University of Southampton | Medicine

Genomic analysis of germline and somatic lesions at the *FCGR* locus; from reference genomes to clinico-biological implications

Sarah Frampton[1], Jade Forster[2], Konstantinos Boukas[2], Khiyam Hussain[3], Mohammadmersad Ghorbani[1], Helen Parker[1], Chantal Hargreaves[4], Dean Bryant[1], Stephen Beers[3], Mark S Cragg[3], Jane Gibson[1] & Jonathan C Strefford[1]

[1]Cancer Genomics, [2] WISH Laboratory, [3]Antibody and Vaccine Group, Cancer Sciences, University of Southampton, UK; [4] NIHR Oxford Biomedical Research Centre, Department of Clinical Immunology, John Radcliffe Hospital, Oxford, UK

Introduction

- Monoclonal antibody (**mAb**) immunotherapy is becoming standard in the treatment of cancer, particularly for B-cell lymphoma patients. However responses are variable and **relapse** is common, thus the molecular mechanisms that underpin resistance need to be defined.
- mAbs interact with the Fc gamma receptors (**FcγRs**), a family of cell surface molecules crucial in the immune response (**Fig 1**).
- Human cells express six FcγRs: the high-affinity FcγRI and 5 low-affinity receptors which have both **activatory** (FcγRIIa, FcγRIIc, FcγRIIIa and FcγRIIIb) and **inhibitory** (FcγRIIb) capabilities.
- Therapeutic mAb response is dictated by the activatory to inhibitory **(A:I) ratio** where unfavourable expression profiles can obstruct mAb efficacy.
- The **low-affinity locus** on chromosome 1, encodes the *FCGR2A*, *FCGR2B*, *FCGR2C*, *FCGR3A* and *FCGR3B* genes and is littered with single nucleotide polymorphisms (**SNPs**) that can influence protein function (**Fig 2**).
- In addition to germline regions of copy number variation (**CNV**), focal somatic amplification of *FCGR2B* have recently been uncovered in lymphoma patients and are associated with inferior outcomes.
- The locus contains **vast sequence homology** so is therefore very **challenging** to study as reads from traditional short-read sequencing technologies fail to be adequately aligned or assembled.
- For the first time, an **accurate genomic map** of the *FCGR* locus can be assembled with the use of Oxford Nanopore's **long-read** technology to elucidate high-resolution sequence, breakpoint, phasing and base modification information.

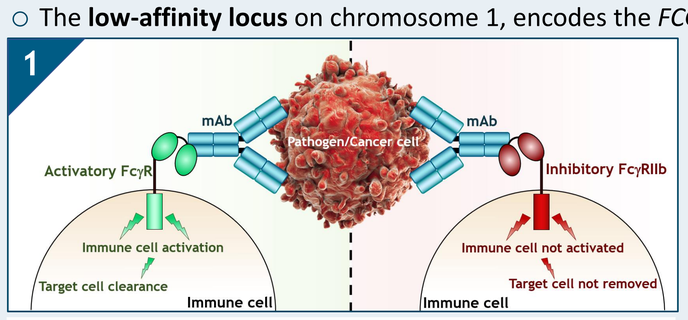


Figure 1 The FcγRs role in the removal of pathogens/malignant cells through the engagement with IgG/mAbs that have opsonised the target. The predominant mechanisms involve antibody-dependent cellular cytotoxicity, antibody-dependent cellular phagocytosis and complement-dependent cytotoxicity. Precise regulation of the activatory FcγRs vs the inhibitory FcγRIIb permits the flexibility and fine-tuning necessary for an appropriate immune response whilst avoiding autoimmunity.

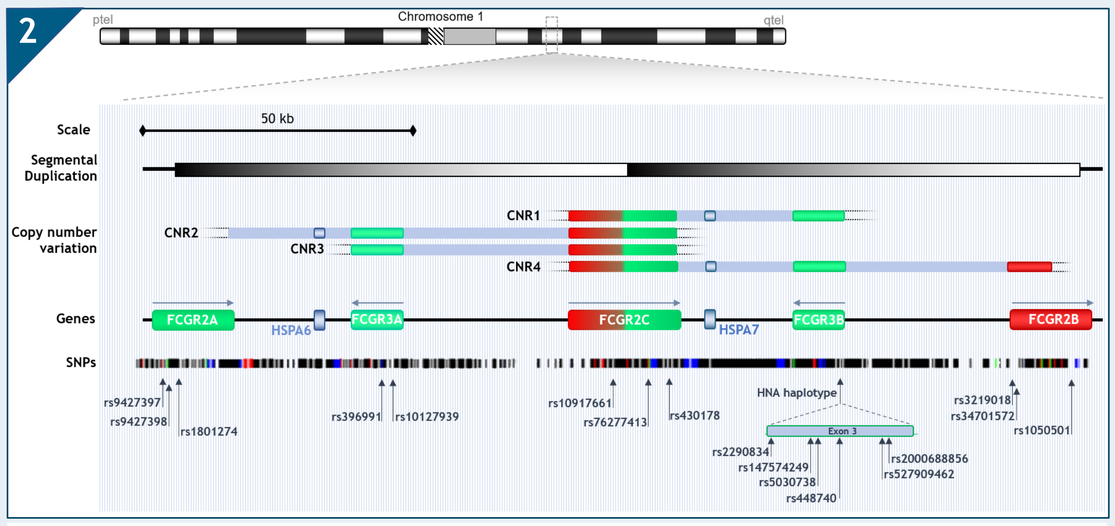


Figure 2 The current understanding of FcγR genomic architecture. The low-affinity FcγR genes are encoded in a 200kb locus on chromosome 1q23-24, which encompasses two tandem 85kb paralogous blocks derived from multiple ancestral recombination and segmental duplication events. *FCGR2C* is a fusion gene that encodes FcγRIIc, a receptor that shares intracellular activatory signalling similar to FcγRIIa but an extracellular domain almost identical to FcγRIIb. The locus is extensively polymorphic (missense SNPs in red, intronic in black and splice site SNPs in green) and SNPs relevant to the efficacy of mAb immunotherapy are highlighted. CNV consistently occurs in regions containing a series of genes, termed CNV regions (**CNR**s), but their precise boundaries are unknown.

Aims

1. Generate **highly-detailed accurate maps** of the low-affinity *FCGR* locus under different CNV states
2. Improve genomic/epigenetic understanding of *FCGR* region and it's **regulation**
3. Provide reference maps for **further omics** studies into **normal and malignant regulation** of the FcγRs

Methods

Plan:

- Stage 1 (**Fig 3**, green arrows) is to conduct whole genome sequencing (**WGS**) on 1 copy neutral sample to optimise library preparation and sequencing techniques.
- Stage 2 (**Fig 3**, magenta arrows) is to assess enrichment strategies to target the *FCGR* locus and apply the optimal method to the remaining cohort.

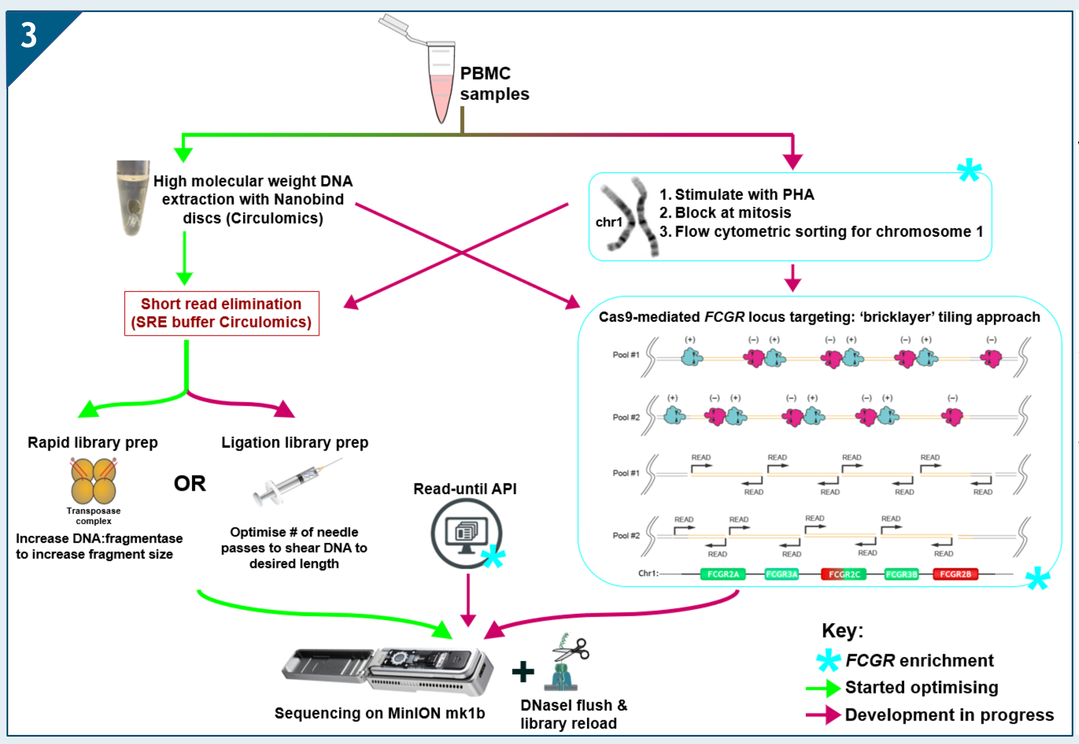


Figure 3 Experimental workflow. A cohort (n=271) of healthy donor peripheral blood mononuclear cell (**PBMC**) samples from volunteers entering National Blood service have all undergone **extensive molecular characterisation** with MLPA, Taqman genotyping and Haloplex^{HS} to best define CNV and SNP structures. For this project samples (**n=21**) have been selected that display a range of genomic complexity. Additionally both germline and **malignant** material is available from a fresh-frozen lymphoma cohort (**n=10**) to investigate somatic changes to the *FCGR* locus. Experimental approaches will provide different coverage across the locus depending on enrichment capacity. In stage 1 of the project only **WGS** has been conducted with the rapid library prep kit (workflow highlighted in green).

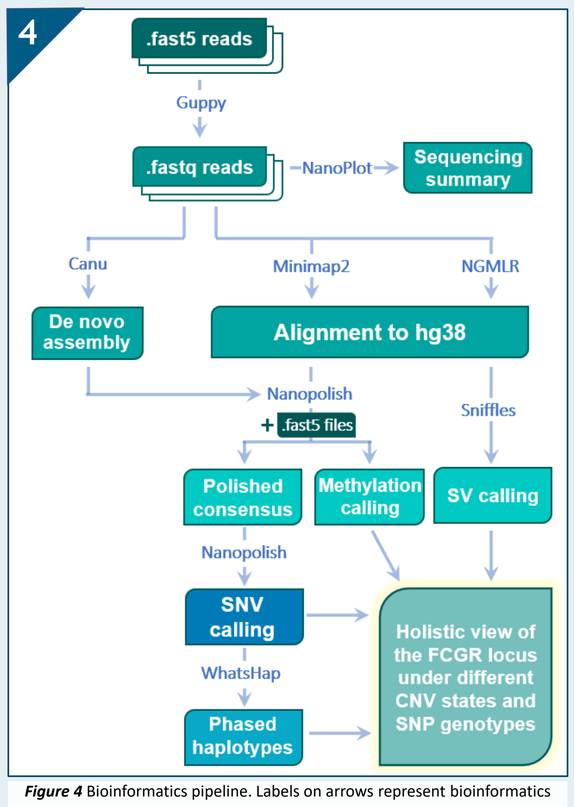


Figure 4 Bioinformatics pipeline. Labels on arrows represent bioinformatics tools. Tools for nanopore analysis are currently developing at a rapid pace and are therefore subject to change as new and improved algorithms are released.

Samples and Sequencing

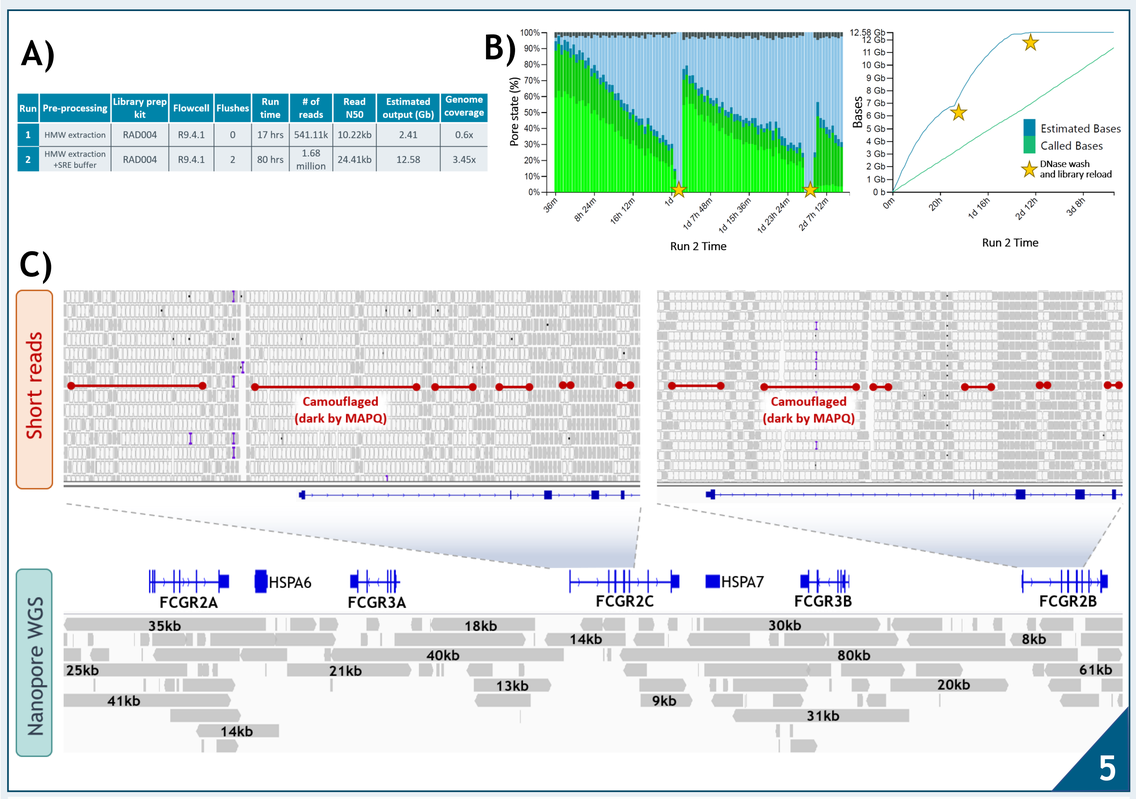
- Genetic material is available from healthy donors (**n=21**) and lymphoma patients (**n=10**)
- In stage 1, one sample was sequenced twice with a **MinION mk1b** device with a WGS approach (**Fig 3**).
- Both runs were with R9.4.1 flowcells but run 2 differed from run 1 by a size selective precipitation stage and DNaseI flushes/library reloads.
- In stage 2, we will assess methods for **enrichment** of the *FCGR* locus; using chr1 sorting, cas9 targeting and the **'Read Until' API** for adaptive sequencing to reject non-*FCGR* reads in real-time.
- As sequencing will only be performed on **native DNA**, additional insight into DNA modifications (primarily 5-methylcytosine and 5-hydroxymethylcytosine) will also be available.

Analysis Plan

- Sequencing data will be processed and analysed using a bioinformatics pipeline (**Fig 4**) to examine sequence polymorphisms, structural variation, breakpoint and phasing definition as well as epigenetics of the *FCGR* locus.

Preliminary Results and Future Work

- So far in the project only WGS has been conducted, showing the addition of **size selection** and **flowcell flush** steps is substantially beneficial to both read length and yield (**Fig 5a&b**).
- From our **preliminary results** of WGS it's clear nanopore long-reads have the capacity to **map uniquely** to the current human reference where short-reads have previously failed (**Fig 5c**).
- This facilitates the exciting opportunity to **overcome the inherent difficulties** of sequence homology and enables a comprehensive view of the *FCGR* genomic/epigenetic landscape.
- One expected outcome from this work is the generation of high-resolution reference maps of the low-affinity *FCGR* locus under different SNP genotype and CNV states in normal cells.
- Subsequently using the same approaches to **characterise primary lymphoma samples** we aim to generate a detailed understanding of the somatic lesions in the *FCGR* locus and their **functional impact**.
- This includes their structure, regulatory effects and their relationship with germline CNV in order to gain insight into the mechanisms by which the **FcγRs drive carcinogenesis**.
- Utilising the reference maps, we hope to be able to perform extensive **screening** of large **clinical trial lymphoma cohorts** using ddPCR to define the clinical significance of the spectrum of *FCGR* lesions.



Run	Pre-processing	Library prep kit	Flowcell	Flushes	Run time	# of reads	Read N50	Estimated output (Gb)	Genome coverage
1	HMW extraction	RAD004	R9.4.1	0	17 hrs	541.11k	10.22kb	2.41	0.6x
2	HMW extraction +SRE buffer	RAD004	R9.4.1	2	80 hrs	1.68 million	24.41kb	12.58	3.45x

Figure 3 Preliminary results from the WGS performed from stage 1. WGS with the rapid library kit (SQK-RAD004) is the simplest approach to set up from the workflow and these results will help guide the development of the enrichment and targeting approaches. A and B show summary metrics and plots from two consecutive sequencing runs. **A)** HMW represents high molecular weight and SRE short read eliminator. Genome coverage is calculated as an average across the whole genome (approx. 3.1Gb) using only the reads that have met the default quality threshold (7). **B)** demonstrates how flushing the flowcell with DNaseI and reloading library can further increase the throughput from one flowcell as it 're-activates' more pores for further sequencing (light green). Unfortunately there was insufficient library left after the second flush hence the low pore occupancy. **C)** shows viewings in IGV of the difference in mapping quality to the hg38 reference genome between short-reads and nanopore long-reads. Dark grey bars indicate sequencing reads with a high mapping quality (MAPQ) whereas white bars indicate reads with low a MAPQ, at the *FCGR* locus this is usually due to mapping equally well to multiple locations. Read lengths have been highlighted on some of the nanopore reads to give an indication of the range.

Discussion

Nanopore sequencing presents, for the first time, the possibility to overcome the inherent challenges of **the highly homologous** nature of the ***FCGR* locus**, facilitating a complete view of the region and it's regulatory control. This in turn will enhance further **multi-omics** work on the FcγRs by the wider scientific community, providing them with a unique resource that enables the interrogation of **gene expression** and **epigenetic modifications** that control the system. In broader terms, this work will catalyse a detailed understanding of FcγR **regulation** and ultimately improve **risk-adapted stratification** of cancer patients. Detailed knowledge of the molecular mechanisms that reinforce mAb resistance will promote the development of novel therapeutics to augment FcγR function and **improve response to mAb immunotherapy**.

CANCER RESEARCH UK

Acknowledgements
The authors gratefully acknowledge all of the patients and clinicians who contributed clinical material and information for this study.
This work was funded by Cancer Research UK (CRUK)

Contact
Sarah Frampton | PhD candidate
Email: smf1g18@soton.ac.uk
: @smframptonsoto1