Genotyping of 500,000 UK Biobank participants

Description of sample processing workflow and preparation of DNA for genotyping

Version 2.0
Date: 11/07/2017
Prepared by: Samantha Welsh

Prepared for: UK Biobank Showcase

Version	Date	Amendments
1.0	27/04/2015	Original document for release of genotype data from 150,000 participants
2.0	11/07/2017	Updated for release of genotype data from 500,000 participants; includes selection of repeats and updated links to Thermo Fisher Scientific and Wellcome Trust Centre for Huiman Genetics documentation

Contents

1.0 Introduction

UK Biobank has undertaken a project to genotype all 500,000 participants. The majority of participants (~450k) are genotyped on the UK Biobank Axiom® Array[1], with 50,000 participants genotyped on the Affymetrix UK BiLEVE Axiom® array[2], which has >95% content overlap with the UK Biobank Axiom® Array. Preparation of the DNA was performed at UK Biobank, genotyping at Affymetrix[3] and quality control of the data at the Wellcome Trust Centre for Human Genetics (WTCHG). The Clinical Trial Service Unit (CTSU) at the University of Oxford is responsible for the storage and distribution of the genotyping data (and all other UK Biobank data) to researchers.

The present document is intended to;

- provide an overview of the sample and data processing workflow – Section 2.0
- describe the preparation of DNA at UK Biobank for genotyping at Affymetrix – Section 3.0
    - Buffy coat (850 µl) was recovered from a 10 ml EDTA vacutainer and stored in an automated sample archive at -80°C
    - Buffy coat aliquots were retrieved from storage using an algorithm to avoid introducing bias
    - DNA was extracted on an automated extraction system
    - Quality of the DNA was assessed prior to shipping.
- provide a summary of the genotyping sample and data workflow at Affymetrix – Section 4.0
- provide a summary of the quality control at WTCHG – Section 5.0

Documents generated by Affymetrix and WTCHG provide further detail on the genotyping sample and data processing and data quality control, respectively (Table 1).

Table 1: Summary of additional documents describing genotyping sample processing and data quality control

Document Name	Description of contents	Reference
UKB_WCSGAX: UK Biobank 500K Samples Genotyping Data Generation	Description of data analysis methods that differ from the standard Affymetrix data analysis methods described in (1)	(2) Thermo Fisher Scientific. UKB_WCSGAX: UK Biobank 500K Samples Genotyping Data Generation by the Affymetrix Researcher Services Laboratory. [Online] June 2017.
UKB_WCSGAX: UK Biobank 500K Samples Processing	Description of sample processing in relation to the genotyping on the Axiom array.	(3) Thermo Fisher Scientific. UKB_WCSGAX: UK Biobank 500K Samples Processing by the Affymetrix Researcher Services Laboratory. [Online] June 2017. .
Genotyping and quality control of UK Biobank, a large-scale, extensively phenotyped, prospective resource	Description of quality control measures undertaken at WTCHG and imputation methodology.	(4) Wellcome Trust Centre for Human Genetics. Genome-wide genetic data on ~500,000 UK Biobank participants. [Online] 2017.
Comparison of DNA quantification methodology used in the DNA extraction protocol for the UK Biobank cohort	Description of laboratory processing and comparison between quantification at UK Biobank and Affymetrix	(5) Welsh, Samantha, et al. Comparison of DNA quantification methodology used in the DNA extraction protocol for the UK Biobank cohort. 26, s.l. : BMC Genomics, 2017, Vol. 18

[1] The UK Biobank Axiom® Array was designed by leading researchers in the fields of epidemiology, human disease and population genetics for use by UK Biobank in the largest genotyping study to date. UK Biobank Axiom® Array has been designed using imputation-aware SNP selection and provides optimized content modules for genome-wide association studies (GWAS) of common and low-frequency variants, biological function, and human disease in populations of European and British ancestry. It includes over 820,000 variants, providing comprehensive coverage and includes rare coding variants, pharmacogenomics markers, copy number regions, HLA, inflammation and eQTL variants (11).

[2] UK BiLEVE was a project undertaken by the University of Leicester using the UK Biobank resource to investigate genetic variants in relation to Chronic Obstructive Pulmonary Disease (COPD). This project was completed prior to the genotyping of the 450,000 UK Biobank participants.

[3] Affymetrix are now part of Thermo Fisher Scientific

2.0 Summary of sample and data workflow

Figure 1 provides an overview of the sample and data processing workflow for the genotyping project with further detail obtained in the hyperlinked sections.

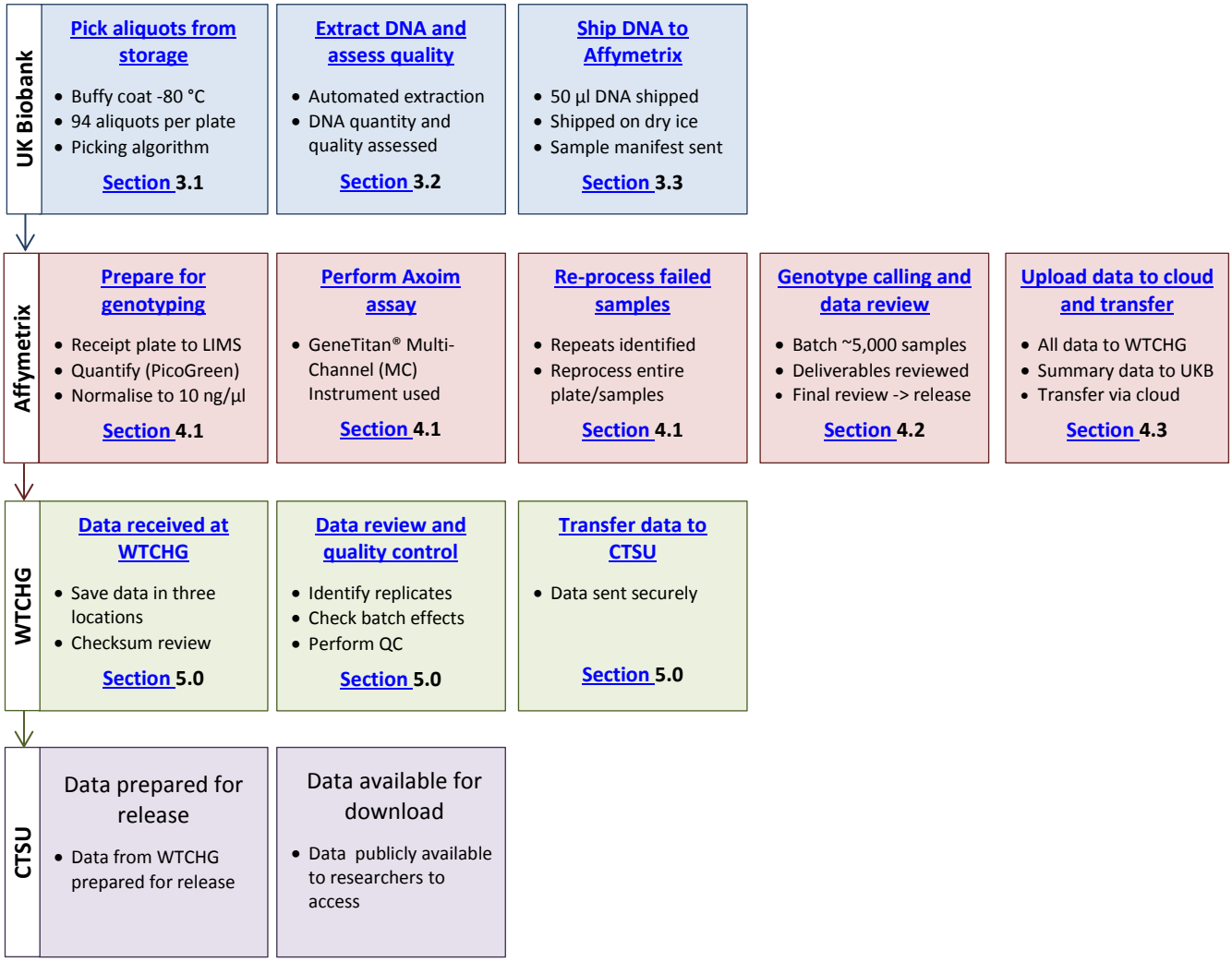

Figure 1: Workflow for UK Biobank Genotyping project

3.0 Sample processing at UK Biobank

UK Biobank was responsible for sample retrieval, DNA extraction and shipping of samples for genotyping. The following sections describe in detail the processes undertaken to prepare DNA for genotyping.

3.1 Pick buffy coat aliquots for DNA extraction

Buffy coat was picked from storage in preparation for DNA extraction.

3.1.1 Collection of buffy coat aliquots from participants

UK Biobank collected two 10 ml EDTA vacutainers per participant. These vacutainers were split into aliquots of 850 µl and the aliquots were split across two storage locations; -80°C and -196°C. Vacutainer one (EDTA1) was aliquotted as follows; three plasma and one buffy coat aliquot to -80°C plate, one plasma and one red blood cell aliquot to LN2 plate. Vacutainer two (EDTA2) was aliquotted as follows; three plasma aliquots to -80°C plate, buffy coat, one plasma and one red blood cell aliquot to LN2 plate (Figure 2).

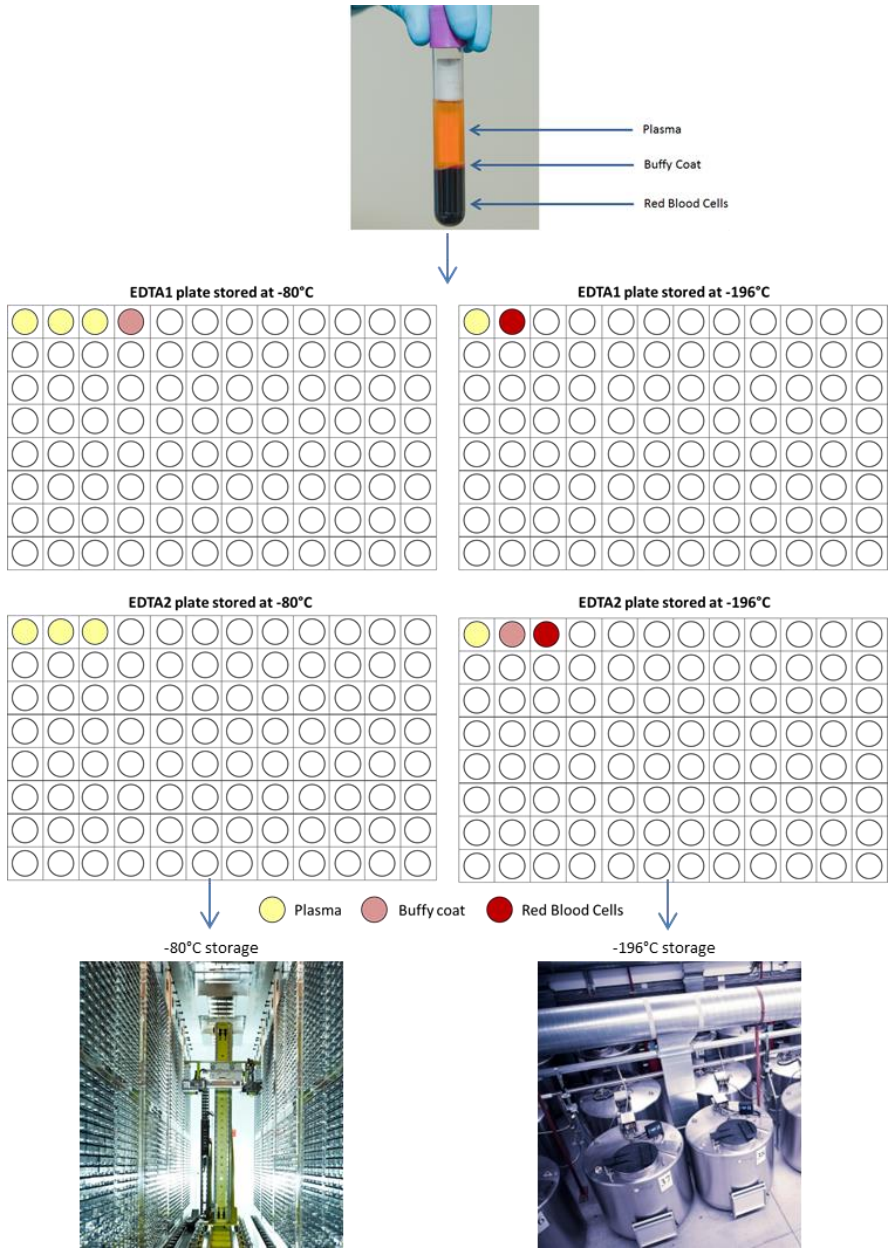

Figure 2: Aliquotting of EDTA vacutainers to generate aliquots for storage

3.1.2 General picking method

Generally, a destination plate containing 94 picked buffy coat aliquots was composed of buffy coat aliquots from four EDTA1 -80°C source plates (~24 buffy coat aliquots from each plate). Aliquots were picked from 'source plates' and placed on the 'destination plate' in rows. The final two positions (H11 and H12) were empty. The aliquots from A12 and E12 were moved to H11 and H12 to create spaces in A12 and E12 for the addition of controls at Affymetrix (Figure 3). The specific source plates selected to create a destination plate were identified via the picking algorithm to reduce clustering of phenotypes (Section 3.1.5).

Picked plates awaiting DNA extraction were stored at -80°C in semi-automated storage.

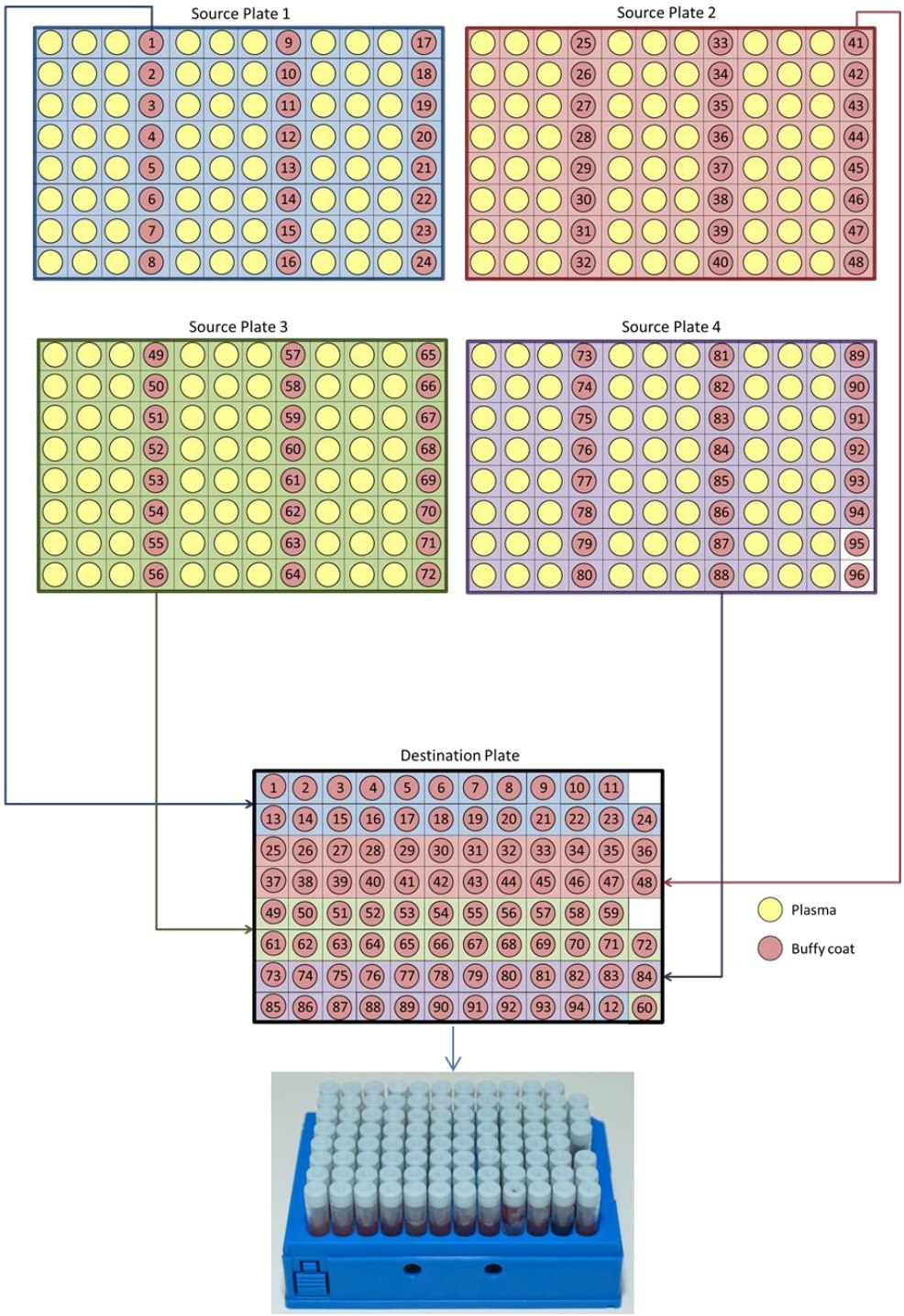

Figure 3: Picking process describing creation of a destination plate from multiple source plates

3.1.3 Picking for the UK BiLEVE project

The picking of samples for the UK BiLEVE project did not follow the picking algorithm developed at UK Biobank, because of the requirement for samples to be selected from a specific phenotype distribution (related to smoking history and lung function).

The UK BiLEVE project required four controls (instead of 2 controls for the UK Biobank project) to be added to the genotyping plates at Affymetrix, resulting in 92 aliquots picked to each plate.

3.1.4 Deviations from the general picking method

There were some deviations from the general picking method described in Section 3.1.2. These are summarised in Table 2.

Table 2: Summary of deviations to the general picking method

Scenario	Deviation
No buffy coat aliquot for a participant available in automated -80°C storage	The aliquot in -196°C was picked and transferred to the -80°C automated sample repository for selection (according to the picking algorithm). Where more than six aliquots on a plate were required from -196°C storage, the entire plate was transferred; where six or fewer aliquots were required from a plate the aliquots were picked. This meant more than four source plates were used to create a destination plate.
Aliquots picked from storage unsuitable for DNA extraction (e.g. low yield)	Unsuitable aliquot(s) removed from destination plate. A plate of aliquots already picked and checked were kept aside and used to fill the spaces on the plate with unsuitable samples to create a plate of 94 aliquots. A replacement strategy was implemented to seek the second buffy coat sample from any participants where the first sample selected was not suitable for extraction.
Requirement for extra space on plate for Blind Spike Duplicate	Approximately one plate in 70 was picked with 93 aliquots rather than 94 to allow for the addition of a Blind Spike Duplicate (added at the DNA extraction stage).

3.1.5 Repeated samples

To minimise the number of participants without genotype data, samples that were unable to provide a satisfactory genotype result from the first sample sent were included in a batch of repeats. DNA was extracted from a second buffy coat sample (where available) for all samples that met any of the following criteria;

- Sample status from genotyping = fail
- Gender mismatch between submitted gender (from UK Biobank records) and genotyped gender (inclusive of samples with suspected XXY genotype)
- Samples that failed Quality Control (QC) at WTCHG
- Samples with suspected sample tracking issue

Samples were shipped and included in the final two batches run at Affymetrix.

3.1.6 Picking algorithm

The algorithm developed to pick buffy coat aliquots from source plates to destination plates for DNA extraction (and subsequently genotyping) needed to fulfil two requirements;

1. **Pick all aliquots required from a source plate during a single visit** – To pick an aliquot, a plate must be retrieved, then the aliquot picked. Plate retrieval takes substantially longer than aliquot picking. To ensure aliquots were picked within the timeframe of the project, a plate must only be visited once. This meant four source plates (each containing ~24 aliquots needed in the project) were required to generate a single destination plate.

2. **Prevent clustering of assessment centre on a plate sent for genotyping** – Source plates generally have aliquots from one or two assessment centres, one or two geographic regions and one collection date. Since the number of source plates per destination plate could not be changed, it was important the algorithm selected source plates from different assessment.

An overview of the picking algorithm is described below:

1. Assess the proportion of the total aliquots contributed by each assessment centre;
2. Pick one aliquot (and therefore plate) at random from a randomly selected assessment centre (seed) and pick all aliquots from that plate;
3. Assess the new proportions and select the assessment centre whose recalculated proportion has the greatest positive deviation from its respective baseline proportion (1), and select one aliquot (and therefore plate) at random from this assessment centre;
4. Repeat from 2 until all aliquots are picked.

3.1.6.1 Assessing picking algorithm prior to project start

Simulations were run with the algorithm using the location of aliquots on plates prior to picking to ensure the algorithm would generate an acceptable spread of phenotypes per plate (Figure 4).

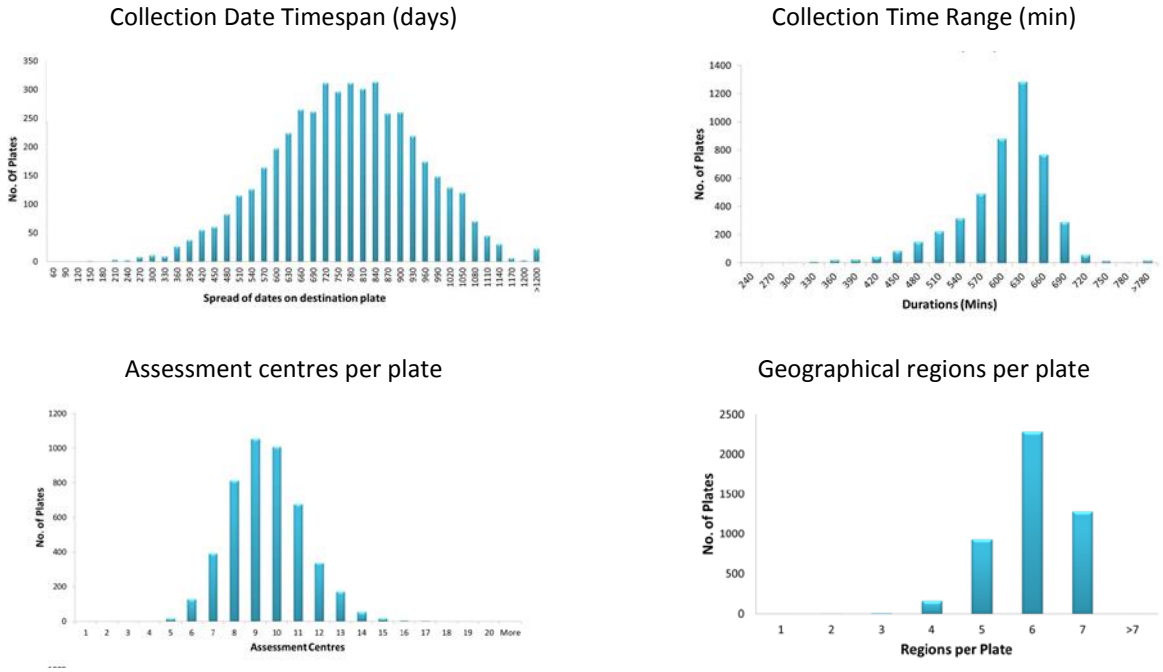

Figure 4: Results from simulation of picking algorithm

3.1.6.2 Assessing picking algorithm success during picking

A check was undertaken of the picked plates to ensure the picking algorithm developed had prevented clustering, as predicted in Section 3.1.6.1.

Three sets of 14 plates picked consecutively were included in the check. One set from the start, middle and end of the picking were selected to ensure the avoidance of clustering was maintained throughout the picking.

The following phenotypes were identified for each sample on the plate and a chi-square test performed to assess the distribution of characteristics across each plate – assessment centre, date of sample collection, time of day of sample collection, age, sex, smoking status, Townsend Score (are-based measure of socio-economic deprivation) and ethnicity. While the selection of samples is not random (some phenotypes have significant differences in distribution between plates), the algorithm does minimise clustering of the phenotypes between

plates, with the distribution of assessment centres across plates being similar to the predicted distribution (Figure 5).

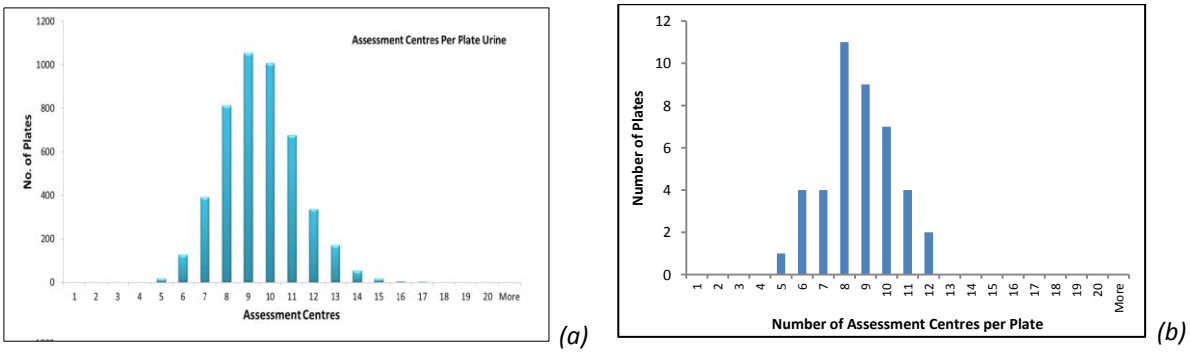

Figure 5: (a) Number of assessment centres predicted from algorithm; (b) observed number of assessment centres picked

3.2 DNA Extraction

DNA was extracted from 850 µl of buffy coat recovered from 10 ml EDTA whole blood. The entire aliquot was extracted to generate a total of 900 µl DNA. Three DNA aliquots were generated per participant; one 50 µl aliquot for genotyping, one 425 µl aliquot for storage at UK Biobank co-ordinating centre and one 425 µl aliquot at the UK Biobank back-up facility (Figure 6).

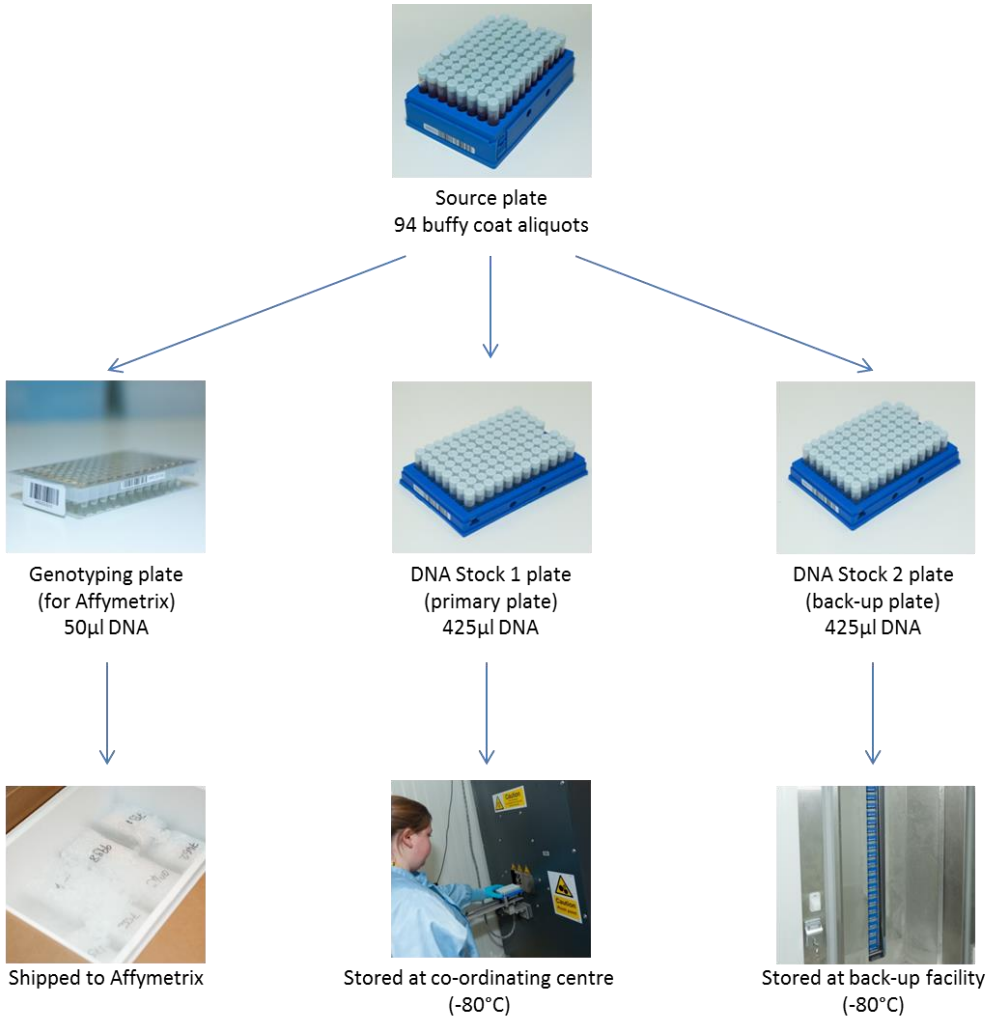

Figure 6: Overview of process to generate DNA plates for genotyping and storage for future projects

3.2.1 DNA extraction system overview

A custom DNA extraction system (Figure 7) designed to extract DNA from the relatively large volume of buffy coat was available for use for the DNA extraction. An additional two systems were purchased to achieve the desired throughput for this project. The UK Biobank DNA Extraction Systems incorporate Ziath® DataPaq High Speed Single Rack Scanner ZTS-A6 for sample tracking, two modified hemi-skinned Promega Maxwell® 16's for DNA extraction, Trinean DropSense® 96 for DNA quantification, Brooks Tube Auditor™ for volume measurement and Liconic STX44 Automated Incubator to keep DNA at 4°C after extraction. Further detail on the DNA extraction system is available elsewhere (6).

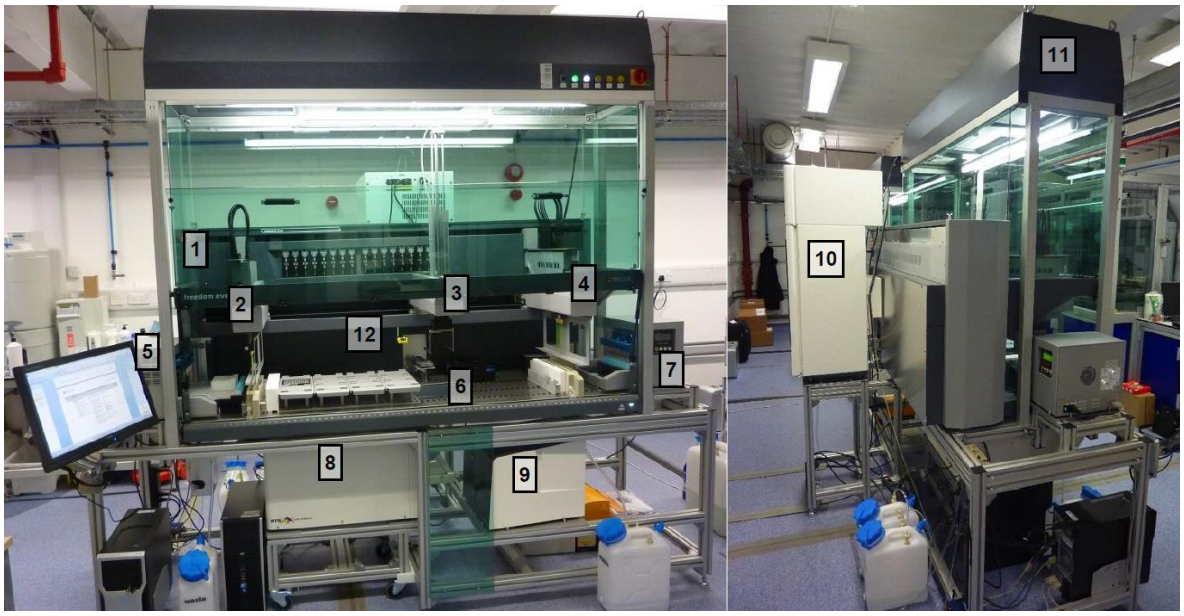

Figure 7: UK Biobank DNA Extraction System

Comprised of: 1. Tecan® Freedom EVO® 200; 2. Liquid Handling arm equipped with 8 fixed tips; 3. Long Robotic Manipulator with eccentric grippers; 4. Liquid Handling arm equipped with 8 fixed tips; 5. Promega® Maxwell® 16 AS2000-HS; 6. 2D BC reader- Ziath® DataPaq High Speed Single Rack Scanner ZTS-A6; 7. Promega Maxwell 16 AS2000-HS; 8. Brooks® Tube Auditor; 9. Trinean® DropSense® 96 microplate reader; 10. Liconic® STX44 deep freezer; 11. HEPA customized laminar flow hood; 12. 2D BC reader - Datalogic® Matrix 200 2D

3.2.2 DNA extraction process overview

Figure 8 describes the process to extract a plate of 94 buffy coat aliquots on the DNA extraction systems. Prior to being loaded onto the DNA extraction system, buffy coat was thawed at room temperature for 2-3 hours (h).

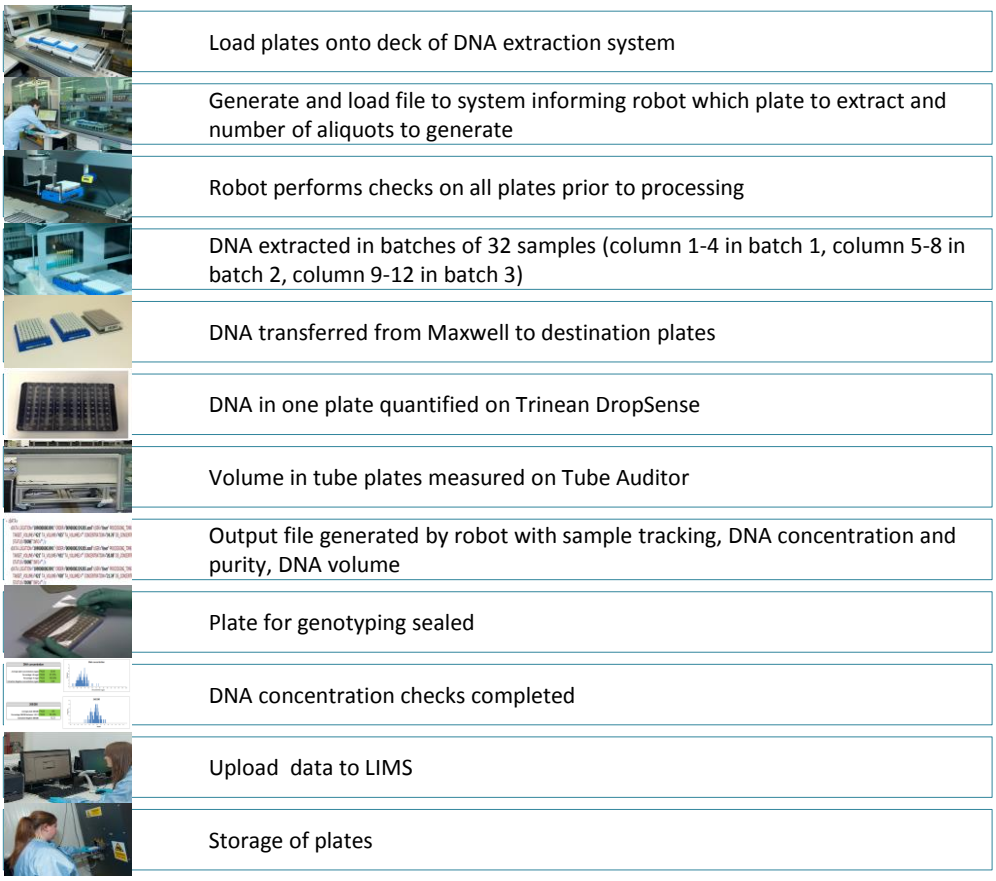

Figure 8: Overview of DNA extraction process on automated system

Within the run, three batches of 32 samples are completed on the Maxwell® 16's (16 samples per Maxwell® 16). An extraction negative was run with each batch in position A12 (Figure 9). A single aliquot of the extraction negative is retained at UK Biobank on the Stock 1 plate.

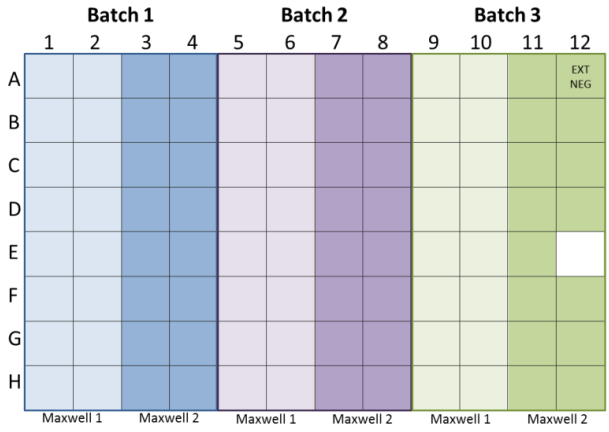

Figure 9: Plate layout displaying the extraction of four columns in a batch and two columns per Maxwell system (e.g. columns 1-4 extracted during batch 1; columns 1-2 extracted on Maxwell 1 (left Maxwell) and columns 3-4 extracted on Maxwell 2 (right Maxwell))

3.2.3 DNA extraction chemistry

DNA was extracted using the Maxwell® 16 Instrument (Promega) and the Maxwell® 16 Blood DNA Purification Kit (Promega – AS1010X). In this method, DNA is purified using paramagnetic particles which provide a mobile solid phase to capture, wash and elute DNA. Reagents for extraction are contained within a cartridge containing seven wells. An elution tube is separate to the cartridge and contains elution buffer. The elution tubes sits in a heated block to assist elution. Magnetic particle handlers (with disposable plungers) transport the magnetic particles through the purification reagents (Figure 10).

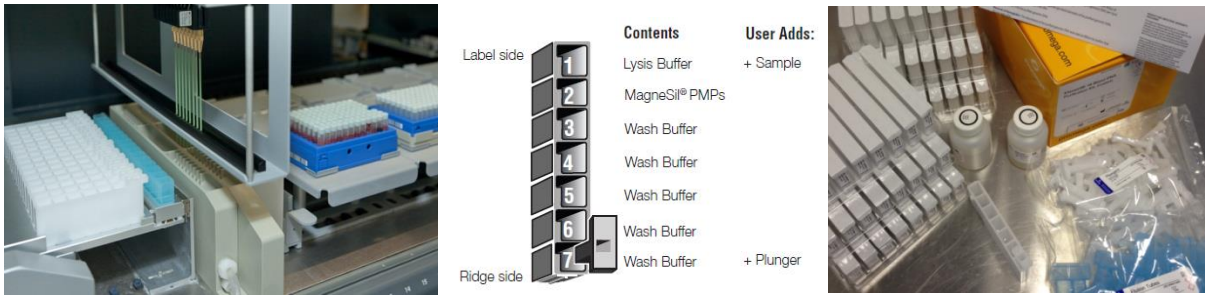

Figure 10: DNA extraction chemistry used to extract DNA from 850 µl of buffy coat

Cells are lysed by guanidine-based lysis buffer along with mechanical lysis from the plungers. DNA adsorbs to silica coated magnetic beads and is moved through a series of wash wells by a plunger. Salts from lysis and other impurities that may inhibit downstream processing (e.g. haem, proteins etc) are removed during the washing. Once purified, DNA is eluted from the beads in TE-based lysis buffer, which is assisted by heating. The TECAN freedom evo pipettes buffy coat to the Maxwell® 16 and pipettes DNA from the Maxwell® 16 once DNA extraction is completed.

3.2.4 DNA extraction for the UK BiLEVE project

DNA extraction for the UK BiLEVE project used the automated DNA extraction systems for a small number of aliquots but the system was still being commissioned for much of the project; many of the samples were processed manually using the Maxwell® 16 Instrument (Promega) and the Maxwell® 16 Blood DNA Purification Kit (Promega – AS1010X).

Samples were processed in batches of 16, as shown in Figure 9. This involved an operator and a witness observing the transfer of each buffy coat sample to the extraction cartridge and transfer of the DNA to the stock DNA plate and genotyping plate. Upload to the UK Biobank LIMS was completed by the operator (with a witness) scanning each tube to the location on the destination plate.

3.2.5 Modifications to extraction chemistry

The standard Promega Maxwell® 16 Blood DNA Purification Kit (AS1010) is designed to extract DNA from up to 250 µl buffy coat. Modifications were made to the extraction kit and Maxwell® 16 protocol to obtain an acceptable DNA yield and sample quality from 850 µl of buffy coat. The volume of lysis buffer (part # A826E) was increased by 600 µl and the wash buffer (MD1412) in two of the five wash wells increased by 1 ml. The elution volume used was 1.2 ml.

An additional pass through the Maxwell® 16 was included to ensure a purified sample. This increased the run on the Maxwell® 16 to 1 h 14 min.

3.2.6 Acceptance of DNA extraction chemistry

Each lot number of Promega Maxwell® 16 Blood DNA Purification Kit (AS1010X) was tested to ensure it was free of contamination and yielded DNA of a suitable concentration and purity. DNA was extracted from three positive controls and the concentration obtained was compared to the expected concentration range.

3.2.7 Quantification checks

The DNA generated was quantified using the Trinean DropSense® 96. The absorbance at 260 nm and 260/280 ratio was imported into the output files from the TECAN and uploaded to the UK Biobank LIMS.

Prior to upload to the UK Biobank LIMS, the concentration (measured by absorbance at 260 nm) and DNA purity (measured by 260/280) metrics of a plate were assessed (Figure 11) for conformance to a set of criteria;

- Average plate concentration at least 20 ng/µl
- Average 260/280 between 1.8-2.2
- At least 80% of samples concentration greater than 10 ng/µl
- At least 80% of samples 260/280 between 1.8-2.2
- The extraction negative concentration less than 1 ng/µl

Deviation from the above criteria was flagged to the Project Manager and an action plan agreed. Generally, when the average 260/280 was >1.7 the plate was shipped, regardless of the percentage of samples with a 260/280 between 1.8 and 2.0. Where the failure to meet the above criteria was due to an issue relating to a batch of samples (maximum of 32 processed during a batch – refer to Figure 9), the plate was not shipped and the samples on the plate without a processing issue were reformatted to generate a full plate for shipment. The second buffy coat aliquot was sought for the samples that failed extraction and were processed as replacements, as described in Section 3.1.4.

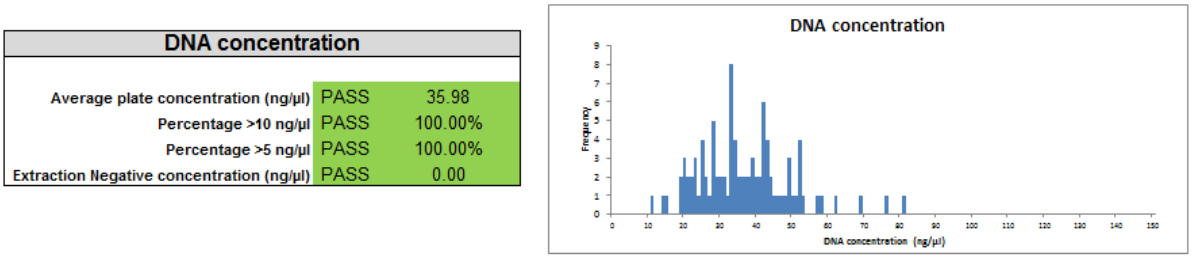

Figure 11: Example of plate check for conformance to DNA quantification criterion

The concentration and 260/280 values were imported to the UK Biobank LIMS by uploading the export file from the DNA extraction system. This data will be made available in the UK Biobank Data Showcase.

3.2.7.1 Quantification checks for the UK BiLEVE projects

At the request of Affymetrix, no samples were sent with a DNA concentration less than 5 ng/µl for the UK BiLEVE project. When a sample was quantified at less than the threshold, the second aliquot from the participant was picked, extracted and included in the genotyping plate, where possible. Where a second aliquot was not available, a replacement from the list provided by the researchers was sought.[4]

[4] Samples with low concentrations were found to perform well in the UK BiLEVE project (<10 ng/µl) and therefore were included in the UK Biobank project.

3.2.8 Volume checks

The volume of DNA in the stock aliquots retained by UK Biobank was accurately determined using a Brooks Tube Auditor™[5]. This ensured that where less than the desired volume of DNA was generated, the LIMS was updated with the actual volume reducing the likelihood of being unable to fulfil sample request for future projects.

3.2.9 Storage of plates

Three plates were generated during DNA extraction; genotyping plate containing 50 µl DNA, stock DNA aliquot for storage at co-ordinating centre containing 425 µl and stock DNA for storage at UK Biobank back-up facility. Upon completion of DNA extraction, the genotyping plate was wrapped in a bubble wrap bag to protect the wells and placed in a -80°C freezer. The stock DNA plate was paced in automated -80°C storage and the back-up plate was temporarily stored in a -80°C and approximately weekly transferred to the back-up site.

3.3 Shipping of DNA from UK Biobank to Affymetrix

Approximately 6,500 DNA samples were shipped weekly from UK Biobank, Stockport, UK to the Affymetrix Research Services Laboratory (ARSL), Santa Clara, CA. For the UK BiLEVE project, approximately 2,000 samples were shipped per week. Samples were shipped on dry ice to ensure the samples remained frozen during transit.

3.3.1 Sample tracking between UK Biobank and Affymetrix

A sample manifest was generated from the UK Biobank LIMS recording the participant identifiers in each of the wells on the plates within the shipment. The sample manifest was sent to CTSU to replace the UK Biobank participant identifiers to a project-specific participant identifier and to add phenotypic data (gender, ethnicity and geographical location for the participant[6]) needed for quality control purposes at Affymetrix.

[5] Volume measurement was not performed on all samples due to upgrades of the Brooks Tube Auditor™ during the project. Aliquots generated when the Brooks Tube Auditor™ was not in use were assigned a volume based on the requested volume and amended if a visual inspection identified significant deviation from the volume.

[6] Fields included in relation to geographical location comprised: UK Biobank assessment centre (data-field 54); Place of birth in UK - east co-ordinate (130); Place of birth in UK - north co-ordinate (129); Country of birth (UK/elsewhere) (1647); Country of Birth (non-UK origin) (20115); Home area population density - urban or rural (20118); Home location at assessment - east co-ordinate (rounded) (20074); Home location at assessment - north co-ordinate (rounded) (20075)

4.0 Sample processing at Affymetrix

Affymetrix were responsible for genotyping on the UK Biobank Axiom® Array (7), generating the genotype data and performing initial QC of the data. Full details are available elsewhere of the sample processing specific to the UK Biobank project (3) and the Axiom array (8).

4.1 Genotyping workflow

The samples were processed on the UK Biobank Axiom® Array according to the Affymetrix Axiom processing standard procedure (on the GeneTitan® Multi-Channel (MC) Instrument). The workflow is outlined in Figure 12.

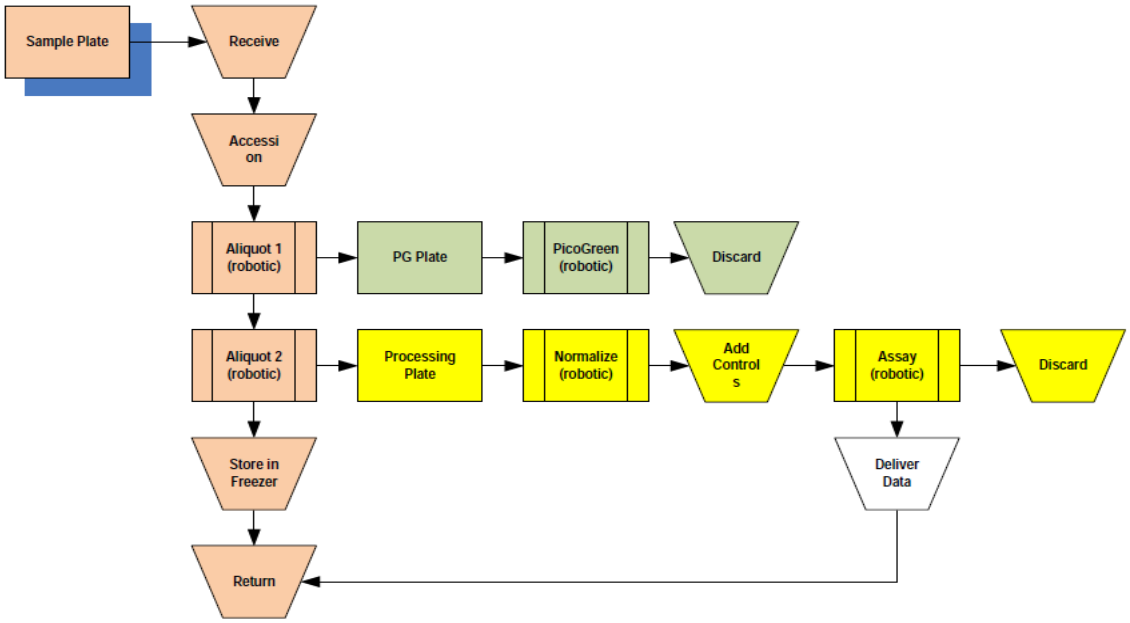

Figure 12: Sample handling workflow at Affymetrix (3)

4.2 Data analysis

Data analysis was performed according to the Best Practices in the Axiom® Genotyping Solution Data Analysis Guide (1). Exceptions to the standard protocols are documented in the Affymetrix UKB_WCSGAX Genotype Data Generation note (2).

Upon generation of the deliverables, a second analyst reviewed the deliverables and a final review completed prior to making the data available to WTCHG and UK Biobank.

Data was delivered in batches of approximately 5,000 samples. Due to reprocessing of samples, data was not always delivered for all samples on a plate in the same batch.

4.3 Transmission of data from Affymetrix

The data was delivered to WTCHG and UK Biobank via a secure cloud (3). UK Biobank was provided with summary information on the genotyping (call rate, dQC, PicoGreen concentration) and WTCHG received all the data generated (3).

5.0 Quality control at WTCHG

The WTCHG, University of Oxford, were responsible for quality control of the data prior to release into the UK Biobank resource for all 500,000 participants (inclusive of the UK BiLEVE samples), performed by Professor Peter Donnelly's group(Dr Colin Freeman, Dr Desislava Petkova and Dr Clare Bycroft).

The processes undertaken in relation to the quality control at WTCHG are documented fully elsewhere (4). In summary, poorly performing SNPs and samples to exclude were identified. The relatedness of samples was assessed and SNP plate/batch effects were considered.

6.0 Bibliography

1. **Thermo Fisher Scientific.** Axiom Genotyping Solution Data Analysis Guide. [Online] http://media.affymetrix.com/support/downloads/manuals/axiom_genotyping_solution_analysis_guide.pdf.

2. **Thermo Fisher Scientific.** UKB_WCSGAX: UK Biobank 500K Samples Genotyping Data Generation by the Affymetrix Researcher Services Laboratory. [Online] June 2017.

3. **Thermo Fisher Scientific**. UKB_WCSGAX: UK Biobank 500K Samples Processing by the Affymetrix Researcher Services Laboratory. [Online] June 2017.

4. **Wellcome Trust Centre for Human Genetics.** Genome-wide genetic data on ~500,000 UK Biobank participants. [Online] 2017.

5. **Welsh, Samantha, et al.,** *Comparison of DNA quantification methodology used in the DNA extraction protocol for the UK Biobank cohort.* 26, s.l. : BMC Genomics, 2017, Vol. 18.

6. **TECAN.** Blood DNA extraction and DNA normalization - 28.11.2014. *TIG - Downloads - Genomics.* [Online] 28 November 2014. http://www.tecan.com/platform/apps/product/index.asp?MenuID=5019&ID=11651&Menu=1&Item=21.1.2.7.4.

7. **UK Biobank.** UK Biobank Axiom Array - Content Summary. [Online] 2015. http://www.ukbiobank.ac.uk/wp-content/uploads/2014/04/UK-Biobank-Axiom-Array-Content-Summary-2014.pdf.

8. **Affymetrix.** Axiom 2.0 Assay Automated Workflow. [Online] 2015. http://media.affymetrix.com/support/downloads/manuals/axiom_2_assay_auto_workflow_user_guide.pdf.

9. **UK Biobank.** UK Biobank Genetic Enclave. [Online] 2015. http://biota.osc.ox.ac.uk/.

10. **Tobin, Martin.** UK Biobank Lung Exome Variant Evaluation. [Online] 26 June 2014. http://www.ukbiobank.ac.uk/wp-content/uploads/2014/06/1225-Tobin-Church-Thurs.pdf.

11. **UK Biobank.** UK Biobank Axiom Array. [Online] http://www.ukbiobank.ac.uk/scientists-3/uk-biobank-axiom-array/.

7.0 Glossary

Name	Description
260/280	Ratio of absorbance from a DNA sample at 260 nm divided by absorbance at 280 nm. This is frequently used as a measure of DNA purity.
Aliquot	Fraction of sample collected from participant, e.g. buffy coat aliquot is an 850 µl fraction from 10 ml EDTA vacutainer.
Blind Spiked Duplicate	Sample placed on plate shipped to Affymetrix already included on another plate. The purpose of this sample was as a control measure to ensure sample tracking was maintained and to assess reproducibility.
CTSU	Clinical Trials Service Unit, Oxford University. Responsible for receipt and distribution of genotype data from WTCHG.
Destination plate	Output of picking. Plate contains 93 or 94 buffy coat aliquots ready for DNA extraction.
Genotyping plate	Semi-skirted PCR plate containing 50µl DNA. This plate was shipped to Affymetrix for genotyping.
Promega Maxwell® 16	DNA extraction platform capable of extracting 16 samples per run (1 h 14 min). Two modified Promega Maxwell® 16 systems were integrated into each UK Biobank DNA Extraction System to enable 32 samples to be extracted in 1 h 14 min.
Source plate	Plate containing buffy coat aliquots required for DNA extraction, amongst other aliquots. Buffy coat aliquots are picked from the source plate to a destination plate.
Trinean DropSense 96	UV/vis-based plate reader used for quantification of DNA. Integrated into UK Biobank DNA Extraction System
UK BiLEVE	UK Biobank Lung Exome Variant Evaluation project. Access project applied to use UK Biobank resource to genotype 50,000 UK Biobank participants on the Affymetrix UK BiLEVE Axiom array. Further detail of the project available from (10)
UK Biobank DNA Extraction System	Automated TECAN liquid handling robot integrated with Promega Maxwell® 16, Trinean DropSense® 96 for DNA quantification, Brooks Tube Auditor™ for volume measurement and Liconic STX44 Automated. System is used to extract DNA, quantify DNA and measure volume of DNA generated. An export file is generated which records the sample hierarchy, DNA concentration and volume of DNA generated. More detail (6)
Vacutainer	Collection tube containing blood donated by participant; typically 10 ml.
WTCHG	Wellcome Trust Centre for Human Genetics, Oxford University. The Centre is performing QC on the data after Affymetrix.