Practical – Genomics England Embassy

Getting at the Data

Matthew Parker
Lead Bioinformatician
Sheffield Diagnostic Genetics Service

Sheffield Children's NHS
NHS Foundation Trust

Information Governance

Sheffield Children's NHS
NHS Foundation Trust

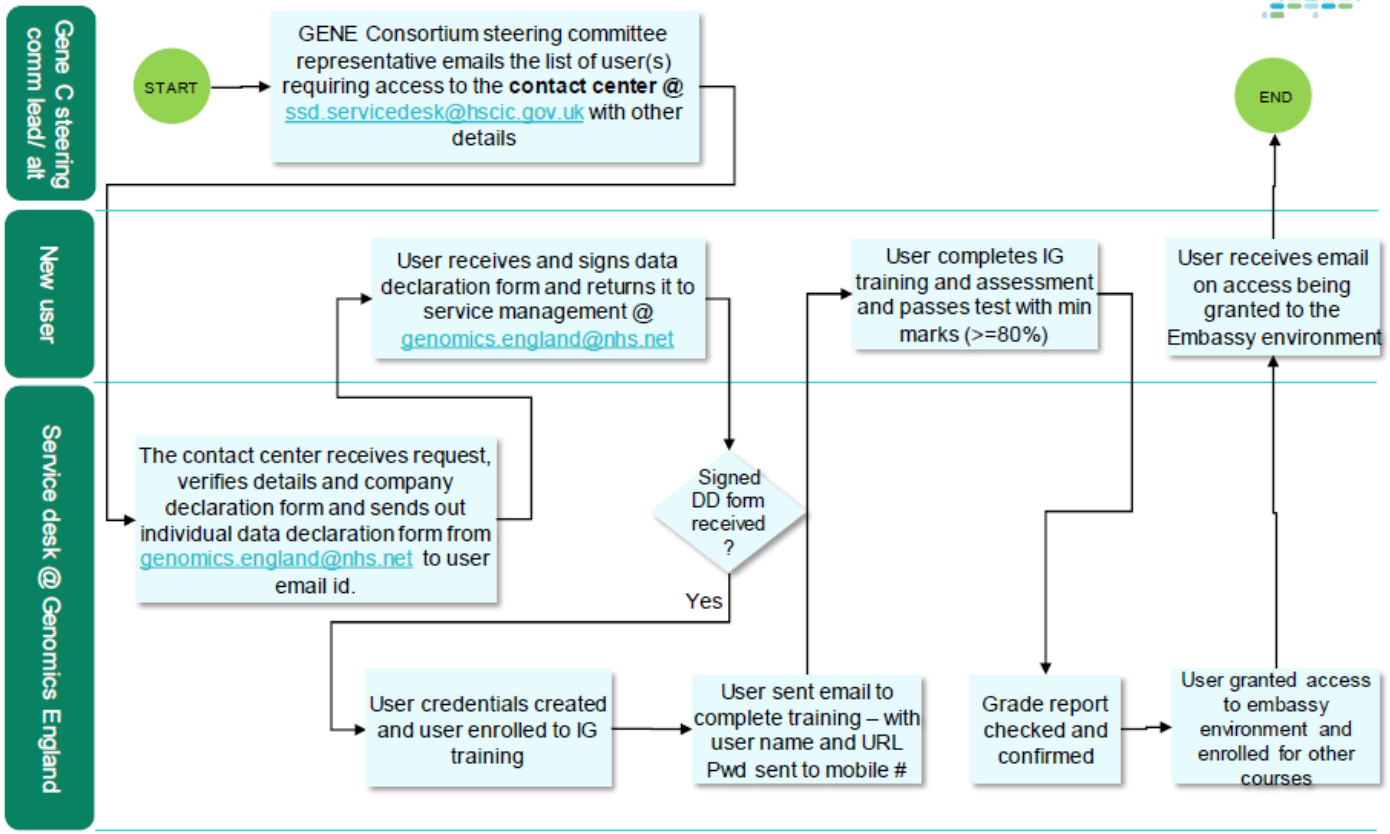
- Patient confidentiality and protection is a key cornerstone of the 100,000 Genomes Project.
- individual level data will not be "released" but will instead be analysed within a secure, monitored environment akin to a reading library
- To protect patient confidentiality access to this environment will be reviewed and granted only for specific, approved purposes in accordance with informed consent and the ethical scope of the Protocol, and any attempted usage beyond the specified purpose may lead to exclusion and possible legal action, supported by automatically-generated electronic evidence.

The Embassy

- Secure area to query GEL repository, store results etc
- Browser based virtual desktop interface
- Access controlled by GEL
- Not everyone has access to the embassy: we'll talk through how to get access and the structure of the data
- The education embassy is currently very limited – not many tools or much data is available

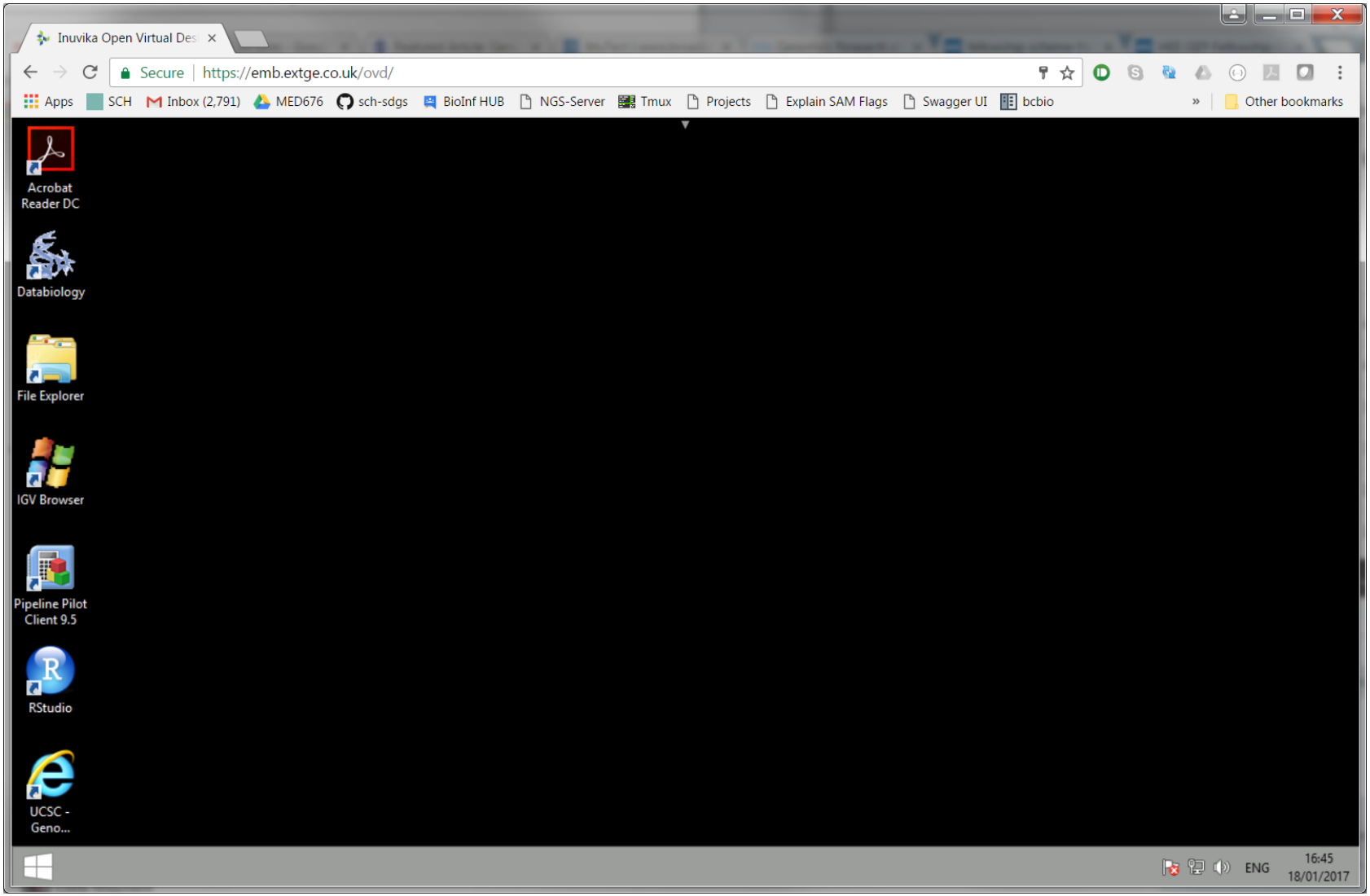
New user Access request process

Genomics england



Virtual Desktop

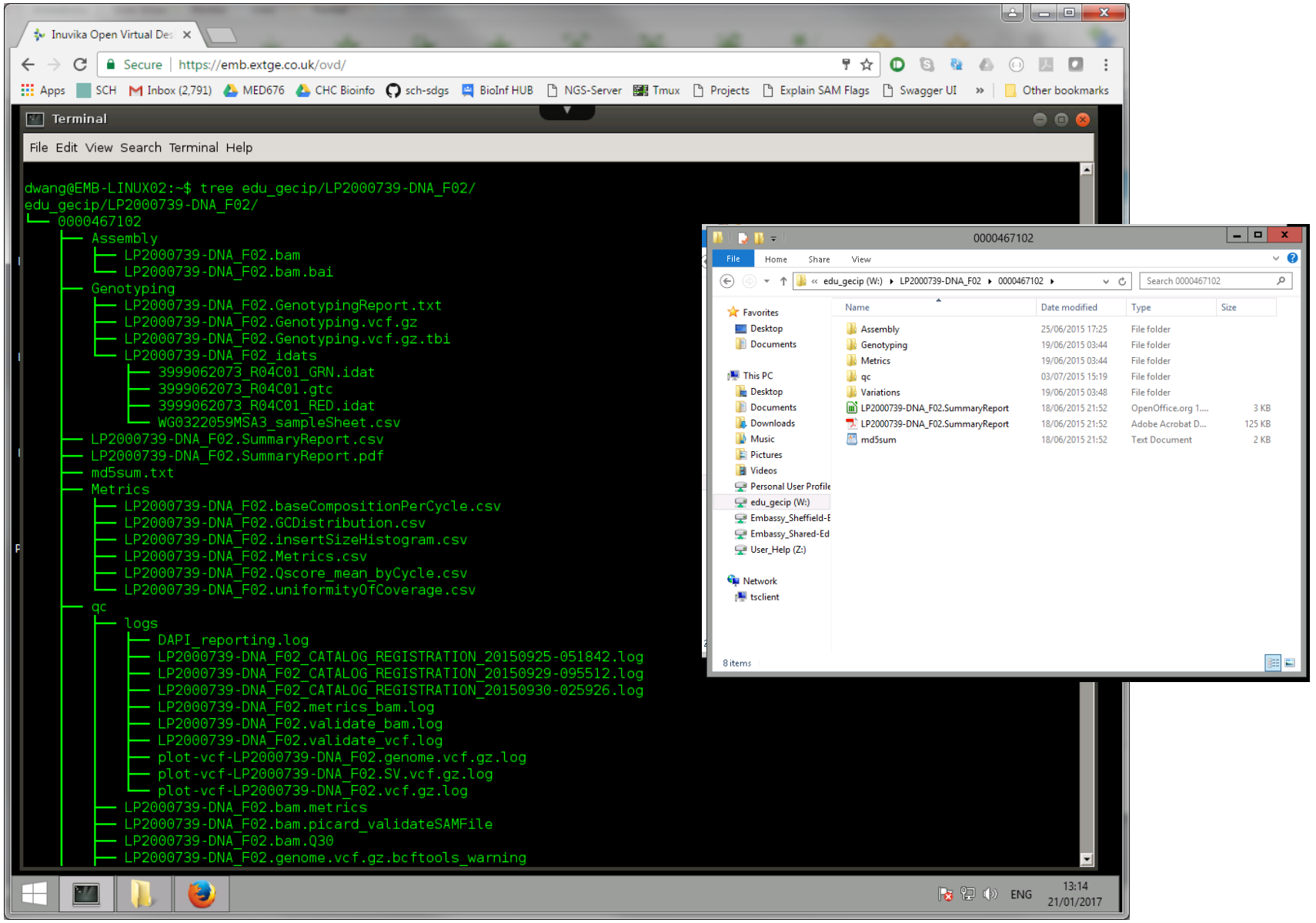
Sheffield Children's NHS
NHS Foundation Trust



https://emb.extge.co.uk/ovd/

Structure of the Data

Sheffield Children's NHS
NHS Foundation Trust



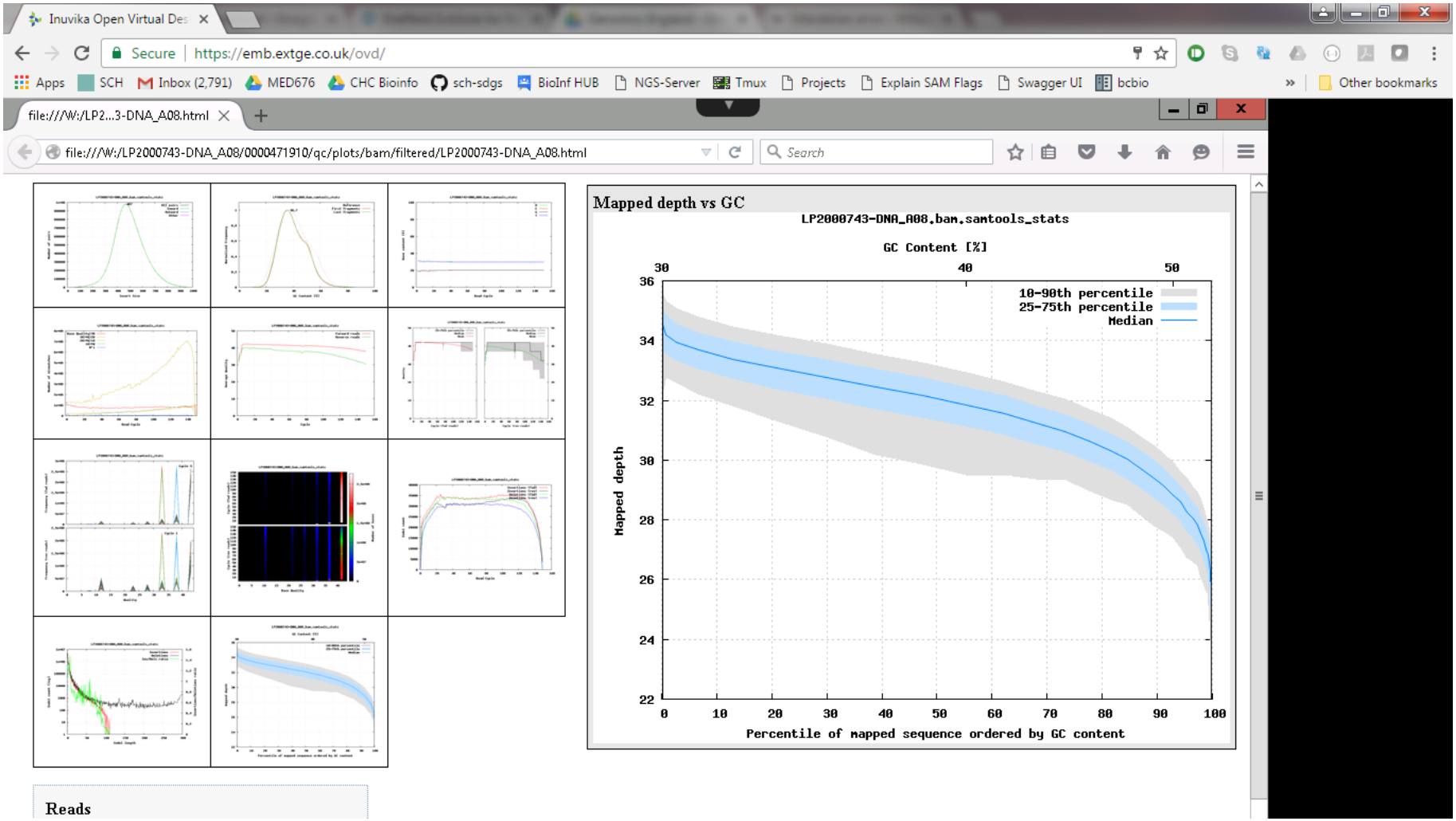
Samtools Stats Plots

Sheffield Children's NHS
NHS Foundation Trust

- Samtools – tool for interrogating sequence files – also includes ability to generate stats on these files
- W:\LP2000743-DNA_A08\0000471910\qc\plts\bam\filtered\LPS000743-DNA_A08
- Low level qc of the sequencing run

QC

Sheffield Children's NHS
NHS Foundation Trust



IGV Viewing BAM & VCF

Sheffield Children's NHS
NHS Foundation Trust