Genomics Services at the University of Edinburgh

Lee Murphy[1] and Javier Santoyo-Lopez[2]

1. Genetics Core, Edinburgh Clinical Research Facility, Western General Hospital, CMVM

2. Edinburgh Genomics, Kings' Buildings, CSE

1. Introduction

This report summarises the Genomics Services available within The University of Edinburgh for the One Health Genomics Edinburgh Executive Committee. It focuses on genomics equipment and services that are located within core facilities, and are open for use by any researcher. It does not cover academic laboratories, who may have equipment, as these may not be available to a wide user base.

There are two large-scale sequencing facilities (Genetics Core, Western General Hospital & Edinburgh Genomics, Kings' Buildings) both with a wide user base. In addition, there are a number of facilities highlighted providing lower-throughput sequencing, as well as those supporting single-cell and spatial profiling.

The University of Edinburgh has a diverse range of core facilities supporting genomics research.

2. Genomics Facilities with Sequencing Capacity

2a. Genetics Core, Edinburgh Clinical Research Facility

https://clinical-research-facility.ed.ac.uk/core-services/genetics

Lead: Lee Murphy (Lee.Murphy@ed.ac.uk)

Main services:

The Genetics Core was established in 2000, as part of the Edinburgh Clinical Research Facility – a joint NHS Lothian/University of Edinburgh facility focusing on clinical research. Based within the College of Medicine and Veterinary Medicine, the Genetics Core provides support for clinicians and researchers with their sample processing, sample storage and genome analysis. The team consists of twelve technicians working to the principles of Good Laboratory Practice for Laboratories, and is an approved supplier by the NHS Lothian/University of Edinburgh ACCORD office. It supports approximately 100 new and ongoing projects a year and recharges for about £2M of work each financial year, this includes an average of £200k of commercial service work p.a. through Edinburgh Innovations.

The main focus of the Genetics Core is to provide equitable access for researchers, at any stage of their career, to cutting-edge genomics. This is implemented through modular workflows supported by Standard Operating Procedures, a Quality Assurance system and a Laboratory Information Management System. A focus of the Genetics Core is providing solutions to difficult and 'near-to-clinic' samples.

Main Genomics Platforms:

1) DNA & RNA extraction: DNA (whole genome and cfDNA) and RNA extraction (total RNA and miRNA) from blood, saliva, FFPE and other tissues, using either automated platforms (3x Revvity Chemagic; 1x Qiagen QiCube) or manual extraction for more difficult samples. Approximately 10k DNA and 2k RNA samples are extracted annually.

2) Sample storage: A Laboratory Information Management System (LIMS) is used to track samples entering the lab and through the workflows, including into the monitored and alarmed freezer suite. Storage currently consists of approximately 160,000 samples held within thirteen -20^{o}C freezers and nine ULT freezers.

3) Genomic Analysis

Stilla Naica six-colour digital PCR platform

Life QuantStudio12K Flex Real-Time PCR system – for Taqman genotyping (used particularly for APOE genotyping) and qPCR

Illumina iScan microarray scanner – for genotyping & methylation arrays. Coupled with the Tecan Evo liquid handling platform and autoloader arm has enabled more than 10k samples to be analysed per year

The sequencing platforms consist of

a) Illumina iSeq – for low-throughput requirements allowing 4M reads at 300 cycles

b) Two Illumina NextSeq2000 platforms – for a wide range of requirements from 100M reads up to 1.8B reads and from 50 cycles up to 600 cycles, with the aim of providing a fast and responsive service. The Illumina platforms were run 350 times in 2023 providing sequencing data on 9,802 samples.

c) Oxford Nanopore MinION (for low-throughput) and PromethION (for high-throughput) long-read single-molecule sequencing that allows epigenetic marks to also be resolved.

Library preparation is supported by Beckman NGeniuS automated library preparation platform – for simpler library preparation techniques. However, most library preparation is processed manually by trained technicians to allow the best data to be produced from difficult samples. This includes a wide-selection of kits and workflows, including Total RNA, mRNA, Gene Expression Profiling, CAGE, Exome sequencing and targeted gene panels.

4) Quality Control and automation

Qubit, Nanodrop, Agilent Bioanalyser and Agilent Fragment Analyser. Beckman Biomek FXp for cherry-picking plates. Covaris E220 for mechanical shearing of DNA.

2b. Edinburgh Genomics

web: https://genomics.ed.ac.uk **email:** genomics@ed.ac.uk

Location: Ashworth's Building, School of Biological Sciences, King's Buildings Campus (KB)

Lead: Javier Santoyo-Lopez, PhD (javier.santoyo@ed.ac.uk)

Background:

The Facility was established in 1996 by Prof Mark Blaxter under the name of GenePool at Ashworth's laboratories, KB. In 2014 Genepool was merged with ARK Genomics, the genomics core facility at The Roslin Institute, College of Medicine and Veterinary Medicine (CMVM), to become Edinburgh Genomics (EG). This rebranded facility had a single lab based at KB, and was part of both colleges, CSE and CMVM.

In 2015 a new Genome Sequencing laboratory was established at the Roslin Institute to fulfil the Scottish Genome Partnership project, led by Prof Tim Aitman. This Genome facility was equipped with 5 Illumina HiSeq X sequencers and a computing infrastructure of 3,000 cores and 3PB of data storage. The facility was operative for nearly five years and generated high-quality data for more than 15,000 genomes, holding an ISO 17025:2005 accreditation for human and non-human genome sequencing.

In January 2020 the University closed down the genome sequencing laboratory at the Roslin institute, re-structuring the facilities. EG became a much smaller group with just 12 staff (less than one third of the 40 staff members were retained) and with one single laboratory based at KB, which is fully managed by the School of Biological Sciences (SBS).

Main services:

Edinburgh Genomics is a Sequencing and Bioinformatics Core Facility at the University of Edinburgh. It is located at the King's Buildings (KB) Campus and is part of the SBS Professional Services, under CSE financial management.

The facility provides access to a range of genomic technologies and services including next-generation sequencing (NGS), third-generation sequencing (TGS), transcriptomics, single-cell sequencing, epigenomics, metagenomics, bioinformatics analysis and advanced genomics & bioinformatics training.

EG has a genomics lab team with 5 technicians, a bioinformatics team with 5 bioinformaticians, a (part time) project & training support officer and a facility manager. The Facility operates under Good Laboratory Practices and is working towards a Silver Award for sustainability.

The facility collaborates with a broad range of research groups and organizations (internal and external) on a variety of research projects, including human health, agriculture and environmental sciences, and have generated data for more than 1,300 peer-reviewed publications.

EG processes more than a hundred projects every year with an approximated turnover of £1.4M. EG delivers about 20 advanced training courses a year, with a yearly total of more than 300 trainees. It also provides genomics project design to a large number of research groups, ERCs and PhD students.

A list of detailed services is:

DNA sequencing:

- Genome de novo sequencing.
- Genome resequencing.
- Target sequencing (exon and gene panels).
- Whole genome methylation sequencing and Methyl-Seq
- Genotyping by sequencing
- ChIP-Seq and ATAC-Seq
- Short-read 16S rRNA sequencing
- Full length 16S rRNA sequencing
- shotgun metagenomics sequencing
- Hi-C library preparations
- Ultralong nanopore sequencing
- Plasmid libraries sequencing (full length insert sequencing)
- Nanopore cfDNA sequencing
- Sequencing of libraries prepared by users (Illumina and nanopore)

RNA sequencing:

- Total RNA, mRNA and cDNA sequencing
- Direct RNA sequencing
- Full-length RNA ISO-seq transcriptome sequencing
- Small RNA sequencing
- Low input RNA sequencing
- Parse Biosciences single-cell RNA library preparation and sequencing
- Short-read and long-read single-cell RNA sequencing

Bioinformatics services:

- *De novo* Genome/transcriptome assembly and annotation
- Variant calling, annotation and discovery (SNV and INDELS)
- Large structural variant analysis (SVs)
- RNA-seq and miRNA differential expression analysis
- Gene pathway and GO analysis
- Gene clustering
- Single cell transcriptomics data processing and analysis
- Pangenome analysis
- Haplotype phasing
- ChIP-seq and ATAC-seq peak detection
- Metabarcoding/metagenomics analysis
- Amplicon analysis
- Bioinformatics bespoke analysis

Bioinformatics training (ELIXIR-UK recognised training):

- Introduction to Python (onsite & online)
- Introduction to Long reads Bioinformatics (onsite & online)
- RNA-Seq data analysis (onsite & online)

- R for Biologists (onsite & online)
- Variant calling (onsite & online)
- Introduction to Linux (onsite & online)
- Metabarcoding and metagenomics (onsite)
- Advanced Python (onsite)
- Data Manipulation and visualization (onsite)
- Snakemake (onsite & online)
- Introduction to single cell RNA-seq analysis (onsite & online)
- Genome Assembly (onsite)
- Structural Variant analysis (onsite)

Genomics lab training

- Short-read DNA-seq library preparations (onsite)
- Short-read RNA-seq library preparation (onsite)
- Long-read library preparation (including DNA extraction and library multiplexing) for Oxford Nanopore sequencing (onsite)

Open access to UoE groups of the following genomics instrumentation:

- PromethION
- MinIONs
- Tapestation
- Bioanalyzer
- BluePippin
- SageELF
- Covaris
- qPCR instruments (e.g. Roche lightcycler 480)

Sequencing Platforms:

- Illumina NovaSeq 6000
- MiSeq (2X)
- PromethION 24
- MinION (3X)
- PacBio Revio and Sequel IIe

Computing infrastructure **(hosted at EPCC)**

- HPC infrastructure with 3,000 cores
- 2PB of lustre storage
- 5PB of tape storage
- Dedicated transfer server (for data delivery)

2c. Genomics Platform, Roslin Institute

https://www.ed.ac.uk/roslin/facilities-resources/genomics-platform

Lead: Dr James Furniss (jfurniss@ed.ac.uk)

Main services:

The production of single-cell-RNA-seq & single-nuclei-RNA-seq library preparations. Currently this utilises two different techniques, either the 10x Chromium platform or the Parse whole transcriptome workflow, a plate-based approach requiring no dedicated equipment.

The facility provides access to Oxford Nanopore sequencing through the GridION (which is able to run five lower-throughput MinION flow cells at a time) and the PromethION P2 Solo (which is able to run two higher-throughput PromethION flow cells at a time).

In addition, the facility has a Promega Maxwell Rapid Sample Concentrator platform that is available for automated DNA and RNA extractions. This platform uses a bead-based reagent-cartridge to allow rapid extraction from up to 48 samples at a time on a wide variety of sample types (blood, saliva, FFPE, cells, tissue). Sample input is low (50µL – 300µL fluids, 5mg – 50mg tissue or $5x10^2$ – $5x10^6$ cells) with extracted DNA and RNA eluted in to 30-100µl final volume.

Genomics Platforms:

DNA & RNA Extraction - Promega Maxwell RSC

Single Cell – 10x Chromium platform

Sequencing – Oxford Nanopore Technologies GridION & PromethION P2

2d. DNA Sequencing Facility, Institute of Genetics and Cancer

https://www.ed.ac.uk/institute-genetics-cancer/facilities/dna-sequencing-facility

Lead: Stephen Brown (stephen.brown@ed.ac.uk)

Main services:

The provision of Sanger sequencing using a Thermo 3730xl DNA Analyzer and newly purchased SeqStudio 24 Flex Genetic Analyzer. The facility can handle samples in single tubes through to 384-well and offers automated PCR purification before Sanger Sequencing. The 3730xl can also be used for user-prepared cell line authentication (STR profiling) and genotyping (Microsatellites, SNPs, ALFPs).

The facility provides a plasmid mini-prep extraction and isolation service processed by Beckman NX^p automation and downstream DNA and RNA quality and quantification using Agilent Tapestation and Agilent Bioanalyzer.

Genomics Platforms:

Sequencing – Thermo Fisher 3730xl DNA Analyzer

Robotics – 2x Beckman Coulter Biomek NX^p automated workstations

3. Other Genomics Facilities

3a. Flow Cytometry Facility, Institute of Genetics and Cancer

https://www.ed.ac.uk/institute-genetics-cancer/facilities/flow-cytometry-facility

Lead: Elisabeth Freyer (Lizzie.Freyer@ed.ac.uk)

Main services:

The facility consists of two cell sorter and two Flow Cytometry analysers. In addition, it also houses the institutes Chromium 10x platform for the production of single-cell-RNA-seq & single-nuclei-RNA-seq library preparations.

Genomics Platforms:

Single Cell – 10x Chromium platform

3b. Host and Tumour Profiling Unit microarray services, Institute of Genetics and Cancer

https://www.ed.ac.uk/institute-genetics-cancer/facilities/htpu-microarray-services

Lead: Alison Munro (alison.munro@ed.ac.uk)

Main services:

The facility provides cutting-edge protein, antibody and transcriptomic microarray facilities for the analysis of complex signalling pathways utilising Quanterix-2470-microarrayer and Innopsys 710 infra-red microarray scanner. The facility also houses two nanoString platforms to study transcriptomics. The nCounter platform is used for bulk gene expression analysis of cell line, FFPE or fresh frozen tissue.

The GeoMx uses barcodes and an NGS-based output to allow spatially resolved tissue analysis, and can image on the nCounter for lower-plex assays.

Genomics Platforms:

nanoString nCounter platform

nanoString GeoMx platform

3c. Single-cell multi-omics Facility, Institute of Regeneration and Repair

https://www.ed.ac.uk/regenerative-medicine/facilities/single-cell-multi-omics

Lead: Dr Andrea Corsinotti (A.Corsinotti@ed.ac.uk)

Main services:

The production of single-cell-RNA-seq & single-nuclei-RNA-seq library preparations. Currently this utilises two different techniques, either the 10x Chromium platform or the Parse whole transcriptome workflow, a plate-based approach requiring no dedicated equipment.

Genomics Platforms:

single-cell/single-nuclei assays using
- a) 10X genomics Chromium X platform (droplet-based)
- b) Parse Biosciences (plate-based) platform.

Spatial transcriptomics
- 10X genomics Visium
- 10X genomics Xenium
- Nanostring CosMx

3d. Edinburgh Genome Foundry, SBS, SCE

https://biology.ed.ac.uk/research/facilities/edinburgh-genome-foundry

Lead: Dr Rennos Fragkoudis

Main services:

Modular assembly of synthetic DNA parts into plasmid constructs. They deliver high-throughput projects involving long constructs (>5kbp), constructs comprising large numbers of genetic parts, or combinatorial libraries.

EGF offers a wide range of automated molecular biology protocols as a service using their robotic platform (e.g., DNA transformation and purification, qPCR, Replica plating, Colony picking, Mini-fermentations). They develop and share industry-grade software for DNA Design, Laboratory Automation and Synthetic Biology.

They also provide cell phenotyping with the Berkeley Light Beacon, first instrument installed in an academic institution in Europe. This high-throughput automated cell selection platform allows high-capacity single cell cloning, analysis, outgrowth and RNA library preparations, with FDA-approved verified clonal origin.

Genomics Platforms:

- DNA assembly
- NGS library miniaturisation (in development)
- Sequencing with Oxford Nanopore MinION and bespoke software analysis of data
- Berkeley light Beacon for cell phenotyping and RNA-seq library prep
- Advanced robotics platform
- Bioinformatics and software development