	Sequencing platform, flow cell and chemistry version recorded as they bound accuracy	Run record	*Platform record*
02	**Basecalling model version lock**	Basecalling model and version recorded since it materially affects accuracy	Basecaller configuration	*Model version record*
03	**Basecalling**	Raw signal converted to basecalls using the validated high-accuracy model	Dorado	*Basecalled reads*
04	**Read quality assessment**	Read length distribution, quality and yield assessed against acceptance criteria	NanoPlot, pycoQC	*Read QC record*
05	**Read filtering**	Short, low-quality and chimeric reads removed with retention accounting	Filtlong, chopper	*Filtered reads*
06	**Long-read alignment**	Reads aligned with a method designed for long reads and high indel error	minimap2, pbmm2	*Aligned BAM*
07	**Alignment quality assessment**	Mapping rate, coverage uniformity and clipping assessed across the genome	samtools stats, mosdepth	*Alignment QC record*
08	**Identity and contamination verification**	Fingerprint captured and contamination estimated from long-read data	somalier, allele fraction review	*Identity clearance*
09	**Small variant calling**	Small variants called with a long-read specific model rather than a short-read caller	Clair3, DeepVariant long-read model	*Small variant VCF*
10	**Structural variant calling**	Structural variants called with breakpoint resolution from spanning reads	Sniffles2, cuteSV	*SV call set*
11	**Repeat expansion sizing**	Repeat loci sized directly from reads spanning the full expansion	Straglr, tandem-genotypes	*Repeat genotypes*
12	**Long-range phasing**	Variants phased across long distances using read-backed phasing	WhatsHap	*Phased VCF*
13	**Haplotagging**	Reads tagged by haplotype to enable allele-resolved downstream analysis	WhatsHap haplotag	*Haplotagged BAM*
14	**Native methylation calling**	Methylation extracted directly from the same reads where chemistry supports it	Dorado modified basecalls, modkit	*Methylation calls*
15	**Difficult locus resolution**	Pseudogene-bearing and segmentally duplicated loci resolved using full-length reads	Targeted locus analysis	*Difficult locus results*
16	**Annotation and population filtering**	Variants annotated with attention to long-read specific representation differences	VEP, gnomAD, gnomAD-SV	*Annotated variants*
17	**Phenotype-driven prioritisation**	Candidates ranked against phenotype across sequence, structural and repeat classes	Exomiser, HPO terms	*Ranked candidates*
18	**Classification**	Findings classified under the framework appropriate to each variant class	ACMG-AMP, ACMG-ClinGen	*Classified findings*
19	**Orthogonal confirmation**	Findings confirmed by class-appropriate independent methods	PCR, Sanger, MLPA, ddPCR	*Confirmed findings*
20	**Report generation and sign-out**	Phased results, resolved loci and methylation issued under dual signature	Validated report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
20	**100 h · 17 d**	**SCD-27**	**2,35,000**

SCD-28
SECTOR 01

Diagnostic RNA-Seq Analysis

(**RNA studies for variant interpretation** · RNA studies · splice confirmation)

WHY IT IS DONE

DNA analysis frequently produces a variant whose effect on splicing or expression is predicted but not proven. RNA analysis converts that prediction into evidence, resolves variants of uncertain significance, and independently identifies causes that DNA analysis missed entirely.

WHAT IT ANSWERS

Does this variant actually affect splicing or expression in the patient's cells, and are there transcript-level abnormalities pointing to a cause that DNA analysis did not reveal?

WHEN IT IS DONE

After a DNA finding of uncertain significance with predicted splice or expression impact, after a negative DNA workup in a strongly suspected monogenic condition, and increasingly as a routine parallel arm in diagnostic pathways.

WHAT TRIGGERS IT

A variant of uncertain significance with predicted splice impact, a negative exome or genome in a compelling phenotype, or a suspected deep intronic or regulatory cause.

HOW IT IS DONE

Tissue expression suitability is confirmed for the genes of interest before sequencing. Reads are aligned with a splice-aware method, splicing is analysed against a matched control cohort for aberrant events, expression outliers and allele-specific expression are assessed, and findings are linked back to the candidate DNA variant as functional evidence for reclassification.

WHAT IT PRODUCES

An aberrant splicing report with event types and supporting reads, expression outlier findings, allele-specific expression results, the tissue expression suitability assessment, and the functional evidence record supporting DNA variant reclassification.

WHAT IT DETECTS

Aberrant splicing including exon skipping, cryptic site activation and intron retention, allele-specific and monoallelic expression, expression outliers indicating loss of function, and the functional consequence of specific candidate variants.

WHAT IT DOES NOT DETECT

Only genes expressed in the tissue sampled can be assessed, and blood or fibroblasts do not express many disease genes at usable levels. Nonsense-mediated decay can remove the very transcript that carries the evidence, and the analysis identifies functional consequence rather than the causal variant itself.

WHAT MAKES IT HARD

Tissue choice determines everything, and analysis is frequently attempted in a tissue that does not express the gene at all. A matched control cohort is essential because aberrant splicing is only recognisable relative to normal, and building that cohort is the main barrier to establishing the service.

WHO DOES THIS JOB

Clinical bioinformatician with transcriptomics competence, mid to senior level. A bridging skill between diagnostic genomics and RNA analysis, and currently in short supply.

WHAT YOU CAN DO AFTERWARDS

Run diagnostic RNA analysis against a control cohort, detect and characterise aberrant splicing, assess allele-specific expression, and generate functional evidence for variant reclassification.

WHO HIRES FOR IT

Advanced hospital genomics laboratories, national rare disease programmes, neuromuscular and metabolic specialist services, and reference laboratories offering functional evidence services.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
18	**88 hours · 15 days**	Prior DNA variant analysis experience	**2,05,000**

SCD-28
WORKFLOW

Diagnostic RNA-Seq Analysis

(**RNA studies for variant interpretation** · RNA studies · splice confirmation)

COMPLETE WORKFLOW — 18 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Tissue expression suitability assessment**	Confirmation that the genes of interest are expressed in the tissue available	Expression reference data	*Suitability record*
02	**Control cohort verification**	A matched control cohort confirmed available since aberrance is defined relative to normal	Internal control cohort	*Control cohort record*
03	**RNA quality assessment**	RNA integrity, ribosomal contamination and library complexity assessed	RNA QC metrics, MultiQC	*RNA QC record*
04	**Read preprocessing**	Adapters and low-quality tails removed with strand information preserved	fastp	*Trimmed FASTQ*
05	**Splice-aware alignment**	Reads aligned with a splice-aware method using an annotation-guided index	STAR two-pass	*Aligned BAM*
06	**Alignment and library metrics**	Mapping rate, duplication, strand specificity and gene body coverage assessed	RSeQC, Qualimap	*Alignment QC record*
07	**Identity verification**	Fingerprint compared against the patient's DNA sample to confirm attribution	somalier extract	*Identity clearance*
08	**Gene and transcript quantification**	Expression quantified at gene and transcript level against the control cohort	featureCounts, Salmon	*Expression matrices*
09	**Expression outlier detection**	Genes with expression far outside the control distribution identified statistically	OUTRIDER	*Expression outliers*
10	**Aberrant splicing detection**	Exon skipping, cryptic site activation and intron retention detected against controls	FRASER, rMATS, LeafCutter	*Splicing outliers*
11	**Splice event characterisation**	Each aberrant event characterised by type, supporting reads and predicted consequence	Event annotation, sashimi plots	*Event characterisation*
12	**Allele-specific expression analysis**	Allelic imbalance assessed at heterozygous sites to detect monoallelic expression	ASE analysis, phASER	*ASE results*
13	**Nonsense-mediated decay assessment**	Evidence considered for transcripts degraded before they could be observed	NMD analysis	*NMD assessment*
14	**DNA variant linkage**	RNA findings linked back to the specific candidate DNA variant under investigation	Variant-event correlation	*Linkage record*
15	**Functional evidence derivation**	RNA evidence translated into the applicable ACMG-AMP functional evidence strength	ACMG-AMP PS3 or BS3	*Functional evidence weight*
16	**Variant reclassification**	The DNA variant reclassified with the RNA functional evidence incorporated	ACMG-AMP reassessment	*Reclassification outcome*
17	**Novel candidate assessment**	Aberrant events in genes not previously implicated assessed as new candidates	Candidate review	*Novel candidates*
18	**Report generation and sign-out**	Splicing, expression and functional evidence issued under dual signature	Validated report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
18	**88 h · 15 d**	**SCD-28**	**2,05,000**

SCD-29
SECTOR 01

Methylation Episignature Analysis

(**Episignature and imprinting methylation analysis** · Episignature · EpiSign)

WHY IT IS DONE

A growing group of Mendelian disorders produces a characteristic genome-wide methylation pattern that identifies the condition directly, independent of the causal variant. This resolves variants of uncertain significance, diagnoses imprinting disorders, and identifies cases where no causal variant is found at all.

WHAT IT ANSWERS

Does this patient's methylation profile match a known disease episignature, and does that confirm a diagnosis or resolve an uncertain variant?

WHEN IT IS DONE

After an uncertain variant in a gene with a known episignature, in suspected imprinting disorders, after a negative sequence workup in a phenotype with a recognised methylation signature, and to distinguish epimutation from uniparental disomy.

WHAT TRIGGERS IT

A variant of uncertain significance in an episignature-associated gene, a clinical suspicion of an imprinting disorder, or a negative sequence analysis in a compatible chromatinopathy phenotype.

HOW IT IS DONE

Methylation data are generated and quality controlled with attention to cell composition and batch, then classified against a reference episignature catalogue using a validated model with confidence scoring. Imprinted loci are assessed individually, epimutation is distinguished from uniparental disomy, and results are integrated with the sequence findings for reclassification.

WHAT IT PRODUCES

An episignature classification report with confidence scores, imprinted locus methylation results, epimutation and uniparental disomy discrimination, quality control covering cell composition and batch, and the reclassification evidence record.

WHAT IT DETECTS

Matches to catalogued disease episignatures with classification confidence, imprinting defects at known loci, methylation-based distinction between epimutation and uniparental disomy, and methylation evidence supporting reclassification of an uncertain variant.

WHAT IT DOES NOT DETECT

Only conditions with a catalogued episignature can be identified, which is a limited and slowly growing set. Signatures are tissue-specific and defined in blood, so other tissues are not interpretable against them. Age, cell composition and technical batch all shift methylation and can produce misclassification if not controlled.

WHAT MAKES IT HARD

It is a classification problem rather than a variant-calling problem, and it depends entirely on reference cohort quality and batch control. Age and blood cell composition shift methylation substantially, and without correction a healthy sample can be misclassified into a disease signature.

WHO DOES THIS JOB

Clinical bioinformatician or clinical scientist with methylation and classification competence, mid to senior level. A specialist niche with very few practitioners.

WHAT YOU CAN DO AFTERWARDS

Run episignature classification with proper batch and composition control, interpret confidence scores appropriately, assess imprinted loci, and integrate methylation with sequence evidence.

WHO HIRES FOR IT

Specialist epigenetics diagnostic centres, national rare disease programmes, advanced hospital genomics laboratories, and imprinting disorder services.

COMMERCIALS

STEPS IN WORKFLOW	TRAINING DURATION	PREREQUISITE	TRAINING FEE (INR)
15	**74 hours · 13 days**	Familiarity with methylation data and classification methods	**1,70,000**

SCD-29
WORKFLOW

Methylation Episignature Analysis

(**Episignature and imprinting methylation analysis** · Episignature · EpiSign)

COMPLETE WORKFLOW — 15 ORDERED STEPS

#	STEP	WHAT HAPPENS	TOOL / METHOD	OUTPUT
01	**Referral scoping**	The candidate condition or variant and the episignature question defined	Requisition, prior report	*Scoped referral*
02	**Sample suitability confirmation**	Blood confirmed as the tissue, since episignatures are defined in blood specifically	Sample provenance	*Suitability record*
03	**Methylation data generation**	Genome-wide methylation measured on the validated platform	Methylation array or EM-seq	*Raw methylation data*
04	**Signal quality control**	Detection rates, control probe performance and technical failures assessed	Array or sequencing QC	*Methylation QC record*
05	**Normalisation**	Data normalised to remove technical variation prior to any classification	Normalisation pipeline	*Normalised beta values*
06	**Batch effect assessment**	Processing batch tested for structure that could distort classification	Batch analysis	*Batch assessment*
07	**Cell composition estimation**	Blood cell proportions estimated since composition shifts methylation substantially	Cell deconvolution	*Composition estimates*
08	**Age and covariate correction**	Age and composition effects corrected before comparison to reference signatures	Covariate correction	*Corrected methylation data*
09	**Reference cohort matching**	Age and composition matched controls selected from the reference cohort	Reference cohort selection	*Matched controls*
10	**Episignature classification**	Profile classified against the catalogued disease episignatures with confidence scoring	Validated classification model	*Classification scores*
11	**Confidence interpretation**	Scores interpreted against validated thresholds with intermediate results declared	Threshold criteria	*Interpreted classification*
12	**Imprinted locus methylation assessment**	Methylation assessed individually at known imprinted loci	Imprinted locus panel	*Imprinting results*
13	**Epimutation versus disomy discrimination**	Methylation-based distinction made between epimutation and uniparental disomy	Comparative analysis	*Mechanism determination*
14	**Sequence evidence integration**	Methylation findings integrated with existing sequence results for reclassification	Combined evidence review	*Reclassification evidence*
15	**Report generation and sign-out**	Classification, confidence, imprinting results and mechanism issued under dual signature	Validated report template	*Signed clinical report*

STEPS IN WORKFLOW	TRAINING DURATION	PROGRAMME CODE	TRAINING FEE (INR)
15	**74 h · 13 d**	**SCD-29**	**1,70,000**

Staff and accredit a *clinical genomics laboratory*

Each of the twenty-nine programmes in this catalogue trains one complete type of NGS data analysis as an accredited laboratory performs it — not a tool, not a step, but the whole analysis from the data that arrives to the report that is signed. Trainees run the pipeline on live clinical data, break it, diagnose the failure, and produce the deliverable report set a laboratory would place in a case file and defend to an assessor. Programmes are named the way laboratories name their own work, so a trainee returns to the bench able to run the analysis on the requisition rather than a generic version of it. Individual programmes may be taken alone, combined into a track, or delivered in-house at the client laboratory against that laboratory's own assays and gene lists.

CATALOGUE

SCD · Sector 01 of 16

SCOPE

29 analysis types · 502 steps

DURATION

2450 contact hours · 419 working days

COMPLETE CATALOGUE FEE

INR 37,02,000 (list 56,95,000)

15%	22%	28%	35%
ANY THREE PROGRAMMES	CORE TRACK	PROFESSIONAL TRACK	COMPLETE CATALOGUE