**NTHRYS**

OPC PVT LTD · HYDERABAD · INDIA

SECTOR CATALOGUE 05 · NGS DATA ANALYSIS

# Biologics Manufacturing & *GMP Quality Control*

Twenty-five industrial training programmes, one for each distinct type of NGS data analysis performed in biologics, cell and gene therapy manufacturing under quality systems — from cell bank characterisation and viral safety through vector and product release testing to method validation and data integrity. Every programme answers the questions a laboratory 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 — and is followed by its complete ordered workflow, step by step.


| 25 | 481 | 2328 | 397 |
|---|---|---|---|
| ANALYSIS TYPES | STEPS | HOURS | DAYS |

CATALOGUE SBQ · 2026 EDITION

INDUSTRIAL TRAINING · GMP QUALITY PRACTICE

# Biologics Manufacturing & GMP Quality Control — *Analysis Types*

Twenty-five distinct types of NGS data analysis performed in regulated biologics manufacturing, 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 analysis type occupies a two-page entry: the first page answers why, when, what it detects, what it does not detect, who does the job and what follows; the second sets out the complete ordered workflow.

| CODE | ANALYSIS TYPE · INDUSTRY NAME & SHORT NAME | STEPS | HOURS | DAYS | FEE (INR) |
|---|---|---|---|---|---|
| SBQ-01 | **Master Cell Bank Genomic Characterisation**<br>Master cell bank genomic characterisation · MCB characterisation | 22 | 106 | 18 | **2,45,000** |
| SBQ-02 | **Working Cell Bank & End-of-Production Cell Analysis**<br>WCB and EOPC genomic testing · WCB · EOPC analysis | 19 | 92 | 16 | **2,10,000** |
| SBQ-03 | **Cell Line Genetic Stability Analysis**<br>Producer cell line genetic stability assessment · Genetic stability analysis | 20 | 96 | 16 | **2,20,000** |
| SBQ-04 | **Cell Substrate Identity & Purity Analysis**<br>Cell substrate identity and cross-contamination testing · Substrate identity testing | 16 | 76 | 13 | **1,75,000** |
| SBQ-05 | **Adventitious Virus Detection by Sequencing**<br>Sequencing-based adventitious virus testing · Viral safety sequencing | 23 | 112 | 19 | **2,55,000** |
| SBQ-06 | **Bacterial & Mycoplasma Detection by Sequencing**<br>Sequencing-based microbial and mycoplasma testing · Microbial detection sequencing | 18 | 86 | 15 | **2,00,000** |
| SBQ-07 | **Endogenous Retroviral Element Analysis**<br>Endogenous retroviral element characterisation · ERV analysis | 18 | 88 | 15 | **2,05,000** |
| SBQ-08 | **Raw Material & Feedstock Screening Analysis**<br>Raw material adventitious agent screening · Raw material screening | 17 | 82 | 14 | **1,90,000** |
| SBQ-09 | **Transgene Integrity & Copy Number Analysis**<br>Transgene integration and copy number characterisation · Transgene characterisation | 19 | 92 | 16 | **2,10,000** |
| SBQ-10 | **Expression Construct Verification Analysis**<br>Expression construct sequence verification · Construct verification | 17 | 82 | 14 | **1,90,000** |
| SBQ-11 | **Producer Cell Line Clone Selection Analysis**<br>Clone selection genomic screening · Clone selection analysis | 20 | 96 | 16 | **2,20,000** |
| SBQ-12 | **Host Cell DNA Residual Analysis**<br>Residual host cell DNA characterisation · Residual DNA analysis | 18 | 86 | 15 | **2,00,000** |
| SBQ-13 | **Viral Vector Genome Integrity Analysis**<br>Vector genome integrity and impurity characterisation · Vector integrity QC | 22 | 106 | 18 | **2,45,000** |

Continued overleaf. Programme codes read SBQ–01 and protocol references PRT/SBQ/01.

# Analysis Types — *14 to 25*

| CODE | ANALYSIS TYPE · INDUSTRY NAME & SHORT NAME | STEPS | HOURS | DAYS | FEE (INR) |
|---|---|---|---|---|---|
| SBQ-14 | **Capsid & Packaging Composition Analysis**<br>Vector packaging composition characterisation · Packaging composition analysis | 20 | 98 | 17 | 2,25,000 |
| SBQ-15 | **Plasmid Identity & Quality Analysis**<br>Plasmid sequence identity and quality testing · Plasmid QC | 17 | 82 | 14 | 1,90,000 |
| SBQ-16 | **mRNA Sequence & Impurity Analysis**<br>mRNA product sequence and impurity characterisation · mRNA QC | 20 | 96 | 16 | 2,20,000 |
| SBQ-17 | **Cell Therapy Product Release Analysis**<br>Cell therapy genomic release testing · Cell therapy release testing | 22 | 108 | 18 | 2,45,000 |
| SBQ-18 | **Microbial Identification for Contamination Investigation**<br>Contamination isolate identification and source tracing · Contamination investigation | 18 | 86 | 15 | 2,00,000 |
| SBQ-19 | **Environmental & Utility Monitoring Metagenomics**<br>Facility environmental monitoring by sequencing · Environmental monitoring metagenomics | 17 | 82 | 14 | 1,90,000 |
| SBQ-20 | **Process Comparability Genomic Analysis**<br>Manufacturing comparability genomic assessment · Comparability analysis | 21 | 102 | 17 | 2,35,000 |
| SBQ-21 | **Stability Study Genomic Analysis**<br>Product stability genomic assessment · Stability genomic analysis | 16 | 78 | 13 | 1,80,000 |
| SBQ-22 | **Analytical Method Validation for Sequencing Assays**<br>Sequencing assay validation for GMP use · Method validation | 24 | 118 | 20 | 2,70,000 |
| SBQ-23 | **Data Integrity & Computerised System Compliance Analysis**<br>Data integrity assessment for computational systems · Data integrity compliance | 20 | 98 | 17 | 2,25,000 |
| SBQ-24 | **Deviation & Out-of-Specification Investigation Analysis**<br>Genomic investigation of manufacturing deviations · Deviation investigation | 19 | 92 | 16 | 2,10,000 |
| SBQ-25 | **Rapid Sterility Assurance by Sequencing Methods**<br>Rapid microbiological method implementation · Rapid sterility methods | 18 | 88 | 15 | 2,05,000 |
| **Complete Sector Catalogue 02 — all thirty-nine analysis types** | | **481** | **2328** | **397** | **53,60,000** |

Each fee is derived independently from the step count, analytical complexity and reporting burden of that specific analysis type. There is no banded pricing. Days are calculated at six contact hours per working day and rounded up. Fees are per participant and exclusive of GST and the one-time registration charge. In-house delivery at the client laboratory and institutional rates are available on application.

**SBQ-01**
SECTOR 01

# Master Cell Bank Genomic Characterisation

(**Master cell bank genomic characterisation** · MCB characterisation)

### WHY IT IS DONE

The master cell bank is the permanent origin of every batch a product will ever produce, so it is characterised once, exhaustively, and never again from the same material. Anything missed at this point propagates through the entire commercial life of the product.

### WHAT IT ANSWERS

What is the complete genomic identity, integrity and purity status of this master cell bank, and does it support use as the origin of an approved product?

### WHEN IT IS DONE

At master cell bank establishment before any further banking, and when a legacy bank requires re-characterisation to modern expectations.

### WHAT TRIGGERS IT

Establishment of a new master cell bank, a regulatory requirement to re-characterise a legacy bank, or a technology transfer requiring full characterisation.

### HOW IT IS DONE

Identity and species are established, the host genome and integrated construct are sequenced and characterised, and integration sites and copy number are determined. Adventitious agent screening runs against comprehensive references with controls processed identically, endogenous retroviral elements are characterised, and the complete profile is archived as the permanent baseline.

### WHAT IT PRODUCES

Identity and species confirmation, host genome characterisation, construct sequence verification, integration site and copy number determination, adventitious agent screening results with controls, endogenous element characterisation, and the archived baseline profile.

### WHAT IT DETECTS

Cell line identity and species, genomic sequence of the host and any integrated construct, transgene integration sites and copy number, adventitious agent sequence, endogenous retroviral element status, and the baseline genomic state against which all later banks are compared.

### WHAT IT DOES NOT DETECT

It does not establish productivity, product quality attributes or viability, which are separate release attributes. Nucleic acid detection of an agent does not establish infectivity, and genomic characterisation cannot anticipate every process-related change that will occur downstream.

### WHAT MAKES IT HARD

This characterisation is done once and referenced for decades, so anything not captured cannot be recovered without re-establishing the bank. Distinguishing genuine agent detection from ubiquitous reagent background requires controls designed in from the start rather than added after a positive appears.

### WHO DOES THIS JOB

Bioinformatician in biologics quality or cell bank characterisation, senior level, working with quality assurance and regulatory affairs.

### WHAT YOU CAN DO AFTERWARDS

Characterise a master cell bank to regulatory expectations, establish a permanent genomic baseline, and produce documentation that supports a marketing application.

### WHO HIRES FOR IT

Biologics manufacturers, cell and gene therapy companies, contract development and manufacturing organisations, and cell bank testing providers.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **22** | **106 hours · 18 days** | Prior genome analysis experience, GMP awareness | **2,45,000** |

**SBQ-01**
WORKFLOW

# Master Cell Bank Genomic Characterisation

(**Master cell bank genomic characterisation** · MCB characterisation)

## COMPLETE WORKFLOW — 22 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Characterisation protocol authorisation** | The characterisation protocol and its regulatory basis approved before work begins | Protocol approval | *Authorised protocol* |
| 02 | **Bank provenance documentation** | Cell line origin, derivation history and all prior handling documented | Provenance records | *Provenance record* |
| 03 | **Sample chain of custody establishment** | Custody from bank vial to analytical sample established and recorded | Custody procedure | *Custody record* |
| 04 | **Species identification** | Species confirmed independently of any claimed identity | Species analysis | *Species confirmation* |
| 05 | **Identity profile generation** | An identity profile generated and archived as the permanent reference | Genotyping | *Identity profile* |
| 06 | **Data receipt and quality assessment** | Sequence integrity and quality verified against acceptance criteria | md5sum, FastQC | *QC acceptance record* |
| 07 | **Host genome alignment and processing** | Reads aligned to the host reference and processed for analysis | bwa-mem2, Picard | *Analysis-ready BAM* |
| 08 | **Genome coverage assessment** | Coverage and uniformity assessed against characterisation requirements | mosdepth | *Coverage record* |
| 09 | **Host genome variant characterisation** | Genomic variation relative to the reference characterised as the baseline | Variant calling | *Host variant profile* |
| 10 | **Karyotypic and copy number assessment** | Chromosomal complement and copy number state characterised | Copy number analysis | *Karyotypic profile* |
| 11 | **Construct sequence verification** | Any integrated construct verified against its design sequence | Sequence comparison | *Construct verification* |
| 12 | **Integration site mapping** | Integration sites mapped to host coordinates with junctions recovered | Junction analysis | *Integration sites* |
| 13 | **Transgene copy number quantification** | Copy number quantified against host genome background | Copy number quantification | *Copy number result* |
| 14 | **Integration structure reconstruction** | Integration architecture reconstructed including concatemeric arrangements | Structure reconstruction | *Integration structure* |
| 15 | **Host depletion for agent screening** | Host sequence depleted with efficiency quantified before agent classification | Host depletion | *Non-host reads* |
| 16 | **Adventitious agent classification** | Non-host reads classified against comprehensive reference databases | Classification | *Agent classification* |
| 17 | **Control comparison** | Detections assessed against reagent and negative controls processed identically | Control analysis | *Control comparison* |
| 18 | **Spike-in sensitivity determination** | Achieved detection sensitivity established by spike-in across agent classes | Spike analysis | *Sensitivity determination* |
| 19 | **Endogenous retroviral element characterisation** | Endogenous elements characterised and distinguished from exogenous sequence | Element analysis | *Endogenous profile* |
| 20 | **Unassigned sequence characterisation** | Sequence matching no reference assembled and characterised | Assembly, characterisation | *Unassigned characterisation* |
| 21 | **Baseline profile archiving** | The complete characterisation archived as the permanent reference baseline | Archive procedure | *Archived baseline* |
| 22 | **Characterisation report assembly** | All results assembled into a regulatory characterisation dossier | Report assembly | *Characterisation dossier* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **22** | **106 h · 18 d** | **SBQ-01** | **2,45,000** |

**SBQ-02**
SECTOR 01

# Working Cell Bank & End-of-Production Cell Analysis

(**WCB and EOPC genomic testing** · WCB · EOPC analysis)

### WHY IT IS DONE

Cells change through expansion, and the population at the end of production has been through many more doublings than the bank it came from. Testing at both points brackets the manufacturing window and demonstrates that the process does not drive the cells outside their characterised state.

### WHAT IT ANSWERS

Do the working cell bank and the end-of-production cells remain consistent with the master bank, and has anything changed across the manufacturing window?

### WHEN IT IS DONE

At working cell bank establishment, at the limit of in vitro cell age during process characterisation, and at end of production for each qualified process.

### WHAT TRIGGERS IT

Working cell bank establishment, process characterisation studies, or a limit of in vitro cell age qualification requirement.

### HOW IT IS DONE

Both banks and end-of-production cells are analysed against the archived master bank baseline, identity concordance is confirmed, and genomic differences are identified and assessed. Transgene copy number and integrity are quantified at each point, adventitious agent screening is repeated, and changes are characterised across the manufacturing window.

### WHAT IT PRODUCES

Identity concordance against the master bank, genomic change assessment across expansion, transgene copy number and integrity at each point, adventitious agent screening results, and a manufacturing window stability statement.

### WHAT IT DETECTS

Identity concordance with the master bank, genomic changes accumulated through expansion, transgene copy number and integrity at the production limit, adventitious agent status at both points, and any population shift relative to the bank.

### WHAT IT DOES NOT DETECT

It cannot establish that changes observed are inconsequential for product quality, which requires product characterisation. Low-frequency subpopulations below detection may still expand under different process conditions.

### WHAT MAKES IT HARD

Deciding which observed changes matter is the difficulty, since some genomic drift is inevitable and not all of it is consequential. The comparison depends entirely on the master bank baseline having been characterised well enough to compare against.

### WHO DOES THIS JOB

Bioinformatician in biologics quality control, mid to senior level, working with process development.

### WHAT YOU CAN DO AFTERWARDS

Compare banks and production cells against a characterised baseline, assess genomic change across the manufacturing window, and support limit of in vitro cell age qualification.

### WHO HIRES FOR IT

Biologics manufacturers, contract manufacturing organisations, cell and gene therapy companies, and quality control testing laboratories.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **19** | **92 hours · 16 days** | Prior cell bank characterisation experience | **2,10,000** |

**SBQ-02**
WORKFLOW

# Working Cell Bank & End-of-Production Cell Analysis

(**WCB and EOPC genomic testing** · WCB · EOPC analysis)

## COMPLETE WORKFLOW — 19 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Testing protocol confirmation** | The applicable testing protocol and acceptance criteria confirmed | Protocol review | *Testing protocol* |
| 02 | **Sample provenance recording** | Bank or production stage, passage and custody recorded for each sample | Sample records | *Provenance record* |
| 03 | **Master bank baseline retrieval** | The archived master bank baseline retrieved for comparison | Archive retrieval | *Retrieved baseline* |
| 04 | **Data receipt and quality assessment** | Sequence integrity and quality verified per sample | md5sum, FastQC | *QC acceptance record* |
| 05 | **Alignment and processing** | Reads aligned and processed identically to the baseline characterisation | bwa-mem2, Picard | *Analysis-ready BAM* |
| 06 | **Identity profile generation** | Identity profile generated for concordance against the master bank | Genotyping | *Identity profile* |
| 07 | **Identity concordance assessment** | Concordance with the master bank baseline assessed against criteria | Concordance analysis | *Concordance determination* |
| 08 | **Coverage assessment** | Coverage assessed against the requirements of each comparison | mosdepth | *Coverage record* |
| 09 | **Genomic difference identification** | Differences from the baseline identified across variant and copy number | Comparative analysis | *Difference set* |
| 10 | **Difference significance assessment** | Identified differences assessed for likely consequence | Assessment framework | *Significance assessment* |
| 11 | **Transgene copy number quantification** | Copy number quantified and compared against the baseline | Copy number quantification | *Copy number comparison* |
| 12 | **Transgene integrity verification** | Construct sequence integrity verified at the production stage | Sequence comparison | *Integrity verification* |
| 13 | **Integration site verification** | Integration sites confirmed unchanged from the baseline | Junction analysis | *Site verification* |
| 14 | **Host depletion and agent screening** | Host depleted and adventitious agent screening performed | Depletion, classification | *Agent screening results* |
| 15 | **Control comparison** | Detections assessed against controls processed identically | Control analysis | *Control comparison* |
| 16 | **Population shift assessment** | Clonal composition compared against the baseline for population shift | Clonal analysis | *Population assessment* |
| 17 | **Manufacturing window assessment** | Findings assessed across the full window from bank to production limit | Window analysis | *Window assessment* |
| 18 | **Acceptance determination** | Results assessed against protocol acceptance criteria | Acceptance criteria | *Acceptance determination* |
| 19 | **Testing report assembly** | Concordance, differences and agent screening assembled into the testing report | Report assembly | *Testing report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **19** | **92 h · 16 d** | **SBQ-02** | **2,10,000** |

**SBQ-03**
SECTOR 01

# Cell Line Genetic Stability Analysis

(**Producer cell line genetic stability assessment** · Genetic stability analysis)

### WHY IT IS DONE

A production cell line that loses transgene copies, rearranges its integration site or accumulates chromosomal change over expansion will eventually produce a different product from the one that was approved. Stability data define how long a line can be used.

### WHAT IT ANSWERS

Is this cell line genetically stable across the intended manufacturing window, and where does instability begin?

### WHEN IT IS DONE

In process development to establish the limit of in vitro cell age, during process qualification, and when investigating a productivity or product quality drift.

### WHAT TRIGGERS IT

Establishment of a production cell line, a limit of in vitro cell age study, or an unexplained drift in productivity or product quality.

### HOW IT IS DONE

Cells are sampled at defined passage intervals across and beyond the intended window, transgene copy number is quantified at each, and integration sites are re-characterised to detect rearrangement. Chromosomal change is assessed, clonal composition is tracked to detect population shift, and the passage at which change appears is identified.

### WHAT IT PRODUCES

Copy number trajectory across passage, integration site integrity at each interval, chromosomal change assessment, clonal composition tracking, the detected onset of instability, and a limit of in vitro cell age recommendation.

### WHAT IT DETECTS

Transgene copy number change across passage, integration site rearrangement or loss, chromosomal-level change including aneuploidy, clonal population shift within the culture, and the passage at which any change becomes detectable.

### WHAT IT DOES NOT DETECT

It does not link genomic change to product quality consequence, which requires parallel product characterisation. Changes affecting only a minority subpopulation may be below detection until that population expands.

### WHAT MAKES IT HARD

Population-level measurement conceals subclonal change until it dominates, so tracking clonal composition rather than only average copy number is what detects instability early. Sampling intervals must extend beyond the intended window or the study cannot demonstrate a margin.

### WHO DOES THIS JOB

Bioinformatician in process development or quality, mid to senior level.

### WHAT YOU CAN DO AFTERWARDS

Design and analyse genetic stability studies, track clonal composition rather than averages, and derive defensible limits of in vitro cell age.

### WHO HIRES FOR IT

Biologics manufacturers, contract development organisations, cell line development companies, and process development functions.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **20** | **96 hours · 16 days** | Prior copy number and structural variant experience | **2,20,000** |

**SBQ-03**
WORKFLOW

# Cell Line Genetic Stability Analysis

(**Producer cell line genetic stability assessment** · Genetic stability analysis)

## COMPLETE WORKFLOW — 20 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Study design authorisation** | Passage intervals, endpoints and the intended window approved before study start | Protocol approval | *Authorised design* |
| 02 | **Passage interval definition** | Sampling intervals defined to extend beyond the intended manufacturing window | Design specification | *Sampling schedule* |
| 03 | **Culture and sampling record capture** | Culture conditions and sampling events recorded for every timepoint | Culture records | *Culture record* |
| 04 | **Baseline characterisation** | The starting population characterised as the comparison baseline | Baseline analysis | *Baseline profile* |
| 05 | **Data receipt and quality assessment** | Sequence integrity and quality verified per timepoint | md5sum, FastQC | *QC acceptance record* |
| 06 | **Alignment and processing** | All timepoints aligned and processed identically | bwa-mem2, Picard | *Analysis-ready BAMs* |
| 07 | **Coverage consistency assessment** | Coverage confirmed consistent across timepoints for valid comparison | mosdepth | *Coverage consistency* |
| 08 | **Transgene copy number quantification** | Copy number quantified at every passage interval | Copy number quantification | *Copy number trajectory* |
| 09 | **Copy number trend modelling** | Copy number modelled as a trend across passage rather than pairwise | Trend modelling | *Copy number model* |
| 10 | **Integration junction verification** | Integration junctions re-verified at each interval for rearrangement | Junction analysis | *Junction verification* |
| 11 | **Integration structure comparison** | Integration architecture compared across intervals | Structure comparison | *Structure comparison* |
| 12 | **Transgene sequence integrity assessment** | Construct sequence integrity assessed at each interval | Sequence comparison | *Integrity assessment* |
| 13 | **Chromosomal change assessment** | Chromosomal complement and copy number change assessed across passage | Copy number analysis | *Chromosomal assessment* |
| 14 | **Clonal composition analysis** | Clonal composition characterised at each interval | Clonal analysis | *Clonal composition* |
| 15 | **Population shift detection** | Shifts in dominant clonal populations detected across the study | Shift analysis | *Population shift* |
| 16 | **Subclonal emergence assessment** | Emerging subpopulations detected before they dominate | Subclonal analysis | *Emergence assessment* |
| 17 | **Instability onset determination** | The passage at which any change becomes detectable determined | Onset analysis | *Onset determination* |
| 18 | **Margin assessment** | Margin between instability onset and the intended window assessed | Margin analysis | *Margin assessment* |
| 19 | **Limit of in vitro cell age recommendation** | A defensible limit recommended from the observed trajectories | Derivation framework | *Limit recommendation* |
| 20 | **Stability report assembly** | Trajectories, onset and limit recommendation assembled | Report assembly | *Stability report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **20** | **96 h · 16 d** | **SBQ-03** | **2,20,000** |

**SBQ-04**
SECTOR 01

# Cell Substrate Identity & Purity Analysis

(**Cell substrate identity and cross-contamination testing** · Substrate identity testing)

### WHY IT IS DONE

A manufacturing facility running multiple cell lines carries a permanent risk of cross-contamination between them, and a contaminated substrate produces a product that is not what the marketing authorisation describes. Identity and purity testing is the control against that.

### WHAT IT ANSWERS

Is this cell substrate the declared line, is it free of other cell lines, and does it match the reference profile held for it?

### WHEN IT IS DONE

On receipt of any substrate, at bank establishment, at defined intervals during campaigns, and in investigating a suspected cross-contamination event.

### WHAT TRIGGERS IT

Substrate receipt, bank establishment, a scheduled interval, or a contamination investigation.

### HOW IT IS DONE

An identity profile is generated and compared against the archived reference for that substrate, species is confirmed independently, and mixed allele signal is analysed for evidence of a second line. Contamination proportion is estimated where present, and the facility's other lines are compared as candidate sources.

### WHAT IT PRODUCES

Identity profile with reference concordance, species confirmation, cross-contamination assessment with proportion, candidate source identification from facility lines, and a purity determination.

### WHAT IT DETECTS

Identity profile against the reference, species confirmation, cross-contamination from other cell lines with proportion estimated, and inter-species contamination.

### WHAT IT DOES NOT DETECT

It cannot detect contamination by a line with an identical profile, and it does not assess microbiological contamination, which requires separate methods. Very low-level cross-contamination falls below the discrimination limit of any marker panel.

### WHAT MAKES IT HARD

Identifying the source line is what converts a detection into a corrective action, and that requires reference profiles for every line the facility handles. Low-level contamination sits near the discrimination limit and must be reported with its detection bound.

### WHO DOES THIS JOB

Bioinformatician or quality control analyst, entry to mid level, within a quality function.

### WHAT YOU CAN DO AFTERWARDS

Perform substrate identity and purity testing, estimate contamination proportion, and identify candidate source lines within a facility.

### WHO HIRES FOR IT

Biologics manufacturers, contract manufacturing organisations, cell banking services, and quality control laboratories.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **16** | **76 hours · 13 days** | Prior genotype analysis experience | **1,75,000** |

**SBQ-04**
WORKFLOW

# Cell Substrate Identity & Purity Analysis

(**Cell substrate identity and cross-contamination testing** · Substrate identity testing)

## COMPLETE WORKFLOW — 16 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Testing scope confirmation** | The substrate, testing occasion and applicable criteria confirmed | Protocol review | *Testing scope* |
| 02 | **Reference profile retrieval** | The archived reference profile for the substrate retrieved | Archive retrieval | *Reference profile* |
| 03 | **Facility line inventory retrieval** | Profiles of all other lines handled in the facility retrieved for comparison | Archive retrieval | *Facility profile set* |
| 04 | **Sample custody verification** | Custody from culture to analytical sample verified | Custody procedure | *Custody record* |
| 05 | **Data receipt and quality assessment** | Data quality verified against acceptance criteria | Quality assessment | *QC acceptance record* |
| 06 | **Identity profile generation** | The identity profile generated across the marker panel | Genotyping | *Identity profile* |
| 07 | **Profile completeness assessment** | Profile completeness and call quality assessed | Quality analysis | *Profile quality* |
| 08 | **Reference concordance assessment** | Profile compared against the archived reference with thresholds applied | Concordance analysis | *Concordance determination* |
| 09 | **Species confirmation** | Species confirmed independently of the identity match | Species analysis | *Species confirmation* |
| 10 | **Mixed allele analysis** | Additional alleles indicating a second cell line identified | Mixture analysis | *Mixture findings* |
| 11 | **Contamination proportion estimation** | Where mixed, the proportion of the contaminant estimated with bounds | Proportion estimation | *Contamination estimate* |
| 12 | **Facility line comparison** | Mixed signal compared against facility line profiles to identify the source | Comparative analysis | *Source identification* |
| 13 | **Detection limit statement** | The contamination detection limit of the marker panel stated | Limit calculation | *Detection limit* |
| 14 | **Inter-species contamination assessment** | Contamination from other species assessed separately | Species analysis | *Inter-species assessment* |
| 15 | **Purity determination** | A purity determination made against the applicable criteria | Determination criteria | *Purity determination* |
| 16 | **Testing report assembly** | Identity, purity, source and detection limit assembled | Report assembly | *Identity and purity report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **16** | **76 h · 13 d** | **SBQ-04** | **1,75,000** |

**SBQ-05**
SECTOR 01

# Adventitious Virus Detection by Sequencing

(**Sequencing-based adventitious virus testing** · Viral safety sequencing)

### WHY IT IS DONE

Conventional viral safety testing detects only what its assays are designed for, and historical contamination events involved viruses those assays could not see. Sequencing detects any viral sequence present, which is why regulators now accept it as a broad complement to targeted testing.

### WHAT IT ANSWERS

Does this material contain any viral sequence, of what identity, at what abundance, and does the evidence support release?

### WHEN IT IS DONE

In cell bank and end-of-production testing, in bulk harvest testing, in raw material qualification, and in investigating a suspected contamination event.

### WHAT TRIGGERS IT

Cell bank or harvest testing, raw material qualification, a suspect observation in cell culture, or a contamination investigation.

### HOW IT IS DONE

Host sequence is depleted with efficiency quantified because it sets the effective sensitivity, non-host reads are classified against comprehensive viral references, and every detection is assessed against reagent and negative controls processed identically. Spike-in controls establish achieved sensitivity, detections are confirmed by targeted analysis, and unassigned sequence is assembled and characterised.

### WHAT IT PRODUCES

Host depletion efficiency, viral detections with identity and abundance, control comparison establishing background, spike-in derived sensitivity per agent class, targeted confirmation of detections, unassigned sequence characterisation, and a viral safety assessment.

### WHAT IT DETECTS

Viral sequence from any family present above the detection threshold, species and strain identification where references exist, viral abundance relative to host, integrated and episomal viral sequence, and novel viral sequence flagged as unassigned.

### WHAT IT DOES NOT DETECT

It detects nucleic acid rather than infectious virus, so significance requires follow-up. Viruses absent from reference databases can only be flagged as unassigned sequence, and detection sensitivity depends entirely on abundance relative to the overwhelming host background.

### WHAT MAKES IT HARD

Sensitivity must be demonstrated rather than claimed, and it varies by an order of magnitude between agents depending on genome size and host background. Reagent-borne viral sequence is ubiquitous, so controls processed in parallel are what separate a finding from laboratory background.

### WHO DOES THIS JOB

Bioinformatician in viral safety or biosafety testing, senior level, working with virology and quality assurance.

### WHAT YOU CAN DO AFTERWARDS

Run sequencing-based viral safety testing with demonstrated sensitivity, distinguish findings from reagent background, and produce regulatory-grade safety assessments.

### WHO HIRES FOR IT

Biologics manufacturers, viral safety testing laboratories, cell and gene therapy companies, and contract testing organisations.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **23** | **112 hours · 19 days** | Prior metagenomic analysis experience | **2,55,000** |

**SBQ-05**
WORKFLOW

# Adventitious Virus Detection by Sequencing

(**Sequencing-based adventitious virus testing** · Viral safety sequencing)

## COMPLETE WORKFLOW — 23 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Testing protocol authorisation** | Scope, agent classes and acceptance criteria approved before testing | Protocol approval | *Authorised protocol* |
| 02 | **Reference database assembly** | Comprehensive viral reference databases assembled with versions recorded | Database assembly | *Reference databases* |
| 03 | **Control design** | Negative, reagent and spike-in controls designed into the run | Control design | *Control plan* |
| 04 | **Spike-in panel preparation** | Spike organisms spanning genome types prepared at defined levels | Spike preparation | *Spike controls* |
| 05 | **Sample and control extraction** | Samples and controls extracted in parallel under identical conditions | Extraction | *Prepared material* |
| 06 | **Nucleic acid type coverage confirmation** | Both DNA and RNA workflows confirmed where the scope requires both | Workflow verification | *Coverage confirmation* |
| 07 | **Data receipt and quality assessment** | Sequence integrity and run quality verified per sample and control | md5sum, FastQC | *QC acceptance record* |
| 08 | **Host sequence depletion** | Host reads removed with the method recorded | Host depletion | *Non-host reads* |
| 09 | **Depletion efficiency quantification** | Depletion efficiency quantified as it determines effective sensitivity | Efficiency analysis | *Depletion metrics* |
| 10 | **Viral classification** | Non-host reads classified against the viral reference databases | Classification | *Viral classification* |
| 11 | **Read distribution assessment** | Read distribution across each detected genome assessed for authenticity | Coverage analysis | *Distribution assessment* |
| 12 | **Reagent background subtraction** | Taxa present in reagent controls treated as background | Background analysis | *Background-corrected results* |
| 13 | **Negative control comparison** | Sample detections compared against parallel negative controls | Control comparison | *Control comparison record* |
| 14 | **Spike-in recovery quantification** | Spike organism recovery quantified per genome type | Recovery analysis | *Recovery metrics* |
| 15 | **Per-class sensitivity determination** | Detection sensitivity determined per viral genome class | Sensitivity calculation | *Per-class sensitivity* |
| 16 | **Detection threshold application** | Detections assessed against the reporting threshold | Threshold criteria | *Threshold decisions* |
| 17 | **Targeted confirmation** | Detections confirmed by targeted analysis or genome assembly | Confirmation analysis | *Confirmation results* |
| 18 | **Integrated viral sequence assessment** | Viral sequence integrated into the host genome identified separately | Integration analysis | *Integration findings* |
| 19 | **Endogenous element discrimination** | Endogenous retroviral sequence distinguished from exogenous virus | Discrimination analysis | *Endogenous discrimination* |
| 20 | **Unassigned sequence assembly** | Unclassified sequence assembled and characterised | Assembly, characterisation | *Unassigned characterisation* |
| 21 | **Viability limitation documentation** | The distinction between nucleic acid and infectious virus documented | Limitation template | *Viability caveat* |
| 22 | **Risk assessment** | Detections assessed for significance against the material's intended use | Risk assessment | *Risk determination* |
| 23 | **Viral safety report assembly** | Detections, sensitivity, controls and risk assembled into a safety report | Report assembly | *Viral safety report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **23** | **112 h · 19 d** | **SBQ-05** | **2,55,000** |

**SBQ-06**
SECTOR 01

# Bacterial & Mycoplasma Detection by Sequencing

(**Sequencing-based microbial and mycoplasma testing** · Microbial detection sequencing)

### WHY IT IS DONE

Mycoplasma contamination is common, invisible in culture and alters cell behaviour profoundly. Sequencing detects it and other bacterial contaminants without culture, in days rather than the weeks compendial methods require, which changes what a manufacturer can do about it.

### WHAT IT ANSWERS

Does this material contain bacterial or mycoplasmal sequence, of what identity, and does it represent contamination or reagent background?

### WHEN IT IS DONE

In routine cell culture monitoring, in bank and harvest testing, in raw material qualification, and in investigating a suspected contamination.

### WHAT TRIGGERS IT

Routine monitoring, bank or harvest testing, an anomalous culture observation, or a contamination investigation.

### HOW IT IS DONE

Host sequence is depleted, non-host reads are classified against curated bacterial and mycoplasmal references, and detections are assessed against reagent and negative controls run in parallel. Species assignment is confirmed by marker gene analysis or assembly, abundance is quantified, and spike-in controls establish sensitivity.

### WHAT IT PRODUCES

Classification results with species assignment, abundance quantification, control comparison establishing reagent background, spike-in derived sensitivity, confirmation analysis for detections, and a contamination determination.

### WHAT IT DETECTS

Bacterial and mycoplasmal sequence with species-level identification where references permit, relative abundance, and the distinction between sample contamination and reagent-derived background.

### WHAT IT DOES NOT DETECT

It detects nucleic acid rather than viable organism, so a positive does not establish live contamination and a negative does not establish sterility. Species absent from references are identifiable only to higher taxonomic rank.

### WHAT MAKES IT HARD

Reagent and environmental bacterial sequence is present in every preparation, so distinguishing genuine contamination from background is the entire analytical problem and requires controls rather than thresholds alone. Low-biomass samples are dominated by background.

### WHO DOES THIS JOB

Bioinformatician in quality control or biosafety, mid level.

### WHAT YOU CAN DO AFTERWARDS

Detect and identify microbial contamination by sequencing, distinguish it from reagent background, and demonstrate sensitivity by spike-in.

### WHO HIRES FOR IT

Biologics manufacturers, quality control laboratories, contract testing organisations, and cell therapy manufacturers.

### COMMERCIALS

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

**SBQ-06**
WORKFLOW

# Bacterial & Mycoplasma Detection by Sequencing

(**Sequencing-based microbial and mycoplasma testing** · Microbial detection sequencing)

## COMPLETE WORKFLOW — 18 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Testing scope confirmation** | Material type, agent scope and acceptance criteria confirmed | Protocol review | *Testing scope* |
| 02 | **Reference database assembly** | Bacterial and mycoplasmal references assembled with versions recorded | Database assembly | *Reference databases* |
| 03 | **Control design and preparation** | Negative, reagent and spike controls designed and prepared | Control preparation | *Control set* |
| 04 | **Sample and control extraction** | Samples and controls extracted in parallel under identical conditions | Extraction | *Prepared material* |
| 05 | **Data receipt and quality assessment** | Sequence integrity and quality verified per sample and control | md5sum, FastQC | *QC acceptance record* |
| 06 | **Host sequence depletion** | Host reads removed with efficiency quantified | Host depletion | *Non-host reads* |
| 07 | **Taxonomic classification** | Non-host reads classified against bacterial and mycoplasmal references | Classification | *Classification results* |
| 08 | **Marker gene analysis** | Detections confirmed and refined by marker gene analysis | Marker analysis | *Marker confirmation* |
| 09 | **Reagent background subtraction** | Taxa present in reagent controls treated as background | Background analysis | *Corrected results* |
| 10 | **Negative control comparison** | Detections compared against parallel negative controls | Control comparison | *Control comparison record* |
| 11 | **Spike-in recovery assessment** | Spike organism recovery quantified to establish sensitivity | Recovery analysis | *Sensitivity determination* |
| 12 | **Abundance quantification** | Detected organism abundance quantified relative to host and total | Quantification | *Abundance results* |
| 13 | **Species assignment confidence** | Assignment confidence assessed with higher-rank assignment where uncertain | Confidence assessment | *Assignment confidence* |
| 14 | **Threshold application** | Detections assessed against the reporting abundance threshold | Threshold criteria | *Threshold decisions* |
| 15 | **Genome assembly for confirmation** | Detected organisms assembled where abundance permits confirmation | Assembly | *Assembly confirmation* |
| 16 | **Viability limitation documentation** | The nucleic acid versus viable organism distinction documented | Limitation template | *Viability caveat* |
| 17 | **Contamination determination** | A contamination determination made against acceptance criteria | Determination criteria | *Contamination determination* |
| 18 | **Testing report assembly** | Detections, sensitivity, controls and determination assembled | Report assembly | *Microbial testing report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **18** | **86 h · 15 d** | **SBQ-06** | **2,00,000** |

**SBQ-07**
SECTOR 01

# Endogenous Retroviral Element Analysis

(**Endogenous retroviral element characterisation** · ERV analysis)

### WHY IT IS DONE

Rodent-derived production cell lines carry endogenous retroviral sequences that are expressed and can produce particles. Characterising them and demonstrating that downstream processing clears them is a defined requirement of viral safety for these products.

### WHAT IT ANSWERS

What endogenous retroviral elements does this cell substrate carry, are they expressed, and what is their contribution to the particle burden the process must clear?

### WHEN IT IS DONE

In master cell bank characterisation, in process viral clearance justification, and in supporting the viral safety section of a regulatory submission.

### WHAT TRIGGERS IT

Master cell bank characterisation, a viral safety assessment, or a regulatory question on endogenous element burden.

### HOW IT IS DONE

Host genome sequence is analysed for retroviral element copies with positions mapped, element structure is assessed for intact reading frames, and expression is quantified from transcriptomic data. Elements are distinguished from any exogenous retroviral sequence, and burden is quantified for viral clearance justification.

### WHAT IT PRODUCES

Element copy inventory with genomic positions, structural integrity assessment per element, expression quantification, exogenous sequence discrimination, and burden quantification supporting clearance justification.

### WHAT IT DETECTS

Endogenous retroviral element copies in the host genome with their integration positions, transcriptional expression of those elements, evidence of intact open reading frames, and the distinction between endogenous elements and exogenous infection.

### WHAT IT DOES NOT DETECT

It does not establish particle infectivity, which requires biological assays. Whether an expressed element produces a competent particle cannot be determined from sequence and expression alone.

### WHAT MAKES IT HARD

Retroviral elements are repetitive and highly similar, so mapping copies individually rather than as an aggregate requires careful handling of multi-mapping reads. Distinguishing an intact expressed element from a degenerate one determines whether it matters at all.

### WHO DOES THIS JOB

Bioinformatician in viral safety, senior level, working with virology and regulatory affairs.

### WHAT YOU CAN DO AFTERWARDS

Characterise endogenous retroviral elements and their expression, discriminate them from exogenous sequence, and support viral clearance justification.

### WHO HIRES FOR IT

Biologics manufacturers using rodent cell lines, viral safety functions, contract testing organisations, and regulatory consultancies.

### COMMERCIALS

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

**SBQ-07**
WORKFLOW

# Endogenous Retroviral Element Analysis

(**Endogenous retroviral element characterisation** · ERV analysis)

## COMPLETE WORKFLOW — 18 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Assessment scope confirmation** | Element classes in scope and regulatory expectations confirmed | Protocol review | *Assessment scope* |
| 02 | **Host genome reference preparation** | Species-appropriate reference with element annotation prepared | Reference preparation | *Prepared reference* |
| 03 | **Element reference assembly** | Retroviral element reference sequences assembled for classification | Reference assembly | *Element references* |
| 04 | **Data receipt and quality assessment** | Sequence integrity and quality verified | md5sum, FastQC | *QC acceptance record* |
| 05 | **Genome alignment** | Reads aligned with multi-mapping handled explicitly given element similarity | bwa-mem2 | *Aligned BAM* |
| 06 | **Multi-mapping resolution** | Reads mapping to multiple element copies resolved or assigned probabilistically | Multi-mapping methods | *Resolved assignments* |
| 07 | **Element copy identification** | Individual element copies identified with genomic positions | Copy identification | *Element inventory* |
| 08 | **Element structure assessment** | Each copy assessed for structural completeness and reading frame integrity | Structure analysis | *Structural assessment* |
| 09 | **Intact element identification** | Copies with intact reading frames identified as those of potential concern | Reading frame analysis | *Intact element set* |
| 10 | **Expression data alignment** | Transcriptomic data aligned with element-aware settings | STAR, element-aware alignment | *Aligned RNA data* |
| 11 | **Element expression quantification** | Expression quantified per element copy where mapping permits | Expression quantification | *Expression profile* |
| 12 | **Aggregate expression assessment** | Total element expression quantified where copies cannot be resolved | Aggregate quantification | *Aggregate expression* |
| 13 | **Exogenous sequence discrimination** | Exogenous retroviral sequence distinguished from endogenous elements | Discrimination analysis | *Exogenous assessment* |
| 14 | **Particle-forming potential assessment** | Structural and expression evidence combined to assess particle potential | Combined assessment | *Potential assessment* |
| 15 | **Burden quantification** | Element burden quantified for viral clearance justification | Burden calculation | *Burden quantification* |
| 16 | **Infectivity limitation documentation** | The inability of sequence data to establish infectivity documented | Limitation template | *Infectivity caveat* |
| 17 | **Clearance justification support** | Burden data assembled to support viral clearance validation | Support assembly | *Clearance support data* |
| 18 | **Assessment report assembly** | Inventory, structure, expression and burden assembled | Report assembly | *Element assessment report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **18** | **88 h · 15 d** | **SBQ-07** | **2,05,000** |

**SBQ-08**
SECTOR 01

# Raw Material & Feedstock Screening Analysis

(**Raw material adventitious agent screening** · Raw material screening)

### WHY IT IS DONE

Animal-derived and complex biological raw materials are the historical route by which adventitious agents entered manufacturing. Screening them before use is cheaper and safer than detecting the consequence in a bulk harvest.

### WHAT IT ANSWERS

Does this raw material lot contain adventitious agent sequence, and does it qualify for use in manufacturing?

### WHEN IT IS DONE

In qualification of new raw material suppliers, in lot-by-lot screening of high-risk materials, and in investigating a contamination traced to materials.

### WHAT TRIGGERS IT

New supplier or material qualification, receipt of a high-risk material lot, or a contamination investigation implicating raw materials.

### HOW IT IS DONE

Material matrix is characterised because it determines extraction and background, nucleic acid is extracted with controls processed identically, and reads are classified against comprehensive references. Source species is verified for animal-derived materials, detections are assessed against controls, and sensitivity is established by spike-in in the same matrix.

### WHAT IT PRODUCES

Matrix characterisation, classification results with species assignment, source species verification, control comparison, matrix-specific spike-in sensitivity, and a material qualification recommendation.

### WHAT IT DETECTS

Viral, bacterial, fungal and mycoplasmal sequence in the material, species identification where references permit, source species verification for animal-derived materials, and abundance relative to the material matrix.

### WHAT IT DOES NOT DETECT

It detects nucleic acid rather than viable agent, and inactivated material may retain detectable sequence from a treated agent. Highly processed materials may carry degraded sequence that is undetectable while the original agent was present.

### WHAT MAKES IT HARD

Sensitivity is matrix-dependent and cannot be transferred between material types, so spike-in must be performed in the actual matrix being tested. Source species verification frequently reveals undeclared substitution in animal-derived materials.

### WHO DOES THIS JOB

Bioinformatician in raw material quality or biosafety, mid level.

### WHAT YOU CAN DO AFTERWARDS

Screen raw materials for adventitious agents with matrix-appropriate sensitivity, verify source species, and support supplier qualification.

### WHO HIRES FOR IT

Biologics manufacturers, raw material suppliers, contract testing organisations, and supply chain quality functions.

### COMMERCIALS

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

**SBQ-08**
WORKFLOW

# Raw Material & Feedstock Screening Analysis

(**Raw material adventitious agent screening** · Raw material screening)

## COMPLETE WORKFLOW — 17 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Material scope and risk assessment** | Material type, source and adventitious agent risk assessed | Risk assessment | *Material risk profile* |
| 02 | **Testing protocol confirmation** | Applicable testing scope and acceptance criteria confirmed | Protocol review | *Testing protocol* |
| 03 | **Matrix characterisation** | The material matrix characterised as it determines extraction and background | Matrix analysis | *Matrix characterisation* |
| 04 | **Extraction method selection** | Extraction selected and qualified for the specific material matrix | Method selection | *Selected extraction* |
| 05 | **Matrix-specific spike preparation** | Spike controls prepared in the actual material matrix | Spike preparation | *Matrix spike controls* |
| 06 | **Sample and control extraction** | Material and controls extracted in parallel under identical conditions | Extraction | *Prepared material* |
| 07 | **Data receipt and quality assessment** | Sequence integrity and quality verified per sample and control | md5sum, FastQC | *QC acceptance record* |
| 08 | **Source species verification** | Declared source species verified from the sequence data | Species analysis | *Species verification* |
| 09 | **Species substitution assessment** | Undeclared species substitution assessed for animal-derived materials | Substitution analysis | *Substitution assessment* |
| 10 | **Host and matrix depletion** | Expected source sequence depleted before agent classification | Depletion | *Depleted reads* |
| 11 | **Agent classification** | Remaining reads classified against comprehensive references | Classification | *Agent classification* |
| 12 | **Reagent background subtraction** | Detections assessed against reagent controls processed identically | Background analysis | *Corrected results* |
| 13 | **Matrix-specific sensitivity determination** | Sensitivity determined from spike recovery in the actual matrix | Recovery analysis | *Matrix sensitivity* |
| 14 | **Detection threshold application** | Detections assessed against the reporting threshold | Threshold criteria | *Threshold decisions* |
| 15 | **Targeted confirmation** | Detections confirmed by targeted analysis where above threshold | Confirmation analysis | *Confirmation results* |
| 16 | **Qualification determination** | A material qualification determination made against criteria | Determination criteria | *Qualification determination* |
| 17 | **Screening report assembly** | Species verification, detections and sensitivity assembled | Report assembly | *Screening report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **17** | **82 h · 14 d** | **SBQ-08** | **1,90,000** |

**SBQ-09**
SECTOR 01

# Transgene Integrity & Copy Number Analysis

(**Transgene integration and copy number characterisation** · Transgene characterisation)

### WHY IT IS DONE

The transgene is the productive element of a recombinant cell line, and its copy number, integration structure and sequence integrity determine both productivity and product consistency. Characterising it is a defined expectation for cell substrate documentation.

### WHAT IT ANSWERS

How many transgene copies does this cell line carry, where are they integrated, and is the transgene sequence intact and unrearranged?

### WHEN IT IS DONE

In master cell bank characterisation, in clone selection during cell line development, in stability studies, and in investigating productivity change.

### WHAT TRIGGERS IT

Cell bank characterisation, clone selection, a stability study timepoint, or an unexplained productivity change.

### HOW IT IS DONE

Copy number is quantified from coverage against the host background, integration junctions are recovered and mapped to host coordinates, and integration structure is reconstructed including tandem arrangements. Transgene sequence is compared against the intended construct across its full length, and flanking context is characterised.

### WHAT IT PRODUCES

Copy number quantification, integration site coordinates with host gene context, integration structure reconstruction, full-length construct sequence verification, and a characterisation record for the substrate documentation.

### WHAT IT DETECTS

Transgene copy number, integration site positions in the host genome, integration structure including concatemers and rearrangements, transgene sequence integrity against the construct, and flanking host sequence context.

### WHAT IT DOES NOT DETECT

It does not measure expression level or productivity, which are separate attributes. Copy number alone does not predict productivity, since integration site context frequently matters more than copy count.

### WHAT MAKES IT HARD

Tandem concatemeric integration is common and is easily reported as multiple independent sites unless junction structure is reconstructed. Copy number from coverage requires careful normalisation because the transgene is present at a very different abundance from the host genome.

### WHO DOES THIS JOB

Bioinformatician in cell line development or quality, mid to senior level.

### WHAT YOU CAN DO AFTERWARDS

Quantify transgene copy number, map and reconstruct integration structure, and verify construct sequence integrity.

### WHO HIRES FOR IT

Biologics manufacturers, cell line development companies, contract development organisations, and quality functions.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **19** | **92 hours · 16 days** | Prior structural variant and coverage analysis experience | **2,10,000** |

**SBQ-09**
WORKFLOW

# Transgene Integrity & Copy Number Analysis

(**Transgene integration and copy number characterisation** · Transgene characterisation)

## COMPLETE WORKFLOW — 19 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Characterisation scope confirmation** | Attributes to be characterised and their acceptance criteria confirmed | Protocol review | *Characterisation scope* |
| 02 | **Construct reference registration** | The intended construct sequence registered as the analytical reference | Sequence registry | *Registered construct* |
| 03 | **Host reference preparation** | Host genome reference prepared and version locked | Reference preparation | *Prepared reference* |
| 04 | **Data receipt and quality assessment** | Sequence integrity and quality verified against criteria | md5sum, FastQC | *QC acceptance record* |
| 05 | **Combined reference alignment** | Reads aligned against host and construct together | bwa-mem2, combined reference | *Aligned BAM* |
| 06 | **Coverage normalisation** | Construct coverage normalised against host genome background | Normalisation | *Normalised coverage* |
| 07 | **Copy number quantification** | Transgene copy number quantified with confidence bounds | Copy number quantification | *Copy number result* |
| 08 | **Junction recovery** | Vector to host junctions recovered from chimeric read evidence | Junction analysis | *Junction set* |
| 09 | **Junction read-level review** | Every candidate junction reviewed at read level to exclude artefact | IGV review | *Reviewed junctions* |
| 10 | **Integration site mapping** | Junctions mapped to host genomic coordinates | Coordinate mapping | *Integration sites* |
| 11 | **Site count determination** | The number of independent integration sites determined | Site analysis | *Site count* |
| 12 | **Concatemer detection** | Tandem concatemeric arrangements detected and distinguished from separate sites | Structure analysis | *Concatemer findings* |
| 13 | **Integration structure reconstruction** | Full integration architecture reconstructed per site | Structure reconstruction | *Integration structure* |
| 14 | **Construct sequence verification** | Integrated construct sequence verified across its full length | Sequence comparison | *Sequence verification* |
| 15 | **Rearrangement and truncation assessment** | Rearranged or truncated construct copies identified | Structural analysis | *Rearrangement findings* |
| 16 | **Flanking context annotation** | Host sequence and genes flanking each site annotated | Annotation | *Flanking context* |
| 17 | **Site stability risk assessment** | Integration site context assessed for stability risk | Risk assessment | *Site risk assessment* |
| 18 | **Baseline archiving** | The characterisation archived as the baseline for stability comparison | Archive procedure | *Archived baseline* |
| 19 | **Characterisation report assembly** | Copy number, sites, structure and sequence verification assembled | Report assembly | *Transgene characterisation report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **19** | **92 h · 16 d** | **SBQ-09** | **2,10,000** |

**SBQ-10**
SECTOR 01

# Expression Construct Verification Analysis

(**Expression construct sequence verification** · Construct verification)

### WHY IT IS DONE

Every recombinant product traces back to a construct sequence, and an undetected error in that sequence propagates into the product itself. Verification confirms that what was built matches what was designed before it becomes the basis of a manufacturing process.

### WHAT IT ANSWERS

Does this expression construct match its intended design sequence exactly, and is every functional element present and correctly positioned?

### WHEN IT IS DONE

At construct completion before cell line development, at plasmid bank establishment, and when a construct is transferred between sites or organisations.

### WHAT TRIGGERS IT

Construct completion, plasmid bank establishment, a technology transfer, or a discrepancy investigation.

### HOW IT IS DONE

The design sequence is registered as the reference, the construct is sequenced with a method resolving the full length including repetitive elements, and the assembled sequence is compared base by base against the design. Functional elements are verified for presence and boundaries, and any discrepancy is characterised and assessed for consequence.

### WHAT IT PRODUCES

Full construct sequence assembly, base-level comparison against the design, functional element verification, discrepancy characterisation with consequence assessment, and a verification record for the construct history.

### WHAT IT DETECTS

Full construct sequence with comparison against the design, sequence variants and their positions, functional element presence and boundaries, backbone integrity, and unintended sequence such as residual cloning scars.

### WHAT IT DOES NOT DETECT

It does not establish that the construct functions as intended, which requires expression testing. Structural features such as supercoiling and methylation state are outside its scope.

### WHAT MAKES IT HARD

Repetitive elements such as tandem promoters and terminators defeat short-read assembly and are exactly where errors occur, so the sequencing method must resolve them. Discrepancies require consequence assessment rather than a pass-fail decision, since many are silent.

### WHO DOES THIS JOB

Bioinformatician in molecular biology or quality, mid level.

### WHAT YOU CAN DO AFTERWARDS

Verify construct sequences against design at base resolution, resolve repetitive elements, and assess the consequence of discrepancies.

### WHO HIRES FOR IT

Biologics and cell line development functions, gene therapy developers, plasmid manufacturers, and contract development organisations.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **17** | **82 hours · 14 days** | Prior sequence analysis and assembly experience | **1,90,000** |

**SBQ-10**
WORKFLOW

# Expression Construct Verification Analysis

(**Expression construct sequence verification** · Construct verification)

## COMPLETE WORKFLOW — 17 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Design sequence registration** | The intended design sequence registered as the verification reference | Sequence registry | *Registered design* |
| 02 | **Functional element annotation** | Functional elements and their expected boundaries annotated on the design | Annotation | *Element map* |
| 03 | **Sequencing method selection** | A method resolving the full construct including repeats selected | Method selection | *Selected method* |
| 04 | **Data receipt and quality assessment** | Sequence integrity and run quality verified | md5sum, run QC | *QC acceptance record* |
| 05 | **Read preprocessing** | Reads preprocessed with settings appropriate to the method | Preprocessing | *Processed reads* |
| 06 | **De novo assembly** | The construct assembled independently of the design to avoid reference bias | Assembly | *Assembled construct* |
| 07 | **Assembly quality assessment** | Assembly completeness and consensus quality assessed | Assembly QC | *Assembly quality* |
| 08 | **Circular structure resolution** | Circular structure resolved with the origin defined consistently | Circularisation | *Resolved structure* |
| 09 | **Design comparison** | The assembly compared base by base against the design sequence | Sequence comparison | *Comparison results* |
| 10 | **Variant identification** | Sequence differences identified with positions and types recorded | Variant identification | *Variant list* |
| 11 | **Repetitive region verification** | Repetitive elements verified specifically since assembly errors concentrate there | Targeted verification | *Repeat verification* |
| 12 | **Functional element verification** | Each functional element verified for presence, sequence and boundaries | Element verification | *Element verification* |
| 13 | **Backbone integrity verification** | Backbone elements verified as intact and correctly positioned | Backbone analysis | *Backbone verification* |
| 14 | **Unintended sequence detection** | Cloning scars and unintended insertions detected | Sequence analysis | *Unintended sequence findings* |
| 15 | **Consequence assessment** | Each discrepancy assessed for functional consequence | Assessment framework | *Consequence assessment* |
| 16 | **Verification determination** | A verification determination made against acceptance criteria | Determination criteria | *Verification determination* |
| 17 | **Verification record assembly** | Assembly, comparison, discrepancies and determination assembled | Report assembly | *Verification record* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **17** | **82 h · 14 d** | **SBQ-10** | **1,90,000** |

**SBQ-11**
SECTOR 01

# Producer Cell Line Clone Selection Analysis

(**Clone selection genomic screening** · Clone selection analysis)

### WHY IT IS DONE

Selecting a production clone commits a product to that cell line for its commercial life, and the clone chosen on early productivity is frequently not the one that remains stable. Genomic screening at selection avoids discovering that years later.

### WHAT IT ANSWERS

Which candidate clone has the integration structure, copy number and clonal purity most likely to remain stable through commercial manufacture?

### WHEN IT IS DONE

During cell line development at the clone selection stage, and when re-selecting after a stability failure.

### WHAT TRIGGERS IT

Completion of a cloning campaign requiring selection, or a stability failure requiring re-selection.

### HOW IT IS DONE

Each candidate clone is characterised for copy number, integration structure and site context, and clonal purity is assessed for evidence that the population derives from a single cell. Integration site genomic environment is assessed for stability risk, clones are compared, and selection risk is ranked alongside productivity data.

### WHAT IT PRODUCES

Per-clone copy number and integration characterisation, integration site context assessment, clonality evidence per clone, comparative stability risk ranking, and a selection recommendation with genomic rationale.

### WHAT IT DETECTS

Transgene copy number and integration structure per clone, integration site genomic context including transcriptional environment, clonal purity and evidence of monoclonality, and early indicators of instability.

### WHAT IT DOES NOT DETECT

It does not measure productivity or product quality, which drive selection alongside genomic criteria. Long-term stability cannot be predicted with certainty from early genomic characterisation, only risk-ranked.

### WHAT MAKES IT HARD

Monoclonality evidence is a regulatory expectation and genomic data contribute to it, but no single method proves it, so the argument is built from converging evidence. Integration site context predicts stability better than copy number and is routinely ignored in favour of the simpler metric.

### WHO DOES THIS JOB

Bioinformatician in cell line development, mid to senior level, working with process development.

### WHAT YOU CAN DO AFTERWARDS

Characterise clones genomically at selection, assess integration site risk, contribute to monoclonality arguments, and rank clones on stability risk.

### WHO HIRES FOR IT

Biologics manufacturers, cell line development companies, contract development organisations, and cell engineering functions.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **20** | **96 hours · 16 days** | Prior copy number and clonality analysis experience | **2,20,000** |

**SBQ-11**
WORKFLOW

# Producer Cell Line Clone Selection Analysis

(**Clone selection genomic screening** · Clone selection analysis)

## COMPLETE WORKFLOW — 20 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Selection criteria definition** | Genomic selection criteria defined alongside productivity criteria | Criteria specification | *Selection criteria* |
| 02 | **Clone inventory and tracking** | Candidate clones inventoried with cloning history and tracking verified | Inventory, tracking | *Clone inventory* |
| 03 | **Data receipt and quality assessment** | Sequence integrity and quality verified per clone | md5sum, FastQC | *QC acceptance record* |
| 04 | **Alignment and processing** | All clones aligned and processed identically | bwa-mem2, Picard | *Analysis-ready BAMs* |
| 05 | **Coverage consistency assessment** | Coverage confirmed comparable across clones for valid comparison | mosdepth | *Coverage consistency* |
| 06 | **Copy number quantification** | Transgene copy number quantified per clone | Copy number quantification | *Per-clone copy number* |
| 07 | **Junction recovery per clone** | Integration junctions recovered for every candidate clone | Junction analysis | *Per-clone junctions* |
| 08 | **Integration site mapping** | Sites mapped to host coordinates for each clone | Coordinate mapping | *Per-clone sites* |
| 09 | **Integration structure reconstruction** | Integration architecture reconstructed per clone | Structure reconstruction | *Per-clone structures* |
| 10 | **Site genomic context annotation** | Genomic and transcriptional context of each site annotated | Annotation | *Context annotation* |
| 11 | **Transcriptional environment assessment** | Site chromatin and transcriptional environment assessed for expression stability | Environment analysis | *Environment assessment* |
| 12 | **Fragile and repetitive site flagging** | Sites in repetitive or unstable genomic regions flagged | Region analysis | *Site risk flags* |
| 13 | **Clonal purity assessment** | Each clone assessed for evidence of a single-cell origin | Clonality analysis | *Clonal purity* |
| 14 | **Heterogeneity detection** | Subpopulation heterogeneity within each clone detected | Heterogeneity analysis | *Heterogeneity findings* |
| 15 | **Monoclonality evidence assembly** | Genomic evidence assembled to contribute to monoclonality assurance | Evidence assembly | *Monoclonality evidence* |
| 16 | **Construct integrity verification** | Construct sequence integrity verified in each clone | Sequence comparison | *Integrity verification* |
| 17 | **Early instability indicator assessment** | Early indicators of instability assessed across clones | Indicator analysis | *Instability indicators* |
| 18 | **Comparative risk ranking** | Clones ranked on combined genomic stability risk | Ranking framework | *Risk ranking* |
| 19 | **Selection recommendation** | A selection recommendation made integrating genomic and productivity evidence | Recommendation framework | *Selection recommendation* |
| 20 | **Selection report assembly** | Per-clone characterisation, evidence and ranking assembled | Report assembly | *Clone selection report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **20** | **96 h · 16 d** | **SBQ-11** | **2,20,000** |

**SBQ-12**
SECTOR 01

# Host Cell DNA Residual Analysis

(**Residual host cell DNA characterisation** · Residual DNA analysis)

### WHY IT IS DONE

Residual host cell DNA in a purified biologic is a defined impurity with limits on both quantity and fragment size, because the risk it poses depends on whether intact functional sequences survive. Sequencing characterises what quantitative assays only measure in bulk.

### WHAT IT ANSWERS

What host cell DNA remains in this purified product, what is its size distribution, and does any functionally relevant sequence survive intact?

### WHEN IT IS DONE

In process characterisation to demonstrate clearance, in supporting residual DNA specifications, and in investigating a residual DNA result that exceeds expectation.

### WHAT TRIGGERS IT

Process characterisation, a specification justification requirement, or an out-of-expectation residual DNA measurement.

### HOW IT IS DONE

Residual DNA is recovered from purified product with a method preserving fragment size, sequenced with size information retained, and fragments are mapped to the host genome. Size distribution is characterised, genomic origin is analysed for regional enrichment, and functionally relevant sequences are assessed for intactness.

### WHAT IT PRODUCES

Fragment size distribution, genomic origin mapping with regional enrichment, intactness assessment for functionally relevant sequences, construct sequence detection, and a characterisation supporting risk assessment.

### WHAT IT DETECTS

Host DNA fragment size distribution, genomic origin of surviving fragments, whether oncogenic or coding sequences survive intact, enrichment of particular genomic regions, and vector or construct sequence carried alongside.

### WHAT IT DOES NOT DETECT

It characterises what is present rather than quantifying total mass, which requires an orthogonal quantitative method. It also cannot establish biological risk, only the sequence characteristics that risk assessment depends on.

### WHAT MAKES IT HARD

Recovery from purified product is technically demanding because the DNA is present at very low mass among abundant protein, and any recovery bias distorts the size distribution that the whole analysis depends on. Fragment size is the risk-relevant attribute and is easily lost through library preparation.

### WHO DOES THIS JOB

Bioinformatician in analytical development or quality control, mid to senior level.

### WHAT YOU CAN DO AFTERWARDS

Characterise residual host DNA including fragment size and genomic origin, assess intactness of relevant sequences, and support risk-based specifications.

### WHO HIRES FOR IT

Biologics manufacturers, analytical development functions, contract testing laboratories, and regulatory affairs groups.

### COMMERCIALS

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

**SBQ-12**
WORKFLOW

# Host Cell DNA Residual Analysis

(**Residual host cell DNA characterisation** · Residual DNA analysis)

**COMPLETE WORKFLOW — 18 ORDERED STEPS**

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Analysis scope confirmation** | Attributes required and their regulatory basis confirmed | Protocol review | *Analysis scope* |
| 02 | **Recovery method qualification** | The recovery method qualified for the product matrix with bias assessed | Method qualification | *Qualified method* |
| 03 | **Size preservation verification** | Library preparation confirmed to preserve fragment size information | Method verification | *Size preservation record* |
| 04 | **Sample and control preparation** | Product and process controls prepared under identical conditions | Preparation | *Prepared material* |
| 05 | **Data receipt and quality assessment** | Sequence integrity and quality verified | md5sum, FastQC | *QC acceptance record* |
| 06 | **Host genome alignment** | Recovered fragments aligned to the host cell genome | bwa-mem2 | *Aligned reads* |
| 07 | **Non-host classification** | Reads not of host origin classified and characterised separately | Classification | *Non-host classification* |
| 08 | **Fragment size distribution analysis** | Fragment length distribution characterised across the recovered population | Length analysis | *Size distribution* |
| 09 | **Size class quantification** | Fragments quantified within regulatory-relevant size classes | Class quantification | *Size class proportions* |
| 10 | **Genomic origin mapping** | Fragment origin mapped across the host genome | Coordinate mapping | *Origin map* |
| 11 | **Regional enrichment analysis** | Enrichment of particular genomic regions assessed statistically | Enrichment analysis | *Enrichment findings* |
| 12 | **Coding sequence intactness assessment** | Whether coding sequences survive intact assessed | Intactness analysis | *Coding intactness* |
| 13 | **Oncogenic sequence assessment** | Survival of sequences of specific regulatory concern assessed | Targeted analysis | *Oncogenic sequence findings* |
| 14 | **Construct sequence detection** | Vector or construct sequence carried with host DNA detected | Sequence detection | *Construct findings* |
| 15 | **Recovery bias assessment** | Method recovery bias assessed against spiked size standards | Bias analysis | *Bias assessment* |
| 16 | **Quantitative method reconciliation** | Sequencing results reconciled against orthogonal quantitative measurement | Reconciliation | *Reconciliation record* |
| 17 | **Risk-relevant characterisation summary** | Attributes relevant to risk assessment summarised | Summary assembly | *Risk characterisation* |
| 18 | **Characterisation report assembly** | Size distribution, origin, intactness and reconciliation assembled | Report assembly | *Residual DNA report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **18** | **86 h · 15 d** | **SBQ-12** | **2,00,000** |

**SBQ-13**
SECTOR 01

# Viral Vector Genome Integrity Analysis

(**Vector genome integrity and impurity characterisation** · Vector integrity QC)

### WHY IT IS DONE

A viral vector preparation is a mixture of intended genomes and a range of incomplete, rearranged and foreign species. The proportion of full-length intended genome is a direct determinant of potency, and the impurities are a direct safety consideration.

### WHAT IT ANSWERS

What proportion of packaged genomes in this vector lot are complete and correct, and what impurity species make up the remainder?

### WHEN IT IS DONE

In vector batch release, in process development and comparability, in investigating potency variability, and in supporting regulatory submissions.

### WHAT TRIGGERS IT

A vector batch requiring release, a process change, a potency deviation, or a submission requirement.

### HOW IT IS DONE

Unencapsidated DNA is removed before extraction so that packaged content is measured, sequencing is performed with a method resolving full-length genomes and terminal repeats, and reads are classified against vector, host and helper references. Genome integrity is assessed across the full construct, truncation positions are mapped, and impurity proportions are quantified against specification.

### WHAT IT PRODUCES

Packaged content classification, full-length genome proportion, truncation position mapping, terminal repeat integrity assessment, reverse packaged and recombinant species quantification, and specification compliance.

### WHAT IT DETECTS

Full-length genome proportion, truncated species with termination positions mapped, terminal repeat integrity, reverse packaged and inverted species, recombinant genomes, and packaged host and helper sequence.

### WHAT IT DOES NOT DETECT

It does not measure infectious titre, capsid content ratio or transduction efficiency, which require separate methods. Encapsidated and surface-associated DNA are indistinguishable without appropriate sample treatment before extraction.

### WHAT MAKES IT HARD

Terminal repeats are structurally difficult to sequence and are precisely where integrity failures occur, so a method that cannot resolve them is not characterising the attribute that matters. Skipping nuclease treatment produces impurity figures that describe the buffer rather than the vector.

### WHO DOES THIS JOB

Bioinformatician in vector analytical development, senior level, working with process development and quality control.

### WHAT YOU CAN DO AFTERWARDS

Characterise packaged vector genomes, resolve terminal repeats, quantify impurity species, and support release and comparability decisions.

### WHO HIRES FOR IT

Gene therapy developers, vector contract manufacturers, analytical development functions, and quality control laboratories.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **22** | **106 hours · 18 days** | Prior sequencing and assembly experience | **2,45,000** |

**SBQ-13**
WORKFLOW

# Viral Vector Genome Integrity Analysis

(**Vector genome integrity and impurity characterisation** · Vector integrity QC)

## COMPLETE WORKFLOW — 22 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Construct and specification registration** | Vector genome sequence and release specifications registered | Registry | *Registered references* |
| 02 | **Reference set assembly** | Vector, host and helper references assembled for classification | Reference assembly | *Reference set* |
| 03 | **Nuclease treatment verification** | Removal of unencapsidated DNA verified before extraction | Treatment verification | *Treatment record* |
| 04 | **Extraction and input quantification** | Packaged nucleic acid extracted and quantified | Extraction, quantification | *Input record* |
| 05 | **Sequencing method selection** | A method resolving full-length genomes and terminal repeats selected | Method selection | *Selected method* |
| 06 | **Data receipt and quality assessment** | Sequence integrity and run quality verified | md5sum, run QC | *QC acceptance record* |
| 07 | **Read classification** | Reads classified against vector, host and helper references | Classification | *Classified reads* |
| 08 | **Vector genome alignment** | Vector-classified reads aligned to the construct reference | Alignment | *Aligned reads* |
| 09 | **Coverage profile analysis** | Coverage analysed across the construct for uniformity and gaps | Coverage analysis | *Coverage profile* |
| 10 | **Full-length species quantification** | Complete genome species quantified as a proportion of packaged content | Length analysis | *Full-length proportion* |
| 11 | **Truncation position mapping** | Termination positions of truncated species mapped | Truncation analysis | *Truncation map* |
| 12 | **Truncated species quantification** | Truncated species quantified by size class | Quantification | *Truncation profile* |
| 13 | **Terminal repeat analysis** | Terminal repeat structure and integrity analysed specifically | Specialised analysis | *Terminal repeat assessment* |
| 14 | **Terminal repeat aberration detection** | Deletions and rearrangements within terminal repeats detected | Aberration analysis | *Aberration findings* |
| 15 | **Reverse packaged species assessment** | Reverse packaged sequence identified and quantified | Orientation analysis | *Reverse packaging findings* |
| 16 | **Recombinant genome detection** | Recombinant and rearranged genome species detected | Recombination analysis | *Recombinant findings* |
| 17 | **Host DNA packaging quantification** | Packaged host sequence quantified with genomic origin characterised | Host analysis | *Host packaging profile* |
| 18 | **Helper sequence quantification** | Helper plasmid packaging quantified | Helper analysis | *Helper profile* |
| 19 | **Sequence variant analysis** | Variants within the vector genome called against error background | Variant analysis | *Variant profile* |
| 20 | **Specification comparison** | All measured attributes compared against release specifications | Specification comparison | *Compliance assessment* |
| 21 | **Batch trend comparison** | Results compared against historical batches for process consistency | Trend analysis | *Trend comparison* |
| 22 | **Release report assembly** | Integrity, impurities, compliance and trends assembled | Report assembly | *Vector integrity report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **22** | **106 h · 18 d** | **SBQ-13** | **2,45,000** |

**SBQ-14**
SECTOR 01

# Capsid & Packaging Composition Analysis

(**Vector packaging composition characterisation** · Packaging composition analysis)

### WHY IT IS DONE

What a vector particle contains determines both how much product reaches the target and what unintended material is delivered with it. Characterising packaging composition is how a manufacturer explains potency variation between lots that look identical by titre.

### WHAT IT ANSWERS

What proportion of packaged material is intended vector genome, and what proportion is host, helper or other sequence?

### WHEN IT IS DONE

In process development to reduce impurity packaging, in batch comparability, in potency investigation, and in supporting specification setting.

### WHAT TRIGGERS IT

A process change affecting packaging, a potency investigation, a comparability exercise, or specification development.

### HOW IT IS DONE

Unencapsidated material is removed, packaged nucleic acid is extracted with size preserved, and reads are classified against all candidate source references. Proportions are quantified with normalisation for genome size, host packaging is analysed for regional origin, and composition is compared across lots and process conditions.

### WHAT IT PRODUCES

Packaged content proportions by source with size normalisation, host DNA regional origin distribution, packaged species size distribution, process condition comparison, and specification support data.

### WHAT IT DETECTS

Relative proportions of vector, host cell, helper plasmid and other packaged sequence, genomic origin distribution of host DNA packaged, size distribution of packaged species, and process-related differences in composition.

### WHAT IT DOES NOT DETECT

It does not measure empty to full particle ratio, which requires biophysical methods, and it does not establish which packaged species are biologically consequential.

### WHAT MAKES IT HARD

Proportions must be normalised for the very different genome sizes of vector, helper and host or the composition figures are meaningless. Recovery bias during extraction distorts the size distribution that composition analysis depends on.

### WHO DOES THIS JOB

Bioinformatician in vector analytical development, mid to senior level.

### WHAT YOU CAN DO AFTERWARDS

Quantify packaging composition with correct normalisation, characterise host DNA packaging, and support process optimisation and specification setting.

### WHO HIRES FOR IT

Gene therapy developers, vector manufacturers, analytical development functions, and process development groups.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **20** | **98 hours · 17 days** | Prior vector analysis experience | **2,25,000** |

**SBQ-14**
WORKFLOW

# Capsid & Packaging Composition Analysis

(**Vector packaging composition characterisation** · Packaging composition analysis)

**COMPLETE WORKFLOW — 20 ORDERED STEPS**

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Analysis scope confirmation** | Packaging attributes in scope and their intended use confirmed | Protocol review | *Analysis scope* |
| 02 | **Reference set assembly** | Vector, host, helper and other candidate references assembled | Reference assembly | *Reference set* |
| 03 | **Nuclease treatment verification** | Unencapsidated material removal verified before extraction | Treatment verification | *Treatment record* |
| 04 | **Size-preserving extraction** | Nucleic acid extracted with fragment size information preserved | Extraction | *Extracted material* |
| 05 | **Data receipt and quality assessment** | Sequence integrity and run quality verified | md5sum, run QC | *QC acceptance record* |
| 06 | **Read classification** | Reads assigned to vector, host, helper or unassigned categories | Classification | *Classified reads* |
| 07 | **Ambiguous read handling** | Reads mapping to more than one source handled under a documented rule | Ambiguity handling | *Ambiguity record* |
| 08 | **Genome size normalisation** | Proportions normalised for the differing genome sizes of each source | Normalisation | *Normalised proportions* |
| 09 | **Packaging proportion quantification** | Proportions of each packaged species quantified with bounds | Quantification | *Packaging proportions* |
| 10 | **Host DNA origin mapping** | Packaged host sequence mapped across the host genome | Coordinate mapping | *Host origin map* |
| 11 | **Host regional enrichment analysis** | Enrichment of particular host regions in packaged DNA assessed | Enrichment analysis | *Regional enrichment* |
| 12 | **Packaged size distribution analysis** | Size distribution of packaged species analysed by source | Length analysis | *Size distribution* |
| 13 | **Helper sequence characterisation** | Packaged helper sequence characterised by region and size | Helper analysis | *Helper characterisation* |
| 14 | **Unassigned sequence characterisation** | Sequence matching no reference assembled and characterised | Assembly, characterisation | *Unassigned findings* |
| 15 | **Recovery bias assessment** | Extraction and library bias assessed against size standards | Bias analysis | *Bias assessment* |
| 16 | **Process condition comparison** | Composition compared across process conditions or batches | Comparative analysis | *Condition comparison* |
| 17 | **Historical trend comparison** | Results compared against historical composition data | Trend analysis | *Trend comparison* |
| 18 | **Specification support analysis** | Data assembled to support composition specification setting | Support analysis | *Specification support* |
| 19 | **Method limitation documentation** | Attributes outside the method's scope documented explicitly | Limitation template | *Limitation statement* |
| 20 | **Composition report assembly** | Proportions, origin, distribution and comparisons assembled | Report assembly | *Packaging composition report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **20** | **98 h · 17 d** | **SBQ-14** | **2,25,000** |

**SBQ-15**
SECTOR 01

# Plasmid Identity & Quality Analysis

(**Plasmid sequence identity and quality testing** · Plasmid QC)

### WHY IT IS DONE

Plasmids are starting materials for vectors, mRNA and cell therapies, and an error or impurity in the plasmid propagates into everything made from it. Sequence-level quality testing is a starting material control with downstream consequences at every stage.

### WHAT IT ANSWERS

Does this plasmid lot match its reference sequence, and what sequence-level impurities are present?

### WHEN IT IS DONE

In plasmid lot release, in plasmid bank establishment, in supplier qualification, and in investigating a downstream deviation traced to plasmid material.

### WHAT TRIGGERS IT

Plasmid lot receipt or release, bank establishment, supplier qualification, or a downstream deviation investigation.

### HOW IT IS DONE

The reference sequence is registered, the plasmid is sequenced with a method resolving the full circular sequence including repetitive regions, and the assembly is compared base by base against the reference. Host genomic contamination is quantified, other plasmid species are detected, and variants are assessed for consequence.

### WHAT IT PRODUCES

Full sequence assembly with reference comparison, variant identification with consequence assessment, structural integrity confirmation, host DNA contamination quantification, foreign plasmid detection, and release recommendation.

### WHAT IT DETECTS

Full plasmid sequence with comparison against the reference, sequence variants and their positions, structural rearrangements and deletions, host genomic DNA contamination, and other plasmid species present.

### WHAT IT DOES NOT DETECT

It does not measure supercoiled fraction, endotoxin or residual host protein, which are separate release attributes. Plasmid topology is outside sequencing scope entirely.

### WHAT MAKES IT HARD

Circular assembly across repetitive backbone elements requires appropriate read lengths, and short-read assembly frequently produces a plausible but incorrect structure. Low-level variant populations within a plasmid preparation are real and require the sequencing error background to be characterised.

### WHO DOES THIS JOB

Bioinformatician in quality control or analytical development, mid level.

### WHAT YOU CAN DO AFTERWARDS

Verify plasmid sequence identity at base resolution, resolve circular and repetitive structure, and quantify sequence-level impurities.

### WHO HIRES FOR IT

Plasmid manufacturers, gene therapy and mRNA developers, contract manufacturers, and quality control laboratories.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **17** | **82 hours · 14 days** | Prior sequence assembly experience | **1,90,000** |

**SBQ-15**
WORKFLOW

# Plasmid Identity & Quality Analysis

(**Plasmid sequence identity and quality testing** · Plasmid QC)

## COMPLETE WORKFLOW — 17 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Reference sequence registration** | The reference plasmid sequence registered for comparison | Sequence registry | *Registered reference* |
| 02 | **Specification confirmation** | Applicable release specifications and attributes confirmed | Specification review | *Specification set* |
| 03 | **Sequencing method selection** | A method resolving the full circular sequence including repeats selected | Method selection | *Selected method* |
| 04 | **Data receipt and quality assessment** | Sequence integrity and run quality verified | md5sum, run QC | *QC acceptance record* |
| 05 | **Read preprocessing** | Reads preprocessed with settings appropriate to the method | Preprocessing | *Processed reads* |
| 06 | **De novo assembly** | The plasmid assembled independently of the reference | Assembly | *Assembled plasmid* |
| 07 | **Circular structure resolution** | Circularity resolved with the origin defined consistently | Circularisation | *Resolved structure* |
| 08 | **Assembly quality assessment** | Assembly completeness and consensus accuracy assessed | Assembly QC | *Assembly quality* |
| 09 | **Reference comparison** | Assembly compared base by base against the registered reference | Sequence comparison | *Comparison results* |
| 10 | **Variant identification and consequence assessment** | Differences identified and assessed for functional consequence | Variant analysis | *Variant assessment* |
| 11 | **Structural integrity verification** | Deletions, insertions and rearrangements assessed against the reference | Structural analysis | *Structural verification* |
| 12 | **Repetitive region verification** | Repetitive backbone elements verified specifically | Targeted verification | *Repeat verification* |
| 13 | **Low-frequency variant analysis** | Subpopulation variants assessed against the sequencing error background | Variant analysis | *Subpopulation findings* |
| 14 | **Host genomic DNA quantification** | Host cell genomic DNA contamination quantified | Contamination quantification | *Host DNA level* |
| 15 | **Foreign plasmid detection** | Other plasmid species present in the preparation detected | Sequence classification | *Foreign plasmid findings* |
| 16 | **Specification comparison** | All measured attributes compared against release specifications | Specification comparison | *Compliance assessment* |
| 17 | **Release report assembly** | Sequence verification, impurities and compliance assembled | Report assembly | *Plasmid QC report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **17** | **82 h · 14 d** | **SBQ-15** | **1,90,000** |

**SBQ-16**
SECTOR 01

# mRNA Sequence & Impurity Analysis

(**mRNA product sequence and impurity characterisation** · mRNA QC)

### WHY IT IS DONE

For an mRNA product the sequence is the active substance, so identity confirmation and impurity characterisation are direct quality attributes rather than supporting information. Truncated and fragmented species affect both potency and immunogenicity.

### WHAT IT ANSWERS

Does this mRNA lot match its intended sequence, and what impurity species are present at what proportions?

### WHEN IT IS DONE

In batch release, in process development and comparability, in stability studies, and in investigating a potency or immunogenicity signal.

### WHAT TRIGGERS IT

A batch requiring release, a process change, a stability timepoint, or a potency investigation.

### HOW IT IS DONE

Sequencing chemistry is selected to resolve full-length species rather than fragment them, reads are aligned to the intended sequence, and identity is confirmed across the full length. Length distributions quantify truncation and fragmentation, variants are called against a characterised error background, tail length is profiled, and residual template is quantified.

### WHAT IT PRODUCES

Full-length identity confirmation, truncation and fragmentation profiles with positions, variant analysis against error background, tail length distribution, residual template quantification, and specification comparison.

### WHAT IT DETECTS

Full-length sequence identity, truncated species with termination positions, fragmented species proportions, sequence variants against the error background, poly-adenosine tail length distribution, and residual template sequence.

### WHAT IT DOES NOT DETECT

It does not measure capping efficiency, modified nucleoside incorporation, secondary structure or lipid formulation attributes, all of which require separate methods. Sequencing chemistry introduces artefacts that mimic product impurities.

### WHAT MAKES IT HARD

Separating genuine product impurity from sequencing artefact requires a characterised control and an explicit error model, because both produce identical signatures. Fragmenting library preparation destroys the length information the impurity analysis depends on.

### WHO DOES THIS JOB

Bioinformatician in analytical development or quality control, mid to senior level.

### WHAT YOU CAN DO AFTERWARDS

Confirm mRNA identity, quantify impurity species against error background, characterise tail length, and support release decisions.

### WHO HIRES FOR IT

mRNA therapeutic and vaccine developers, contract manufacturers, analytical development functions, and quality control laboratories.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **20** | **96 hours · 16 days** | Prior sequencing analysis and long-read basics | **2,20,000** |

**SBQ-16**
WORKFLOW

# mRNA Sequence & Impurity Analysis

(**mRNA product sequence and impurity characterisation** · mRNA QC)

**COMPLETE WORKFLOW — 20 ORDERED STEPS**

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Construct sequence registration** | The intended product sequence registered as the analytical reference | Sequence registry | *Registered construct* |
| 02 | **Quality attribute confirmation** | Sequence-related quality attributes and specifications confirmed | Specification review | *Attribute set* |
| 03 | **Sequencing method selection** | A method resolving full-length species selected | Method selection | *Selected method* |
| 04 | **Control material preparation** | Reference material processed identically to characterise method background | Control preparation | *Control data* |
| 05 | **Data receipt and quality assessment** | Sequence integrity and run quality verified | md5sum, run QC | *QC acceptance record* |
| 06 | **Read preprocessing** | Reads preprocessed with length information preserved | Preprocessing | *Processed reads* |
| 07 | **Alignment to intended sequence** | Reads aligned to the registered construct sequence | Alignment | *Aligned reads* |
| 08 | **Coverage profile analysis** | Coverage analysed across the full construct for gaps and bias | Coverage analysis | *Coverage profile* |
| 09 | **Full-length identity confirmation** | Identity confirmed across the entire construct length | Identity analysis | *Identity confirmation* |
| 10 | **Read length distribution analysis** | Length distribution analysed to characterise product species | Length analysis | *Length distribution* |
| 11 | **Truncation position mapping** | Termination positions of truncated species mapped | Truncation analysis | *Truncation map* |
| 12 | **Truncated species quantification** | Truncated species quantified as a proportion of total | Quantification | *Truncation profile* |
| 13 | **Fragmentation quantification** | Fragmented species quantified with size classes defined | Fragmentation analysis | *Fragmentation profile* |
| 14 | **Method error background characterisation** | Sequencing error background characterised from control material | Background analysis | *Error background* |
| 15 | **Sequence variant analysis** | Variants called against the characterised error background | Variant analysis | *Variant profile* |
| 16 | **Poly-adenosine tail characterisation** | Tail length distribution characterised where applicable | Tail analysis | *Tail profile* |
| 17 | **Residual template quantification** | Residual DNA template or plasmid sequence quantified | Template analysis | *Residual template* |
| 18 | **Impurity classification** | All impurity species classified by type and quantified | Classification | *Impurity profile* |
| 19 | **Specification comparison** | Measured attributes compared against specifications | Specification comparison | *Compliance assessment* |
| 20 | **Release report assembly** | Identity, impurities and compliance assembled | Report assembly | *mRNA QC report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **20** | **96 h · 16 d** | **SBQ-16** | **2,20,000** |

**SBQ-17**
SECTOR 01

# Cell Therapy Product Release Analysis

(**Cell therapy genomic release testing** · Cell therapy release testing)

### WHY IT IS DONE

A cell therapy product is manufactured per patient and released against specification within days, which compresses characterisation into a window no other biologic faces. Genomic release attributes must be measurable, reproducible and fast.

### WHAT IT ANSWERS

Does this cell therapy product meet its genomic release specifications, and is it safe and consistent enough to infuse?

### WHEN IT IS DONE

At batch release for every patient product, in process comparability, and in investigating an out-of-specification release result.

### WHAT TRIGGERS IT

A patient product requiring release, a process comparability exercise, or an out-of-specification investigation.

### HOW IT IS DONE

Release attributes and their specifications are confirmed, patient identity concordance is verified between starting material and product, and vector copy number and transduction are quantified against specification. Adventitious agent and replication competent vector screening are performed within the release window, and composition attributes are measured with results assembled against specification.

### WHAT IT PRODUCES

Patient identity concordance, vector copy number and transduced proportion against specification, adventitious agent screening results, replication competent vector assessment, composition attributes, and a release recommendation.

### WHAT IT DETECTS

Vector copy number per cell, transduced cell proportion, product identity against the patient, adventitious agent status, replication competent vector signal where applicable, and product composition attributes.

### WHAT IT DOES NOT DETECT

It does not measure potency or cytotoxic function, which require functional assays. The release window limits characterisation depth, so some attributes are assessed retrospectively rather than before infusion.

### WHAT MAKES IT HARD

Turnaround is the binding constraint, since a result arriving after the infusion window has no release value. Patient identity concordance between apheresis material and final product is the control against the most serious possible error and must be built into the workflow rather than added.

### WHO DOES THIS JOB

Bioinformatician in cell therapy quality control, senior level, operating within a release timeline.

### WHAT YOU CAN DO AFTERWARDS

Operate genomic release testing for cell therapy within a compressed window, verify patient identity concordance, and make specification-based release recommendations.

### WHO HIRES FOR IT

Cell therapy manufacturers, hospital-based manufacturing units, contract manufacturing organisations, and quality control functions.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **22** | **108 hours · 18 days** | Prior single-cell and vector analysis experience | **2,45,000** |

**SBQ-17**
WORKFLOW

# Cell Therapy Product Release Analysis

(**Cell therapy genomic release testing** · Cell therapy release testing)

## COMPLETE WORKFLOW — 22 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Release specification confirmation** | Genomic release attributes and their specifications confirmed | Specification review | *Release specifications* |
| 02 | **Release window recording** | The available time from sampling to infusion recorded as the operating constraint | Clinical protocol | *Release window* |
| 03 | **Starting material identity capture** | Patient identity captured from apheresis starting material | Identity analysis | *Starting material identity* |
| 04 | **Product sample custody verification** | Custody from product to analytical sample verified | Custody procedure | *Custody record* |
| 05 | **Data receipt and quality assessment** | Sequence integrity and quality verified within the release window | md5sum, FastQC | *QC acceptance record* |
| 06 | **Patient identity concordance** | Product identity confirmed to match the starting material and patient | Concordance analysis | *Identity concordance* |
| 07 | **Vector sequence verification** | The transgene sequence verified against the intended construct | Sequence verification | *Construct verification* |
| 08 | **Vector copy number quantification** | Vector copy number per cell quantified against specification | Copy number quantification | *Copy number result* |
| 09 | **Transduced proportion determination** | Proportion of transduced cells determined | Transduction analysis | *Transduced proportion* |
| 10 | **Integration site assessment** | Integration site distribution assessed where the specification requires it | Integration analysis | *Integration assessment* |
| 11 | **Clonal dominance assessment** | Dominant integration clones assessed against safety criteria | Dominance analysis | *Dominance assessment* |
| 12 | **Replication competent vector screening** | Replication competent vector sequence screened for where applicable | Screening analysis | *RCV screening result* |
| 13 | **Adventitious agent screening** | Rapid adventitious agent screening performed within the window | Screening analysis | *Agent screening result* |
| 14 | **Microbial detection** | Bacterial and mycoplasmal sequence screened for within the window | Detection analysis | *Microbial result* |
| 15 | **Control comparison** | All detections assessed against controls processed in parallel | Control analysis | *Control comparison* |
| 16 | **Product composition assessment** | Cell composition attributes measured where specified | Composition analysis | *Composition result* |
| 17 | **Specification comparison** | Every attribute compared against its release specification | Specification comparison | *Compliance assessment* |
| 18 | **Out-of-specification flagging** | Any attribute outside specification flagged for immediate investigation | Flagging procedure | *OOS flags* |
| 19 | **Turnaround verification** | Result delivery verified against the release window commitment | Timing audit | *Turnaround record* |
| 20 | **Release recommendation** | A release recommendation made against the specification set | Recommendation framework | *Release recommendation* |
| 21 | **Retrospective characterisation planning** | Attributes for retrospective analysis after infusion identified | Planning procedure | *Retrospective plan* |
| 22 | **Release report assembly** | Identity, attributes, compliance and recommendation assembled | Report assembly | *Release report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **22** | **108 h · 18 d** | **SBQ-17** | **2,45,000** |

**SBQ-18**
SECTOR 01

# Microbial Identification for Contamination Investigation

(**Contamination isolate identification and source tracing** · Contamination investigation)

### WHY IT IS DONE

When a contamination event occurs, identifying the organism is only the first question — the decisive one is where it came from. Whole genome comparison against environmental and historical isolates traces the route in a way that species identification alone never can.

### WHAT IT ANSWERS

What organism caused this contamination, and does its genome match any isolate previously recovered from the facility, personnel or materials?

### WHEN IT IS DONE

Immediately following a contamination event, in root cause investigation, and in demonstrating that corrective actions eliminated the source.

### WHAT TRIGGERS IT

A sterility failure, a bioburden excursion, an environmental monitoring alert, or a media fill failure.

### HOW IT IS DONE

The isolate is sequenced and assembled, species and strain are identified, and the genome is compared against the facility's environmental monitoring isolate archive. Relatedness is quantified at core genome level with thresholds interpreted against organism-specific diversity, persistence across time is assessed, and resistance determinants are characterised.

### WHAT IT PRODUCES

Species and strain identification, genomic comparison against the isolate archive with relatedness quantified, persistence assessment, resistance determinant characterisation, source hypothesis with evidence, and corrective action support.

### WHAT IT DETECTS

Species and strain identification of the isolate, genomic relatedness to environmental monitoring isolates, relatedness to historical facility isolates, antimicrobial and biocide resistance determinants, and persistence of a strain across time.

### WHAT IT DOES NOT DETECT

It cannot establish the transmission direction between two related isolates, only that they are related. Absence of a match means the source was not sampled rather than that no source exists.

### WHAT MAKES IT HARD

The comparison is only as good as the archive, so facilities without a systematic environmental isolate collection cannot trace anything. Relatedness thresholds differ by organism, and applying a single cutoff produces both false links and missed ones.

### WHO DOES THIS JOB

Bioinformatician in microbiology or quality investigation, mid to senior level, working with microbiology and quality assurance.

### WHAT YOU CAN DO AFTERWARDS

Identify contamination isolates to strain level, compare against facility archives, quantify relatedness correctly, and construct evidence-based source hypotheses.

### WHO HIRES FOR IT

Biologics and sterile manufacturers, contract manufacturing organisations, quality microbiology functions, and pharmaceutical microbiology consultancies.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **18** | **86 hours · 15 days** | Prior microbial genomics experience | **2,00,000** |

**SBQ-18**
WORKFLOW

# Microbial Identification for Contamination Investigation

(**Contamination isolate identification and source tracing** · Contamination investigation)

## COMPLETE WORKFLOW — 18 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Event scope and timeline capture** | The contamination event, affected material and timeline captured | Event record | *Event record* |
| 02 | **Isolate custody verification** | Custody of the isolate from recovery to sequencing verified | Custody procedure | *Custody record* |
| 03 | **Archive retrieval** | Environmental and historical facility isolate genomes retrieved for comparison | Archive retrieval | *Comparison archive* |
| 04 | **Data receipt and quality assessment** | Sequence integrity and quality verified for the isolate | md5sum, FastQC | *QC acceptance record* |
| 05 | **Genome assembly** | The isolate genome assembled and quality assessed | Assembly, assembly QC | *Assembled genome* |
| 06 | **Species identification** | Species identified from whole genome comparison | Taxonomic analysis | *Species identification* |
| 07 | **Strain typing** | Strain assigned by the appropriate typing scheme for the organism | Typing analysis | *Strain assignment* |
| 08 | **Core genome comparison** | Core genome comparison performed against archive isolates | Core genome analysis | *Core comparison* |
| 09 | **Relatedness quantification** | Genomic distance quantified against archive isolates | Distance analysis | *Relatedness metrics* |
| 10 | **Organism-specific threshold application** | Relatedness assessed against thresholds appropriate to the organism's diversity | Threshold criteria | *Relatedness determination* |
| 11 | **Phylogenetic placement** | The isolate placed phylogenetically among archive and reference isolates | Phylogenetic analysis | *Phylogenetic placement* |
| 12 | **Persistence assessment** | Evidence of the strain persisting across time in the facility assessed | Temporal analysis | *Persistence assessment* |
| 13 | **Resistance determinant characterisation** | Antimicrobial and biocide resistance determinants characterised | Resistance analysis | *Resistance profile* |
| 14 | **Biofilm and survival trait assessment** | Traits supporting persistence in the facility environment assessed | Trait analysis | *Survival trait findings* |
| 15 | **Source hypothesis construction** | A source hypothesis constructed from the genomic and facility evidence | Hypothesis framework | *Source hypothesis* |
| 16 | **Archive gap documentation** | Locations and materials not represented in the archive documented | Gap analysis | *Archive gap record* |
| 17 | **Corrective action support** | Genomic evidence assembled to support corrective and preventive action | Support assembly | *Corrective action support* |
| 18 | **Investigation report assembly** | Identification, relatedness, persistence and source hypothesis assembled | Report assembly | *Investigation report* |

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

**SBQ-19**
SECTOR 01

# Environmental & Utility Monitoring Metagenomics

(**Facility environmental monitoring by sequencing** · Environmental monitoring metagenomics)

### WHY IT IS DONE

Culture-based environmental monitoring recovers only what grows on the chosen media, which is a small fraction of what is present. Sequencing characterises the facility's full microbial landscape, which is what makes trend analysis and source attribution possible.

### WHAT IT ANSWERS

What microbial populations are present across this facility and its utilities, how do they change over time, and where do recurrent organisms persist?

### WHEN IT IS DONE

In routine monitoring programmes alongside compendial methods, in facility qualification, in water system characterisation, and after a contamination event.

### WHAT TRIGGERS IT

A routine monitoring cycle, facility or utility qualification, a contamination investigation, or a trend excursion.

### HOW IT IS DONE

Sampling locations and controls are designed with the facility layout in mind, samples are processed with controls run identically because low-biomass environmental samples are dominated by reagent background. Communities are profiled taxonomically, background is subtracted, and results are analysed spatially and temporally against historical data.

### WHAT IT PRODUCES

Community profiles by location with control-corrected results, organisms not recovered by culture, temporal trend analysis, spatial pattern mapping, recurrent organism identification, and monitoring programme recommendations.

### WHAT IT DETECTS

Microbial community composition across sampled locations, organisms undetectable by the culture media in use, temporal trends and recurrent populations, water system biofilm signatures, and spatial patterns across the facility.

### WHAT IT DOES NOT DETECT

It detects nucleic acid rather than viable organisms, so it does not replace compendial sterility or bioburden testing. Dead organisms and residual nucleic acid from cleaning are detected and can be misread as live contamination.

### WHAT MAKES IT HARD

Environmental samples are low biomass and reagent background dominates them, so controls processed in parallel are the difference between a facility profile and a kit profile. Distinguishing residual nucleic acid from cleaned surfaces from live contamination is not possible by sequencing alone.

### WHO DOES THIS JOB

Bioinformatician in quality microbiology, mid level, working with environmental monitoring programmes.

### WHAT YOU CAN DO AFTERWARDS

Run facility environmental metagenomics with proper background control, analyse spatial and temporal patterns, and support monitoring programme design.

### WHO HIRES FOR IT

Sterile and biologics manufacturers, contract manufacturing organisations, quality microbiology functions, and facility qualification consultancies.

### COMMERCIALS

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

**SBQ-19**
WORKFLOW

# Environmental & Utility Monitoring Metagenomics

(**Facility environmental monitoring by sequencing** · Environmental monitoring metagenomics)

## COMPLETE WORKFLOW — 17 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Monitoring programme scope confirmation** | Locations, frequency and objectives of the monitoring programme confirmed | Programme review | *Programme scope* |
| 02 | **Sampling design review** | Sampling locations reviewed against facility layout and risk zones | Design review | *Sampling design* |
| 03 | **Control design** | Field, reagent and negative controls designed into every sampling round | Control design | *Control plan* |
| 04 | **Sample and control processing** | Samples and controls processed in parallel under identical conditions | Processing | *Prepared libraries* |
| 05 | **Data receipt and quality assessment** | Sequence integrity and quality verified per sample and control | md5sum, FastQC | *QC acceptance record* |
| 06 | **Low-biomass handling verification** | Processing confirmed appropriate for low-biomass environmental samples | Method verification | *Low-biomass record* |
| 07 | **Taxonomic profiling** | Community composition profiled per sampling location | Classification | *Community profiles* |
| 08 | **Reagent background characterisation** | Background taxa characterised from reagent and negative controls | Background analysis | *Background profile* |
| 09 | **Background subtraction** | Sample profiles corrected for reagent-derived background | Decontamination | *Corrected profiles* |
| 10 | **Culture comparison** | Sequencing results compared against parallel culture-based monitoring | Comparative analysis | *Culture comparison* |
| 11 | **Non-culturable organism identification** | Organisms undetected by the culture media in use identified | Comparative analysis | *Non-culturable findings* |
| 12 | **Spatial pattern analysis** | Community patterns analysed across facility zones and locations | Spatial analysis | *Spatial patterns* |
| 13 | **Temporal trend analysis** | Composition trends analysed against historical monitoring data | Trend analysis | *Temporal trends* |
| 14 | **Recurrent organism identification** | Organisms recurring across time and locations identified | Recurrence analysis | *Recurrent organisms* |
| 15 | **Water system signature assessment** | Biofilm and water system associated organisms assessed specifically | Targeted analysis | *Water system findings* |
| 16 | **Viability limitation documentation** | The nucleic acid versus viable organism distinction documented | Limitation template | *Viability caveat* |
| 17 | **Programme recommendation** | Monitoring programme and media recommendations derived from the findings | Recommendation framework | *Programme recommendations* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **17** | **82 h · 14 d** | **SBQ-19** | **1,90,000** |

**SBQ-20**
SECTOR 01

# Process Comparability Genomic Analysis

(**Manufacturing comparability genomic assessment** · Comparability analysis)

### WHY IT IS DONE

Every process change, site transfer and scale-up requires demonstration that the product is unchanged. Genomic attributes are increasingly part of that demonstration for vectors, cells and nucleic acid products, and comparability failures are among the costliest delays in manufacturing.

### WHAT IT ANSWERS

Are the genomic quality attributes of product made by the new process equivalent to those made by the old one?

### WHEN IT IS DONE

In process changes, site transfers, scale-up, and raw material or supplier changes affecting the product.

### WHAT TRIGGERS IT

A planned process change, site transfer, scale-up, or a supplier change requiring comparability demonstration.

### HOW IT IS DONE

Comparability attributes and acceptance criteria are pre-specified before any post-change material is analysed, historical batch variation is characterised to establish the comparison range, and post-change batches are analysed by identical methods. Differences are assessed statistically against historical variation, and any difference outside range is investigated for cause and consequence.

### WHAT IT PRODUCES

Pre-specified comparability protocol with acceptance criteria, historical variation characterisation, attribute-by-attribute comparison with statistical assessment, difference investigation where applicable, and a comparability conclusion.

### WHAT IT DETECTS

Differences in genomic quality attributes between pre-change and post-change material, whether differences fall within historical batch variation, attribute-by-attribute equivalence assessment, and trends indicating a systematic process shift.

### WHAT IT DOES NOT DETECT

It does not establish clinical comparability, which is a broader assessment. Attributes not measured cannot be compared, so the comparability protocol determines what the exercise can conclude.

### WHAT MAKES IT HARD

Historical variation must be characterised from enough batches to define a meaningful range, and most programmes have too few. Pre-specifying acceptance criteria before seeing post-change data is what makes the conclusion credible and is frequently reversed in practice.

### WHO DOES THIS JOB

Bioinformatician in analytical development or quality, senior level, working with statistics and regulatory affairs.

### WHAT YOU CAN DO AFTERWARDS

Design comparability protocols with pre-specified criteria, characterise historical variation, and assess attribute equivalence statistically.

### WHO HIRES FOR IT

Biologics manufacturers, cell and gene therapy companies, contract manufacturers, and regulatory affairs functions.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **21** | **102 hours · 17 days** | Prior analytical and statistical analysis experience | **2,35,000** |

**SBQ-20**
WORKFLOW

# Process Comparability Genomic Analysis

(**Manufacturing comparability genomic assessment** · Comparability analysis)

## COMPLETE WORKFLOW — 21 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Change scope definition** | The process change and its potential product impact defined | Change control record | *Change scope* |
| 02 | **Comparability attribute selection** | Genomic attributes relevant to the change selected and justified | Attribute selection | *Selected attributes* |
| 03 | **Acceptance criteria pre-specification** | Acceptance criteria specified in writing before post-change analysis | Protocol approval | *Pre-specified criteria* |
| 04 | **Historical batch identification** | Pre-change batches identified for characterising historical variation | Batch selection | *Historical batch set* |
| 05 | **Method consistency verification** | Analytical methods confirmed identical across pre and post-change material | Method verification | *Method consistency record* |
| 06 | **Historical variation characterisation** | Batch-to-batch variation characterised from pre-change material | Statistical analysis | *Historical variation* |
| 07 | **Variation adequacy assessment** | Whether enough batches exist to define a meaningful range assessed | Adequacy assessment | *Adequacy determination* |
| 08 | **Data receipt and quality assessment** | Data quality verified for all batches under identical criteria | Quality assessment | *QC acceptance record* |
| 09 | **Post-change batch analysis** | Post-change batches analysed by the identical analytical method | Analysis execution | *Post-change results* |
| 10 | **Attribute-by-attribute comparison** | Each attribute compared between pre and post-change material | Comparative analysis | *Attribute comparisons* |
| 11 | **Statistical equivalence assessment** | Differences assessed statistically against historical variation | Statistical testing | *Equivalence assessment* |
| 12 | **Acceptance criteria evaluation** | Results evaluated against the pre-specified acceptance criteria | Criteria evaluation | *Criteria evaluation* |
| 13 | **Out-of-range difference investigation** | Differences outside historical range investigated for cause | Investigation procedure | *Difference investigation* |
| 14 | **Consequence assessment** | Any confirmed difference assessed for product quality consequence | Assessment framework | *Consequence assessment* |
| 15 | **Trend assessment** | Post-change batches assessed for systematic trend rather than scatter | Trend analysis | *Trend assessment* |
| 16 | **Additional testing determination** | Whether further characterisation is required determined | Determination criteria | *Additional testing decision* |
| 17 | **Comparability conclusion** | A comparability conclusion drawn against the pre-specified criteria | Conclusion framework | *Comparability conclusion* |
| 18 | **Risk assessment** | Residual risk from any unresolved difference assessed | Risk assessment | *Risk determination* |
| 19 | **Regulatory strategy support** | Findings assembled to support the regulatory strategy for the change | Support assembly | *Regulatory support* |
| 20 | **Comparability report assembly** | Criteria, variation, comparisons and conclusion assembled | Report assembly | *Comparability report* |
| 21 | **Change control closure support** | Evidence assembled to support closure of the change control record | Documentation | *Closure support* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **21** | **102 h · 17 d** | **SBQ-20** | **2,35,000** |

**SBQ-21**
SECTOR 01

# Stability Study Genomic Analysis

(**Product stability genomic assessment** · Stability genomic analysis)

### WHY IT IS DONE

Shelf life is set by the attribute that degrades first, and for nucleic acid and vector products that is frequently a genomic attribute rather than a protein one. Stability data determine storage conditions, shelf life and shipping constraints.

### WHAT IT ANSWERS

Do the genomic quality attributes of this product change over storage, at what rate, and what shelf life do they support?

### WHEN IT IS DONE

In formal stability programmes at defined timepoints, in accelerated and stress studies, and in shipping and handling qualification.

### WHAT TRIGGERS IT

A stability programme timepoint, an accelerated study, a shipping qualification, or an excursion investigation.

### HOW IT IS DONE

Stability timepoints and conditions are executed with identical analytical methods throughout, attributes are measured at each timepoint, and trends are modelled rather than assessed pairwise. Degradation rates are estimated, shelf life is projected from the trend, and accelerated and real-time data are compared for pathway consistency.

### WHAT IT PRODUCES

Attribute measurements across timepoints and conditions, trend models with degradation rate estimates, shelf life projection with confidence, accelerated against real-time comparison, and storage condition recommendations.

### WHAT IT DETECTS

Change in sequence integrity over time, fragmentation and degradation trends, impurity species increase, attribute trajectories against time and condition, and the timepoint at which any attribute approaches specification.

### WHAT IT DOES NOT DETECT

It does not measure potency or biological activity directly, and genomic degradation does not always correlate with potency loss. Accelerated conditions may produce degradation pathways that differ from real-time storage.

### WHAT MAKES IT HARD

Method drift across a multi-year study is indistinguishable from product degradation unless controls are analysed alongside every timepoint. Trend modelling rather than pairwise comparison is what supports a shelf life claim, and pairwise testing is what most studies actually do.

### WHO DOES THIS JOB

Bioinformatician in analytical development, mid level, working with stability programme management.

### WHAT YOU CAN DO AFTERWARDS

Analyse stability data as trends, control for method drift, estimate degradation rates, and support shelf life determination.

### WHO HIRES FOR IT

Biologics, vector and mRNA manufacturers, analytical development functions, and contract testing organisations.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **16** | **78 hours · 13 days** | Prior analytical development experience | **1,80,000** |

**SBQ-21**
WORKFLOW

# Stability Study Genomic Analysis

(**Product stability genomic assessment** · Stability genomic analysis)

**COMPLETE WORKFLOW — 16 ORDERED STEPS**

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Stability protocol confirmation** | Timepoints, conditions and attributes confirmed against the approved protocol | Protocol review | *Stability protocol* |
| 02 | **Method consistency planning** | Analytical method consistency across the study duration planned | Method planning | *Consistency plan* |
| 03 | **Control material designation** | Control material designated for analysis alongside every timepoint | Control designation | *Study controls* |
| 04 | **Data receipt and quality assessment** | Data quality verified per timepoint against identical criteria | Quality assessment | *QC acceptance record* |
| 05 | **Control result review** | Control results reviewed at each timepoint to detect method drift | Control analysis | *Method drift assessment* |
| 06 | **Attribute measurement** | Each stability attribute measured at every timepoint and condition | Analysis execution | *Timepoint measurements* |
| 07 | **Method drift correction** | Measurements corrected or qualified where control drift is detected | Correction procedure | *Corrected measurements* |
| 08 | **Trend modelling** | Attributes modelled as trends across time rather than compared pairwise | Trend modelling | *Trend models* |
| 09 | **Degradation rate estimation** | Degradation rates estimated per attribute and condition | Rate estimation | *Degradation rates* |
| 10 | **Condition comparison** | Degradation compared across storage conditions | Comparative analysis | *Condition comparison* |
| 11 | **Accelerated to real-time comparison** | Accelerated and real-time degradation pathways compared for consistency | Pathway comparison | *Pathway consistency* |
| 12 | **Specification approach projection** | The timepoint at which each attribute approaches specification projected | Projection modelling | *Approach projection* |
| 13 | **Shelf life derivation** | Shelf life derived from the limiting attribute with confidence stated | Derivation framework | *Shelf life projection* |
| 14 | **Storage condition recommendation** | Storage and handling conditions recommended from the findings | Recommendation framework | *Storage recommendations* |
| 15 | **Excursion impact assessment** | Impact of temperature or handling excursions assessed where relevant | Impact assessment | *Excursion assessment* |
| 16 | **Stability report assembly** | Measurements, trends, rates and shelf life projection assembled | Report assembly | *Stability report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **16** | **78 h · 13 d** | **SBQ-21** | **1,80,000** |

**SBQ-22**
SECTOR 01

# Analytical Method Validation for Sequencing Assays

(**Sequencing assay validation for GMP use** · Method validation)

### WHY IT IS DONE

A sequencing method used in a regulated release decision must be validated to the same standard as any other analytical method, and the validation parameters that apply to chemistry do not translate straightforwardly to sequencing. Getting this right is what allows a method to be used at all.

### WHAT IT ANSWERS

Does this sequencing method meet the validation requirements for its intended use, across every applicable parameter?

### WHEN IT IS DONE

Before any sequencing method is used for release, comparability or stability decisions, and when a validated method is changed.

### WHAT TRIGGERS IT

A method progressing to regulated use, a method change requiring revalidation, or a regulatory question on method suitability.

### HOW IT IS DONE

Intended use and required performance are defined first because they determine every study, reference materials are selected, and each validation parameter is executed as a designed study rather than an observation. Bioinformatic pipeline contribution is validated separately from wet-lab performance, robustness is tested by deliberate variation, and the validated operating range is defined explicitly.

### WHAT IT PRODUCES

Intended use and requirement definition, accuracy, precision, specificity, detection limit, linearity and robustness study results, separate pipeline validation evidence, defined operating range, and a validation report with acceptance against criteria.

### WHAT IT DETECTS

Method accuracy against reference material, precision within and between runs and operators, specificity and interference, limits of detection and quantitation, linearity and range, robustness to deliberate variation, and the boundaries of the validated method.

### WHAT IT DOES NOT DETECT

Validation establishes that the method performs as characterised, not that the attribute it measures is the right one. It also cannot anticipate every sample condition encountered in routine use.

### WHAT MAKES IT HARD

The bioinformatic pipeline is part of the method and must be validated as such, including version control and change impact, which is where most sequencing method validations are weakest. Detection limit in sequencing depends on depth and background and must be established per matrix rather than claimed generally.

### WHO DOES THIS JOB

Bioinformatician or analytical scientist in method validation, senior level, working with quality assurance and regulatory affairs.

### WHAT YOU CAN DO AFTERWARDS

Validate sequencing methods to regulated standards, validate the analytical pipeline as part of the method, and define operating ranges defensibly.

### WHO HIRES FOR IT

Biologics and advanced therapy manufacturers, contract testing laboratories, analytical development functions, and regulatory consultancies.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **24** | **118 hours · 20 days** | Prior assay development and statistical experience | **2,70,000** |

**SBQ-22**
WORKFLOW

# Analytical Method Validation for Sequencing Assays

(**Sequencing assay validation for GMP use** · Method validation)

## COMPLETE WORKFLOW — 24 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Intended use definition** | The intended use of the method defined since it determines every requirement | Use specification | *Intended use statement* |
| 02 | **Regulatory expectation review** | Applicable validation expectations for the method type reviewed | Guidance review | *Expectation record* |
| 03 | **Performance requirement derivation** | Required performance derived from intended use and specification limits | Requirement derivation | *Performance requirements* |
| 04 | **Method description documentation** | The complete method including wet-lab and computational steps documented | Method documentation | *Method description* |
| 05 | **Reference material selection** | Reference and control materials selected and characterised | Material selection | *Reference materials* |
| 06 | **Validation protocol authorisation** | The validation protocol and acceptance criteria approved before execution | Protocol approval | *Authorised protocol* |
| 07 | **Specificity study execution** | Specificity established including interference and cross-reactivity | Specificity study | *Specificity results* |
| 08 | **Accuracy study execution** | Accuracy established against characterised reference material | Accuracy study | *Accuracy results* |
| 09 | **Repeatability study execution** | Within-run precision established across replicates | Repeatability study | *Repeatability results* |
| 10 | **Intermediate precision study execution** | Between-run, operator and instrument precision established | Precision study | *Intermediate precision* |
| 11 | **Detection limit determination** | Limit of detection established per matrix rather than claimed generally | Detection limit study | *Detection limits* |
| 12 | **Quantitation limit determination** | Limit of quantitation established where the method is quantitative | Quantitation study | *Quantitation limits* |
| 13 | **Linearity and range study** | Linearity and the working range established | Linearity study | *Linearity results* |
| 14 | **Robustness study execution** | Robustness tested by deliberate variation of critical parameters | Robustness study | *Robustness results* |
| 15 | **Computational pipeline validation** | The analytical pipeline validated independently of wet-lab performance | Pipeline validation | *Pipeline validation record* |
| 16 | **Pipeline version control establishment** | Version control and change control established for the pipeline | Version control | *Version control record* |
| 17 | **Reference data reproducibility testing** | Reference datasets re-analysed to confirm pipeline reproducibility | Reproducibility testing | *Reproducibility results* |
| 18 | **Sample stability assessment** | Sample and extract stability under handling conditions assessed | Stability study | *Sample stability* |
| 19 | **Carryover and contamination assessment** | Run-to-run carryover and cross-contamination assessed | Carryover study | *Carryover results* |
| 20 | **Operating range definition** | The validated operating range defined explicitly with boundaries stated | Range definition | *Operating range* |
| 21 | **Acceptance criteria evaluation** | All results evaluated against pre-specified acceptance criteria | Criteria evaluation | *Acceptance evaluation* |
| 22 | **System suitability criteria derivation** | Routine system suitability criteria derived from validation data | Derivation framework | *System suitability criteria* |
| 23 | **Revalidation trigger definition** | Conditions requiring revalidation defined and documented | Trigger definition | *Revalidation triggers* |
| 24 | **Validation report assembly** | All studies, results and conclusions assembled into a validation report | Report assembly | *Validation report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **24** | **118 h · 20 d** | **SBQ-22** | **2,70,000** |

**SBQ-23**
SECTOR 01

# Data Integrity & Computerised System Compliance Analysis

(**Data integrity assessment for computational systems** · Data integrity compliance)

### WHY IT IS DONE

Regulatory findings against manufacturers are dominated by data integrity rather than science, and computational systems handling sequencing data are increasingly within inspection scope. A correct result that cannot be shown to be attributable, legible and contemporaneous is a finding regardless of its correctness.

### WHAT IT ANSWERS

Do the computational systems producing this data meet data integrity expectations, and can every reported result be traced and reproduced?

### WHEN IT IS DONE

In inspection readiness preparation, in qualifying a new computational system, after a data integrity observation, and in periodic system review.

### WHAT TRIGGERS IT

Inspection readiness, new system qualification, a regulatory observation, or a scheduled periodic review.

### HOW IT IS DONE

System scope and the records within it are defined, audit trail coverage and review practice are assessed, and traceability is tested by selecting reported results and tracing them back to raw data. Access control, change control and backup arrangements are reviewed, reproducibility is tested by re-execution, and gaps are documented with remediation.

### WHAT IT PRODUCES

System scope and record inventory, audit trail coverage assessment, traceability test results, access and change control review, backup and retention assessment, reproducibility test results, and a gap and remediation register.

### WHAT IT DETECTS

Gaps in audit trail coverage, breaks in traceability from raw data to reported result, uncontrolled analytical changes, access control weaknesses, backup and retention gaps, and results that cannot be reproduced from retained data.

### WHAT IT DOES NOT DETECT

It assesses control and documentation rather than scientific correctness. A fully compliant system can produce a scientifically wrong answer, and that is a separate review.

### WHAT MAKES IT HARD

Retrospective assessment finds gaps that cannot be closed retrospectively, because contemporaneous records cannot be recreated. Analytical scripts and manual data handling steps outside validated systems are where traceability usually breaks and are the hardest to bring under control.

### WHO DOES THIS JOB

Bioinformatician or computational quality specialist, senior level, working with quality assurance and information technology.

### WHAT YOU CAN DO AFTERWARDS

Assess computational systems against data integrity expectations, test traceability and reproducibility, and construct remediation plans.

### WHO HIRES FOR IT

Biologics and pharmaceutical manufacturers, contract organisations, quality assurance functions, and compliance consultancies.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **20** | **98 hours · 17 days** | Prior pipeline and quality systems experience | **2,25,000** |

**SBQ-23**
WORKFLOW

# Data Integrity & Computerised System Compliance Analysis

(**Data integrity assessment for computational systems** · Data integrity compliance)

## COMPLETE WORKFLOW — 20 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **System scope definition** | Computational systems and records within scope identified | Scope definition | *System scope* |
| 02 | **Record inventory** | Electronic records produced by each system inventoried and classified | Inventory compilation | *Record inventory* |
| 03 | **Regulatory expectation mapping** | Applicable data integrity expectations mapped to each record type | Expectation mapping | *Expectation map* |
| 04 | **Audit trail coverage assessment** | Audit trail coverage assessed against the records requiring it | Coverage assessment | *Audit trail assessment* |
| 05 | **Audit trail review practice assessment** | Whether audit trails are reviewed and by whom assessed | Practice review | *Review practice record* |
| 06 | **Access control review** | User access, privileges and segregation of duties reviewed | Access review | *Access control assessment* |
| 07 | **Change control assessment** | Change control over analytical software and configuration assessed | Change control review | *Change control assessment* |
| 08 | **Version control assessment** | Version control over pipelines, scripts and references assessed | Version control review | *Version control assessment* |
| 09 | **Traceability test sampling** | Reported results sampled for end-to-end traceability testing | Sampling procedure | *Traceability sample* |
| 10 | **Raw data traceability testing** | Each sampled result traced back to its raw data through every step | Traceability testing | *Traceability results* |
| 11 | **Intermediate step documentation assessment** | Manual steps and scripts outside validated systems identified | Gap analysis | *Manual step inventory* |
| 12 | **Reproducibility testing** | Sampled results re-executed to confirm reproducibility | Re-execution | *Reproducibility results* |
| 13 | **Discrepancy investigation** | Any failure to reproduce investigated and explained | Investigation procedure | *Discrepancy record* |
| 14 | **Backup and retention assessment** | Backup, archival and retention arrangements assessed against requirements | Retention review | *Retention assessment* |
| 15 | **Restoration testing** | Restoration from backup tested and verified | Restoration test | *Restoration results* |
| 16 | **Data lifecycle assessment** | The full data lifecycle from generation to disposal assessed | Lifecycle review | *Lifecycle assessment* |
| 17 | **Gap identification and classification** | Identified gaps classified by severity and regulatory exposure | Gap classification | *Gap register* |
| 18 | **Remediation planning** | Remediation actions planned with owners and timelines | Remediation planning | *Remediation plan* |
| 19 | **Retrospective limitation documentation** | Gaps that cannot be closed retrospectively documented explicitly | Limitation documentation | *Retrospective limitations* |
| 20 | **Assessment report assembly** | Scope, findings, gaps and remediation assembled | Report assembly | *Data integrity report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **20** | **98 h · 17 d** | **SBQ-23** | **2,25,000** |

**SBQ-24**
SECTOR 01

# Deviation & Out-of-Specification Investigation Analysis

(**Genomic investigation of manufacturing deviations** · Deviation investigation)

### WHY IT IS DONE

When a batch fails specification the investigation determines whether it is released, reworked or discarded, and whether the process is fit to continue. Genomic evidence frequently distinguishes a laboratory error from a genuine product problem, which are handled entirely differently.

### WHAT IT ANSWERS

What caused this out-of-specification or deviation result, was it laboratory error or a genuine product attribute, and what is the impact on this and other batches?

### WHEN IT IS DONE

Immediately following an out-of-specification result, in deviation investigation, and in assessing impact across other batches.

### WHAT TRIGGERS IT

An out-of-specification result, a manufacturing deviation, an unexpected trend, or a complaint investigation.

### HOW IT IS DONE

The original result is assessed for analytical validity before any retesting, sample identity is verified to exclude mix-up, and the analysis is repeated under controlled conditions. Contamination and error signatures are sought specifically, candidate root causes are tested against the evidence, and impact on other batches is assessed by retrospective analysis.

### WHAT IT PRODUCES

Original result validity assessment, sample identity verification, controlled retest results, error signature analysis, root cause evidence assessment, batch impact analysis, and an investigation conclusion supporting disposition.

### WHAT IT DETECTS

Whether the original result is reproducible, laboratory error signatures such as sample mix-up or contamination, genuine product attribute deviation, the extent of impact across batches, and evidence supporting or refuting candidate root causes.

### WHAT IT DOES NOT DETECT

It cannot establish root cause on its own, since the analytical evidence must be combined with process and laboratory records. An unreproducible result does not prove laboratory error, only that the original observation could not be confirmed.

### WHAT MAKES IT HARD

Retesting without first establishing analytical validity of the original result is the classic investigation failure, because it converts an investigation into a search for a passing result. Sample identity verification is what most often resolves an investigation and is most often done last.

### WHO DOES THIS JOB

Bioinformatician in quality investigation, mid to senior level, working with quality assurance and manufacturing.

### WHAT YOU CAN DO AFTERWARDS

Investigate out-of-specification results systematically, distinguish laboratory error from product deviation, and assess batch impact.

### WHO HIRES FOR IT

Biologics and advanced therapy manufacturers, contract manufacturing organisations, quality assurance functions, and quality control laboratories.

### COMMERCIALS

| STEPS IN WORKFLOW | TRAINING DURATION | PREREQUISITE | TRAINING FEE (INR) |
|---|---|---|---|
| **19** | **92 hours · 16 days** | Prior analytical and investigative experience | **2,10,000** |

**SBQ-24**
WORKFLOW

# Deviation & Out-of-Specification Investigation Analysis

(**Genomic investigation of manufacturing deviations** · Deviation investigation)

## COMPLETE WORKFLOW — 19 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Event capture and scope definition** | The deviation or out-of-specification result captured with full context | Event record | *Event record* |
| 02 | **Immediate containment assessment** | Whether affected material requires immediate quarantine assessed | Containment procedure | *Containment decision* |
| 03 | **Original result validity assessment** | The original result assessed for analytical validity before any retesting | Validity assessment | *Validity determination* |
| 04 | **Analytical record review** | Instrument, run and analytical records reviewed for the original result | Record review | *Record review findings* |
| 05 | **Sample identity verification** | Sample identity verified to exclude mix-up as the cause | Identity analysis | *Identity verification* |
| 06 | **Sample integrity assessment** | Sample handling and integrity assessed as a candidate cause | Integrity assessment | *Integrity findings* |
| 07 | **Contamination signature analysis** | The result assessed for signatures characteristic of contamination | Signature analysis | *Contamination assessment* |
| 08 | **Laboratory error hypothesis testing** | Laboratory error hypotheses tested against the available evidence | Hypothesis testing | *Error hypothesis results* |
| 09 | **Controlled retest design** | Retesting designed with controls and criteria specified in advance | Retest design | *Retest protocol* |
| 10 | **Controlled retest execution** | Retesting executed under the specified protocol | Retest execution | *Retest results* |
| 11 | **Retest interpretation** | Retest results interpreted against the original without bias toward passing | Interpretation framework | *Retest interpretation* |
| 12 | **Product attribute hypothesis testing** | Genuine product deviation hypotheses tested against the evidence | Hypothesis testing | *Product hypothesis results* |
| 13 | **Process record correlation** | Findings correlated against manufacturing and process records | Correlation analysis | *Process correlation* |
| 14 | **Root cause evidence assessment** | Candidate root causes assessed against all assembled evidence | Assessment framework | *Root cause assessment* |
| 15 | **Batch impact analysis** | Impact on other batches assessed by retrospective analysis | Retrospective analysis | *Batch impact* |
| 16 | **Trend review** | Historical data reviewed for prior occurrences of the same signal | Trend analysis | *Trend findings* |
| 17 | **Corrective action support** | Evidence assembled to support corrective and preventive action | Support assembly | *Corrective action support* |
| 18 | **Disposition recommendation** | A disposition recommendation made for the affected material | Recommendation framework | *Disposition recommendation* |
| 19 | **Investigation report assembly** | Validity, testing, root cause and impact assembled into the investigation record | Report assembly | *Investigation report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **19** | **92 h · 16 d** | **SBQ-24** | **2,10,000** |

**SBQ-25**
SECTOR 01

# Rapid Sterility Assurance by Sequencing Methods

(**Rapid microbiological method implementation** · Rapid sterility methods)

### WHY IT IS DONE

Compendial sterility testing takes fourteen days, which is longer than the shelf life of a cell therapy product. Sequencing-based rapid methods provide an answer within the release window, and demonstrating their equivalence is what allows them to be used.

### WHAT IT ANSWERS

Does this product show microbial sequence within the release window, and is the rapid method demonstrably equivalent to the compendial one?

### WHEN IT IS DONE

In release testing of short-shelf-life products, in method implementation and equivalence demonstration, and in supporting regulatory acceptance of a rapid method.

### WHAT TRIGGERS IT

A short-shelf-life product requiring release, a rapid method implementation project, or a regulatory submission for method acceptance.

### HOW IT IS DONE

The method is challenged against the required organism panel at low inoculum with detection limits established per organism, and equivalence against the compendial method is demonstrated across the panel. Routine samples are analysed with controls processed identically, detections are confirmed and identified, and results are reported within the release window.

### WHAT IT PRODUCES

Per-organism detection limits from challenge studies, compendial equivalence demonstration, routine sample results with control comparison, detection identification and confirmation, and a release recommendation within the window.

### WHAT IT DETECTS

Microbial sequence at the sensitivity the method is validated to, species identification of detections, and the achieved detection limit for the panel of challenge organisms.

### WHAT IT DOES NOT DETECT

It detects nucleic acid rather than viable organisms, which is the central obstacle to accepting it as a sterility method, and dead organisms from upstream processing are detected. Very low-level contamination near a single organism per container remains at the edge of what any method resolves.

### WHAT MAKES IT HARD

Demonstrating equivalence to a fourteen-day culture method requires challenge across a defined organism panel at very low inoculum, and sensitivity differs substantially between organisms. The viable-versus-detectable distinction is the argument regulators press hardest on.

### WHO DOES THIS JOB

Bioinformatician in quality microbiology or rapid methods, senior level, working with microbiology and regulatory affairs.

### WHAT YOU CAN DO AFTERWARDS

Implement and validate rapid sequencing-based microbiological methods, demonstrate compendial equivalence, and operate them within release windows.

### WHO HIRES FOR IT

Cell and gene therapy manufacturers, sterile product manufacturers, rapid method technology providers, and quality microbiology functions.

### COMMERCIALS

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

**SBQ-25**
WORKFLOW

# Rapid Sterility Assurance by Sequencing Methods

(**Rapid microbiological method implementation** · Rapid sterility methods)

## COMPLETE WORKFLOW — 18 ORDERED STEPS

| # | STEP | WHAT HAPPENS | TOOL / METHOD | OUTPUT |
|---|---|---|---|---|
| 01 | **Method scope and intended use definition** | The products and release decisions the method will support defined | Use specification | *Intended use statement* |
| 02 | **Compendial comparison basis establishment** | The compendial method against which equivalence will be shown identified | Comparison basis | *Comparison basis* |
| 03 | **Challenge organism panel definition** | The organism panel for challenge studies defined against requirements | Panel definition | *Challenge panel* |
| 04 | **Low inoculum challenge preparation** | Challenge organisms prepared at the low inoculum levels required | Inoculum preparation | *Challenge inocula* |
| 05 | **Matrix-specific challenge execution** | Challenges performed in the actual product matrix | Challenge execution | *Challenge results* |
| 06 | **Per-organism detection limit determination** | Detection limits established for every organism in the panel | Limit determination | *Per-organism limits* |
| 07 | **Compendial parallel testing** | Compendial testing performed in parallel on the same challenged material | Parallel testing | *Compendial results* |
| 08 | **Equivalence assessment** | Rapid method performance assessed against compendial across the panel | Equivalence analysis | *Equivalence assessment* |
| 09 | **Inhibition and interference assessment** | Product matrix inhibition of detection assessed | Interference study | *Interference assessment* |
| 10 | **Control design for routine use** | Positive, negative and process controls designed for routine operation | Control design | *Routine control plan* |
| 11 | **Routine sample analysis** | Product samples analysed with controls processed in parallel | Analysis execution | *Sample results* |
| 12 | **Detection assessment against controls** | Detections assessed against parallel controls before interpretation | Control comparison | *Detection assessment* |
| 13 | **Organism identification** | Detected organisms identified to species where possible | Identification analysis | *Identification results* |
| 14 | **Confirmation analysis** | Detections confirmed by an orthogonal or targeted approach | Confirmation analysis | *Confirmation results* |
| 15 | **Viability limitation documentation** | The viable versus detectable distinction documented explicitly | Limitation template | *Viability caveat* |
| 16 | **Turnaround verification** | Result delivery verified against the release window requirement | Timing audit | *Turnaround record* |
| 17 | **Release recommendation** | A release recommendation made within the available window | Recommendation framework | *Release recommendation* |
| 18 | **Method report assembly** | Challenge results, equivalence, detections and recommendation assembled | Report assembly | *Rapid method report* |

| STEPS IN WORKFLOW | TRAINING DURATION | PROGRAMME CODE | TRAINING FEE (INR) |
|---|---|---|---|
| **18** | **88 h · 15 d** | **SBQ-25** | **2,05,000** |

# Build and staff a *GMP genomic quality function*

Each of the twenty-five programmes in this catalogue trains one complete type of NGS data analysis as a regulated manufacturing quality function 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

SBQ · Sector 05 of 16

SCOPE

25 analysis types · 481 steps

DURATION

2328 contact hours · 397 working days

COMPLETE CATALOGUE FEE

INR 34,84,000 (list 53,60,000)

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