Application of DNA Microarrays to Assess DNA Replication Timing and Chromosomal Aberrations.

Kathryn Woodfine
Darwin College

This dissertation is submitted for the Degree of Doctor of Philosophy

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration, except where specifically indicated in the text.

This Dissertation does not exceed the word limit set by the Biology Degree Committee.

Kathryn Woodfine.

Abstract

I have developed a directly quantitative method to assess the replication timing of sequences during the S phase of the cell cycle utilizing genomic clone DNA microarrays. This is achieved by the co-hybridisation of differentially labelled S and G1 phase DNA to the arrays. The genomic resolution of the replication timing measurements is limited only by the genomic clone size and density on the arrays.

I have demonstrated the power of this approach by constructing a genome wide map of replication timing in human lymphoblastoid cells using an array with clones spaced at 1 Mb intervals. I also constructed an array using chromosome 22 tile path clones and produced a high resolution replication timing map of 22q. Tile path resolution replication timing maps have also been produced for chromosomes 1 and 6.

I have shown a positive correlation, both genome wide and at a tiling path resolution, between replication timing and a range of genome parameters including GC content, gene density and transcriptional activity.

I have further developed the replication timing assay by using an array of PCR products spanning 4.5Mb at a resolution of 10kb, and an array spanning 20Kb using overlapping 500bp PCR products. This will allow the study of correlations with sequence features at a high resolution.

Using the Chromosome 22 tile path array I have also been able to show changes in replication timing in a cell line which contains a balanced translocation between chromosomes 17 and 22. I have also used the chromosome 22 tile path array to analyse deletions in DiGeorge patients and to detect VJ recombination at the immunoglobulin light chain λ lymphoblastoid cell lines.

Acknowledgments

I would like to thank all the people, who without which the work in this thesis would not be possible.

Firstly, all the external collaborators;
Silvana Debanardi and Bryan Young from the Molecular Oncology Unit, St Bartholomew's Hospital, Katrina Prescott from the Institute of Child Health, Richard Mott from the Wellcome Centre for Human Genetics and Charles Shaw-Smith from Addenbrookes Hospital.

The Human Genetics Department at The Wellcome Trust Sanger Institute. Especially; Carol Carder, Paul Hunt, Sean Humphray, Lisa French, Carol Scott, Rob Andrew, Pawendeep Dharmi, Cordelia Langford, Oliver Dovey and David Vetrie.

The Molecular Cytogenetics group (past and present) at the Sanger Institute, paticularly;
Heike Fiegler, Philippa Carr and Eleanor Douglas for the construction of the 1 Mb resolution genomic array, Shelia Clegg for teaching me FISH and Elena Prigmore for discussion and support.

The Molecular Genetics and Proteomics group at the Sanger Institute, especially; David Beare, for his immense help with sequence analysis of chromosome 22, John Collins for advice on analysis of low copy repeats and Owen McCann for help in the assembly of the chromosome 22 clone set

I would like to thank all the people who have given me a personal input into this thesis. In particular I would like to thank;
Gary Woodfine for his support, Andy Mungall, Lisa Rickman, and Deb Burford for proof reading and their suggestions and Susan Gribble for all her help, too great to mention.

Finally I would like to thank my two supervisors, Nigel Carter and Ian Dunham for all their help, support and patience.

Table of Contents

List of Figures:

List of Tables:

List of Abbreviations:

approx.	Approximately
ATP	Adenosine Triphosphate
BAC	Bacterial artificial chromosome
bp	base pair
BrdU	Bromodeoxyuridine
cdk	cyclin dependant kinase
cDNA	Complementary DNA
CpG	Cytosine and Guanosine dinucleotide
Cy	Cyanine dye
D. melanogaster	*Drosophila melanogaster*
dATP	2'-Deoxyadenosine 5'-triphosphate
dCTP	2'-Deoxycytosine 5'-triphosphate
dGTP	2'-Deoxyguanosine 5'-triphosphate
DMSO	Dimethyl sulphoxide
DNase	Deoxyribonuclease
dNTP	2'-Deoxynucleoside 5'-triphosphate
DOP	Degenerate Oligonucleotide Primer
dsDNA	Double stranded DNA
dTTP	2'-Deoxythymidine 5'-triphosphate
E. Coli	*Escherichia coli*
EDTA	Ethylenedinediaminetetraacetic acid
EST	Expressed sequence tag
Fig.	Figure
FISH	Florescence *in-situ* hybridisation
G	Giemsa
G1	Growth 1 phase of the cell cycle
G2	Growth 2 phase of the cell cycle
GC	Guanosine + Cytosine
H	Histone
HCl	Hydrochloric acid
HPLC	High
IgH	Immunoglobulin heavy chain
IgL	Immunoglobulin light chain
J	Joining region (IgL)
K	Lysine
Kb	kilobase
LB Agar	Luria-Bertani agar
log	logarithmic
M	molar
M	Mitosis phase of the cell cycle
Mb	Megabase
MCM	Mini chromosome maintainance

μg	microgram
mg	miligram
μl	microlitre
μM	micromolar
mM	milimolar
mRNA	Messenger RNA
NaAc	Sodium Acetate
NaCl	Sodium Chloride
nm	namometer
ORC	Origin Recognition Complex
Ori	Origin of Replication
PAC	P1-derived artificial chromosome
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PI	Propidium iodide
Pre-RC	Pre-Replication Complex
r.p.m	revolutions per minute
RNA	ribonucleic acid
RNase	Ribonuclease
S	Synthesis phase of the cell cycle
SDS	Sodium dodecyl sulphate
SSC	Sodium chloride/citrate solution
STS	Sequence tagged site
TDP	Timing Decision Point
TE	Tris (hydroxymethyl) aminomethane-Ethylenedinediaminetetraacetic acid
V	Variable region (IgL)