CENTER *of* EXCELLENCE
for ENGINEERING BIOLOGY

Pilot Project Proposal

Name of Project:
Rational and Combinatorial Refactoring of Recombining Human Chromosomal Clusters

Proposer and Contact Information:
Tom Ellis and Paul Freemont, Imperial College London

Background:
The human genome contains several regions of research and healthcare importance where groups of genes with similar functions are clustered together along with sequences and motifs that give structural instabilities. As the chromosomes replicate and divide in mitosis and meiosis these areas are hotspots for genotypic diversity, where genes and other DNA sequences can be deleted, duplicated or rearranged. These structural rearrangements lead to diversity within the human population which is often problematic but can also provide advantageous in specific cases. Our work on the synthetic yeast genome project and its inducible SCRaMbLE system for genome rearrangement has shown us the power of genomic rearrangement in eukaryote. Dramatic new phenotypes can be generated in hours and new information about gene content and the roles of genomic sequence can be combinatorially determined, especially now with routine long-read DNA sequencing. As we scale from the Sc2.0 project to the human genome, it will be interesting to apply the SCRaMbLE tools to human chromosomal regions that naturally recombine and rearrange, so that we can learn the roles of sequence within these loci.

Technical Idea:
Working with UK funding agencies and healthcare research charities, we plan to synthesise two (or more) different megabase-scale regions of the human genome following the design principles of the Sc2.0 project and then investigate them using SCRaMbLE. We will remove all DNA sequences within these regions that are known or suspected to cause structural instabilities and replace these with appropriate located loxPsym sequences and high-efficiency CRISPR target sites. We plan to construct 100 to 200 kb regions at a time in yeast, and transfer these into appropriate human cell lines or iPSCs to gradually replace the native loci, learning from failures on the way. Once the loci are constructed and functional in human cells, the SCRaMbLE system will be used to generate rearrangements and a combination of RNAseq, nanopore sequencing and appropriate phenotypic screening will be used to determine the relationship between genotype and phenotype. A key goal will be to understand what DNA regions and genes within each locus are not essential for function and can be removed or replaced. This information will be used to enable ‘minimal, essential’ versions of these loci to be built.

CENTER *of* EXCELLENCE
for ENGINEERING BIOLOGY

Locus 1: Human MHC Region: Chromosome 6 – Type II section ~950 kb, Full Region 3.5 Mb.
The MHC locus is a hotspot for diversification by recombination and inheritance. In mammalian cells the locus is >3 Mb in length and the roles of much of the DNA and genes within this region are unknown. As synthetic version of the Type II section, which has the highest concentration of diversity, would be a starting point for a complete synthetic MHC region as this is just under 1 Mb in size. Alternatively, (or in parallel), synthesis could be used to make a complete, rationally-reduced MHC region initially containing only an essential set of genes. Further genes and DNA regions could then be added to this to determine their function. Intriguingly, the MHC region is only 92 kb in chickens, which are viewed as having a natural 'minimal' MHC. This could provide the template for a synthetic minimal version of the human MHC locus, and/or efforts to increase immunodiversity in the chicken genome by adding genes to its minimal set. How these regions will be designed and assessed will require expert input from the immunology field.

Locus 2: Human AZF Region: Y chromosome – Full region ~850 kb
Azoospermia Factor (AZF) region on the Y chromosome is known to undergo structural rearrangements, deletions and amplifications that lead to male infertility via failure to generate sperm. A synthetic version of the AZF region with natural hotspots for structural rearrangements replaced by inducible rearrangement (lox/CRISPR sites) could be made to produce large synthetic AZF libraries to test in a high-throughput manner for sperm-generation phenotypes. Input from fertility research will be essential. This would fit in well with recent advances in sperm generation from cultured cells.

Utility:
As well as providing pilot work towards the design and synthesis of complete mammalian chromosomes and genomes, the work has direct implications in healthcare research due to the choice of the two genomic loci. The MHC is one the key parts of immunity and also of central importance for organ transplant and xenotransplantation. Understanding the minimal MHC in chickens is also important for avian health, the food industry and combatting the spread of avian viral threats. The AZF locus is a key player in male fertility and will need to be understood in greater depth as efforts to generate sperm from cell cultures advance.

“Fit” For GP-write:
Pilot work described here will enable testing of megabase-scale construction and design in the human genome and investigate the utility of SCRaMbLE and genome refactoring in medically relevant target loci.

Submit to:

info@engineeringbiologycenter.org