



***Rearrangement breakpoints and viral
integration sites in the human genome***

Deborah Burford

Darwin College

October 2007

This dissertation is submitted for the degree of Doctor of Philosophy

Rearrangement breakpoints and viral integration sites in the human genome.

Abstract

It is estimated that 1 in 2000 live births have an apparently balanced translocation and that the risk of congenital abnormalities in these carriers is approximately double that seen amongst individuals with normal karyotypes. To date, less than 50 de novo reciprocal balanced translocations have been sequenced across the breakpoint junctions and little is understood about the mechanisms that may cause these non-recurrent rearrangements. The aim of this thesis is to investigate methods of translocation breakpoint mapping. The results showed that array painting onto a Whole Genome Tile Path array followed by array painting onto a custom-made NimbleGen oligonucleotide array and subsequent amplification of the translocation junctions by long range PCR was the most efficient method tested for generating sequence across translocations. Analysis of the sequence generated showed that in the three patients studied, three of the six breakpoints directly disrupted a gene; CENTG2 in a patient with a t(2;7)(q37.3;p15.1) and PTPRZ1 and DACH1 in a patient with a t(7;13)(q31.3;q21.3). As an alternative approach, generation and screening of a custom-made fosmid library using flow sorted derivative chromosomes from an additional patient was used to generate sequence across the translocation breakpoints. This approach is useful for translocations in which additional complexity at the breakpoints confounds other methods for breakpoint mapping and sequencing. The fosmid library approach developed for the mapping of translocation breakpoints has also been applied to the mapping of viral integration into the human genome furthering our understanding of the integrated virus.

Table of contents

1	Introduction	1
1.1	The discovery of human chromosomes.....	1
1.2	Chromosome rearrangements.....	1
1.2.1	Chromosome translocations.....	1
1.2.1.1	Frequency of translocations.....	3
1.2.1.2	Associated phenotypic risk.....	6
1.2.2	Deletions, duplications, inversions.....	8
1.2.2.1	Phenotypic effect of deletions, duplications and inversions.....	10
1.2.2.2	Genomic disorders.....	10
1.2.3	Analysis of chromosome rearrangements.....	11
1.2.3.1	Metaphase chromosome banding analysis.....	12
1.2.3.2	FISH.....	12
1.2.3.3	Fluorescence activated chromosome sorting.....	13
1.2.3.4	Somatic cell hybrids.....	14
1.2.3.5	Southern Blotting.....	14
1.2.3.6	Comparative genomic hybridisation.....	14
1.2.3.7	Array painting.....	18
1.2.3.8	Custom-made libraries.....	20
1.2.3.9	Sequence analysis.....	21
1.2.4	Public databases of constitutional rearrangements.....	22
1.2.5	Normal variation within the human genome.....	22
1.3	Mechanisms behind chromosomal rearrangements.....	24
1.3.1	Double strand breaks.....	24
1.3.2	Homologous recombination.....	25
1.3.3	Non-homologous end joining.....	28
1.3.4	Genome architecture associated with rearrangements.....	29
1.3.4.1	Chi sequences.....	29
1.3.4.2	Translin associated motifs.....	29
1.3.4.3	Topoisomerase I and II sites.....	29
1.3.4.4	Immunoglobulin heptamers.....	30
1.3.4.5	Mini satellite sequences.....	30
1.3.4.6	Purine/pyrimidine tracts.....	30
1.3.4.7	SINE elements.....	31
1.3.4.8	LINE elements.....	31
1.3.4.9	Long terminal repeats.....	31
1.3.4.10	Class II transposons.....	32
1.3.4.11	Palindromes.....	32
1.3.4.12	Segmental duplications.....	32
1.4	Viral Integration into the human genome.....	34
1.4.1	The Human Herpes Virus.....	34
1.4.2	Evidence of integration and inheritance.....	35
1.4.3	The structure of the HHV-6B Virus.....	36
1.5	Aims of this thesis.....	37

1.5.1	Use and adaptation of current techniques for the mapping of rearrangement breakpoints in three patients with abnormal phenotypes.....	37
1.5.2	Development of techniques for the investigation of rearrangement breakpoints.....	37
1.5.3	Bioinformatic analysis of the genomic architecture surrounding genomic rearrangement breakpoints.....	38
1.5.4	Application of techniques developed for the mapping of rearrangement breakpoints to the investigation of viral integration in the human genome.....	38
2	Materials and methods.....	40
2.1	Cell culture.....	40
2.2	Isolation of Patient DNA.....	40
2.2.1	Genomic DNA extraction.....	40
2.2.2	Flow sorting of derivative chromosomes.....	41
2.3	Amplification of template DNA.....	42
2.3.1.1	REPLI-g.....	42
2.3.1.2	DOP PCR.....	43
2.4	Construction of microarrays.....	44
2.4.1	Genomic clone microarrays (fosmid and plasmid microarrays).....	44
2.4.1.1	Clone selection.....	44
2.4.1.2	Picking.....	44
2.4.1.3	DNA prepping.....	44
2.4.1.4	DOP PCR amplification of plasmid, fosmid and BAC clones.....	45
2.4.1.5	Amino-linking of DOP PCR products.....	46
2.4.1.6	Spotting onto slides.....	47
2.4.2	PCR product arrays.....	48
2.4.2.1	PCR primer design.....	48
2.4.2.2	PCR amplification of products for microarray production.....	48
2.5	Hybridisation and analysis of microarrays.....	48
2.5.1	Hybridisation and analysis of in-house arrays (PCR, fosmid, WGTP).....	48
2.5.1.1	Labelling.....	49
2.5.1.2	Precipitation of samples.....	49
2.5.1.3	Hybridisation and washing.....	50
2.5.1.4	Scanning and image quantification.....	50
2.5.2	Hybridisation and analysis of NimbleGen oligonucleotide microarrays (carried out by NimbleGen Systems).....	51
2.6	Fluorescence in situ hybridisation.....	51
2.6.1	Preparation of slides for FISH hybridisation.....	51
2.6.1.1	Metaphase spreads.....	51
2.6.1.2	Extended chromatin fibres.....	52
2.6.2	Preparation of clone insert DNA.....	53
2.6.2.1	Preparation of fosmid clone insert DNA.....	53

2.6.2.2	Preparation of plasmid clone insert DNA	54
2.6.3	Fluorescence labelling of probes	55
2.6.3.1	Whole genome amplification labelling of clone DNA.....	55
2.6.3.2	Nick translation of clone DNA	57
2.6.3.3	DOP PCR labelling of flow sorted chromosomes.....	57
2.6.4	Pre-treatment of metaphase spreads and DNA fibre slides.....	58
2.6.5	Hybridisation of FISH slides	58
2.6.6	Detection of FISH slides	58
2.6.7	Microscope analysis of FISH slides	59
2.7	PCR.....	60
2.7.1	STS PCR.....	60
2.7.2	Colony PCR.....	61
2.7.3	LR PCR amplification of junction fragments	61
2.8	Sequencing.....	62
2.8.1	Sequencing of LR PCR products.....	62
2.8.2	End sequencing of fosmid clones	62
2.8.3	Full sequencing of fosmid clones.....	62
2.9	Custom-made library production.....	63
2.9.1	In-house cosmid library	63
2.9.1.1	Extraction of high molecular weight DNA.....	63
2.9.1.2	Preparation of vector arms.....	63
2.9.1.3	Partial digestion of DNA.....	64
2.9.1.4	Recovery of digested DNA.....	65
2.9.1.5	Ligation of vector arms to digested DNA.....	66
2.9.1.6	Packaging of ligated DNA.....	66
2.9.1.7	Preparation of plating cells.....	66
2.9.1.8	Plating of library	66
2.9.2	EpiFOS fosmid library.....	67
2.9.2.1	Size modification of DNA	67
2.9.2.2	End-Repair of sheared DNA	67
2.9.2.3	Size selection of end-repaired DNA	67
2.9.2.4	Ligation of size-selected DNA.....	68
2.9.2.5	Preparation of E.Coli host strain	68
2.9.2.6	Packaging of ligated DNA.....	68
2.9.2.7	Plating of library	69
2.10	Generation of fosmid library filters	69
2.10.1	Printing of polygrid filters from 384 well plates	69
2.10.2	Replication of clones onto round filters from agar plates.....	69
2.10.3	Preparation of clone DNA on filters for hybridisation	69
2.11	Radioactive screening of fosmid library filters.....	70
2.11.1	Generation of probe for library screening	70
2.11.1.1	Generation of probe by Inter-ALU PCR.....	70
2.11.1.2	Generation of probe by STS PCR	71
2.11.2	Hybridisation of radioactive DNA probe to fosmid library filters	71
2.11.3	Washing and detection of fosmid library filters.....	71

3	Results: Investigation of chromosome rearrangements by array painting, comparative genomic hybridisation and PCR analysis	72
3.1	Introduction.....	72
3.2	Summary of methods.....	74
3.3	Results; Refinement of chromosome translocation breakpoints using custom-made fosmid microarrays.....	76
3.3.1	Testing the custom-made fosmid microarray.....	77
3.3.1.1	Self versus self hybridisation.....	77
3.3.1.2	769p hybridisation.....	80
3.3.2	Array painting analysis of patient derivative chromosomes on custom-made fosmid microarrays.....	84
3.3.2.1	Patient t(2;7)(q37.3;p15.1).....	85
3.3.2.2	Patient t(3;11)(q21;q12).....	90
3.3.2.3	Patient t(7;13)(q31.3;q21.3).....	92
3.4	Further refinement of chromosome translocation breakpoints by STS PCR, PCR microarrays and oligonucleotide microarrays.....	94
3.4.1	Refinement of breakpoints by STS PCR.....	94
3.4.1.1	Patient t(2;7)(q37.3;p15.1).....	94
3.4.1.2	Patient t(3;11)(q21;q12).....	97
3.4.1.3	Patient t(7;13)(q31.3;q21.3).....	99
3.4.1.4	Summary of STS PCR data.....	101
3.4.2	Refinement of breakpoints by custom-made PCR product microarrays.....	101
3.4.3	Refinement of breakpoints by custom-made oligonucleotide microarrays.....	103
3.5	Chromosome translocation breakpoint amplification by LR PCR.....	104
3.6	Sequence of junction fragments.....	106
3.7	Most rapid method of translocation breakpoint mapping.....	107
3.8	Investigation of further rearrangement around translocation breakpoints using fosmid microarrays.....	110
3.9	Investigation of additional imbalance in patient t(2;7)(q37.3;p15.1)....	111
3.9.1	Delineation of chromosome 3 duplication breakpoints in patient t(2;7)(q37.3;p15.1) by microarray analysis.....	112
3.9.2	Amplification and sequence of chromosome 3 duplication junction in patient t(2;7)(q37.3;p15.1).....	114
3.10	Conclusions.....	116
4	Results: Bioinformatic investigation of rearrangement breakpoints..	118
4.1	Introduction.....	118
4.2	Breakpoints and phenotypes.....	118
4.2.1	Patient phenotypes (as published in (Gribble et al. 2005)).....	119
4.2.1.1	Patient t(2;7)(q37.3;p15.1) phenotype.....	119
4.2.1.2	Patient t(3;11)(q21;q12) phenotype.....	119
4.2.1.3	Patient t(7;13)(q31.3;q21.3) phenotype.....	119
4.2.2	Direct disruption of a gene by the translocation breakpoint.....	119

4.2.2.1	Direct gene disruption in patient t(2;7)(q37.3;p15.1)	120
4.2.2.2	Direct gene disruption in patient t(3;11)(q21;q12)	121
4.2.2.3	Direct gene disruption in patient t(7;13)(q31.3;q21.3)	122
4.2.3	Translocation breakpoints and position effect	124
4.2.3.1	Position effect in patient t(2;7)(q37.3;p15.1)	124
4.2.3.2	Position effect in patient t(3;11)(q21;q12)	126
4.2.3.3	Position effect in patient t(7;13)(q31.3;q21.3)	128
4.2.4	Duplication in patient t(2;7)(q37.3;p15.1) and phenotype	129
4.3	Breakpoints and genome architecture	131
4.3.1	Analysis of sequence at breakpoints	131
4.3.1.1	Patient t(2;7)(q37.3;p15.1)	133
4.3.1.2	Patient t(3;11)(q21;q12)	133
4.3.1.3	Patient t(7;13)(q31.3;q21.3)	134
4.3.1.4	Summary of basepair sequence analysis at the breakpoints ...	134
4.3.2	Analysis of repeat structures and recombination motifs at breakpoints	134
4.3.2.1	Analysis of repeat structures around t(2;7)(q37.3;p15.1) breakpoints	136
4.3.2.2	Analysis of repeat structures around t(3;11)(q21;q12) breakpoints	137
4.3.2.3	Analysis of repeat structures around t(7;13)(q31.3;q21.3) breakpoints	137
4.3.2.4	Summary of sequence motifs found at translocation breakpoints	138
4.4	Mechanisms underlying genomic rearrangements	138
4.4.1	Translocations in patients t(2;7)(q37.3;p15.1), t(3;11)(q21;q12) and t(7;13)(q31.3;q21.3)	138
4.4.2	Duplication in patient t(2;7)(q37.3;p15.1)	139
4.5	Conclusions	140
5	Results: Investigation of chromosome translocations using custom- made libraries	141
5.1	Introduction	141
5.2	Comparison of custom-made library protocols	141
5.2.1	Cosmid library	142
5.2.2	EpiFOS fosmid library	143
5.2.2.1	Using whole genomic DNA	143
5.2.2.2	Using flow sorted derivative chromosomes	145
5.2.3	Conclusions	147
5.3	Using the EpiFOS kit to generate a library from a patient with a characterised translocation as a test case for developing screening methods	148
5.4	Comparison of screening methods using gridded library fosmid clones	148
5.4.1	Inter ALU PCR screening	149

5.4.2	PCR product pools	150
5.4.3	End sequencing.....	150
5.4.4	Specific breakpoint PCR product.....	153
5.4.5	Summary of fosmid clones identified by the different screening methods.....	154
5.4.6	Conclusions.....	157
5.4.7	Screening of round library plates.....	160
5.5	. Mapping translocation breakpoints using a custom-made fosmid library from a patient with a t(2;6)(q21.1;q25.1) translocation.....	162
5.5.1	Isolation of derivative chromosomes by flow sorting.....	163
5.5.2	Verification of flow sorted derivative chromosomes by reverse chromosome painting.	164
5.5.3	Mapping translocation breakpoints by array painting.....	165
5.5.4	Isolation of breakpoint spanning fosmid clones using a custom-made library.....	167
5.5.4.1	Production of the t(2;6)(q21.1;q25.1) fosmid library	167
5.5.4.2	Screening of the t(2;6)(q21.1;q25.1) library using PCR product pools.....	168
5.5.5	Identification of translocation breakpoints by sequence analysis	169
5.6	Conclusions	170
6	Results: Investigation of viral integration using microarrays and custom-made libraries.....	172
6.1	Introduction.....	172
6.1.1	The HHV-6 virus structure	172
6.1.2	Patient details	173
6.1.2.1	Flow karyograms.....	174
6.2	FISH analysis of viral integration sites in 3 patients.....	175
6.3	Investigation of integration using genomic microarrays	177
6.3.1	Verification of the microarray	177
6.3.2	Comparison between the 3 patients	178
6.4	Investigation of integration using a custom-made library	179
6.4.1	Generation of custom-made library	179
6.4.2	Screening of custom-made library	180
6.5	Investigation of integration using FISH onto extended DNA fibres	183
6.6	Conclusions	184
7	Discussion.....	185
7.1	Methods of rearrangement breakpoint mapping	185
7.1.1	Breakpoint mapping using microarray based techniques	185
7.1.2	Translocation breakpoint mapping using custom-made fosmid libraries.....	188
7.1.3	Application of techniques developed for the mapping of rearrangement breakpoints to the investigation of viral integration in the human genome	190

7.2	Bioinformatic analysis of genomic rearrangement breakpoints.....	191
7.2.1	Genotype to phenotype correlations.....	192
7.2.1.1	Identification of non-pathological variation	192
7.2.1.2	Gene disruption.....	193
7.2.2	Mechanisms underlying genomic rearrangements	195
7.2.2.1	Predisposition	196
7.2.2.2	Occurrence of rearrangements	196
7.3	Recent developments and the future of breakpoint mapping.....	197
7.4	Conclusions	199