4. Results 2
Preparation and Verification of the Genomic Microarrays

4.1: Introduction

To investigate replication timing and to correlate this with genome features at high resolution, a tile path genomic array using large insert clone DNA was constructed to cover the whole of 22q. The tile path resolution utilised overlapping sequencing clones giving an average resolution of 78Kb. After construction the array was verified extensively to assess reproducibility and response to copy number changes.

The replication timing assay entails the flow sorting of nuclei from the G1 and S phases of the cell cycle. To allow rapid sorting, the number of S phase cells within the population of nuclei to be sorted was optimised by adjusting the time of growth between sub-culture and harvest of the cells. This is described in section 4.2.

The construction and verification of the 22q tile path array is described in section 4.3. Array verification experiments were also performed on pre-constructed arrays assessing the entire genome at a 1Mb resolution. This is described in section 4.4.

At a later stage in the project, an array was also constructed, with a 10Kb resolution over a 4.5Mb region of chromosome 22q with 500bp PCR products. This array also contained a 200Kb region covered with overlapping 500bp PCR products. The verification of this array is described in section 4.5.

A further stage of array verification was to test whether each loci on the 22q tile path array responded to chromosome 22 copy number change. This was achieved by adding DNA from flow sorted chromosome 22 to one half of a G1:G1 hybridisation. This is described in section 4.6.

4.2: Optimisation of S phase fractions

The assessment of replication timing in cells developed in this study is dependant on the ability to flow sort S and G1 phase nuclei. In any unsynchronised cell population the majority of the cells are in the G1 phase. The time taken to sort the S phase fraction is thus a limiting factor. To minimise the number of cells and the time required for flow sorting, the optimum time to yield the maximum number of cells in S phase after subculture was assessed.

The male lymphoblastoid cell line HRC 575 (46, XY) was harvested at different intervals after subculture and passed through a flow sorter to obtain a cell cycle profile as described in 2.3.2.2. The percentage of cells in S phase was plotted against the time between sub culture and harvest (Figure 4.1).

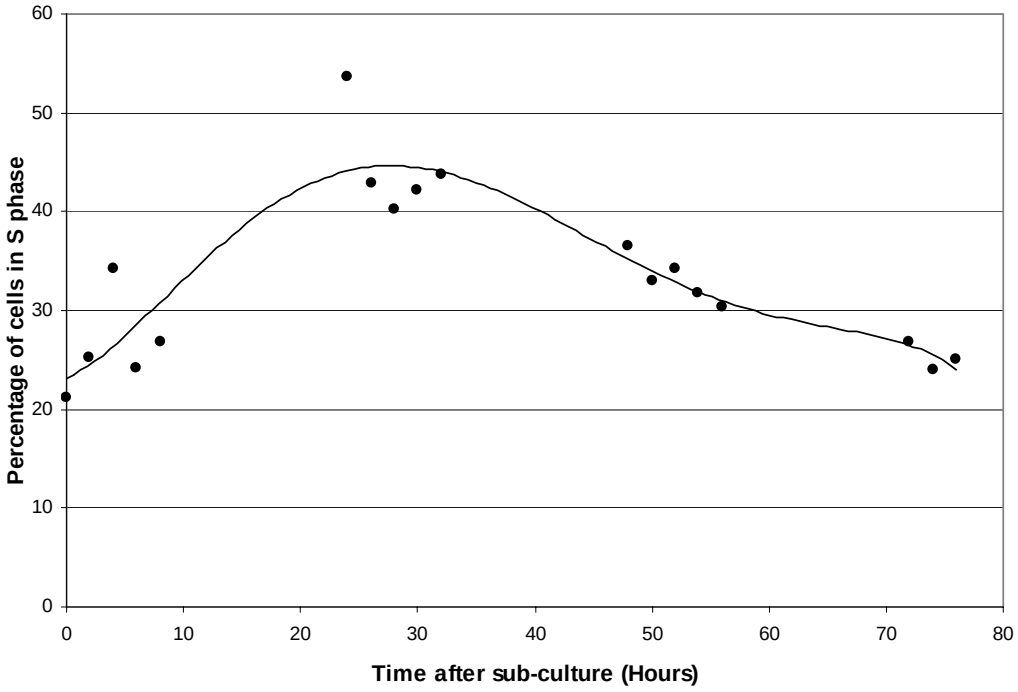


Figure 4.1: The change in the proportion of the cells in S phase at times after subculture for a lymphoblastoid cell line.

The optimal time between sub-culture and harvest of the lymphoblastoid cell line for a maximal S phase fraction was approximately 26 hours. The flow sort profile 26

hours after subculture is shown in Figure 4.2. This shows a high proportion of cells in S phase.

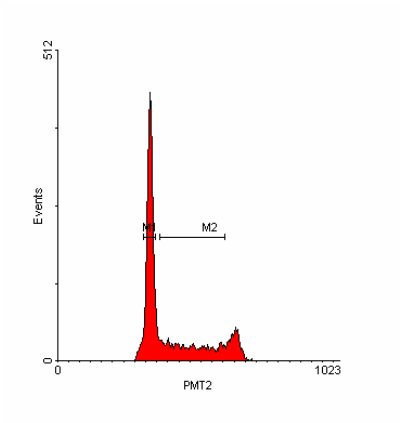


Figure 4.2: Lymphoblastoid nuclei flow sort profile after harvest 26 hours from subculture.

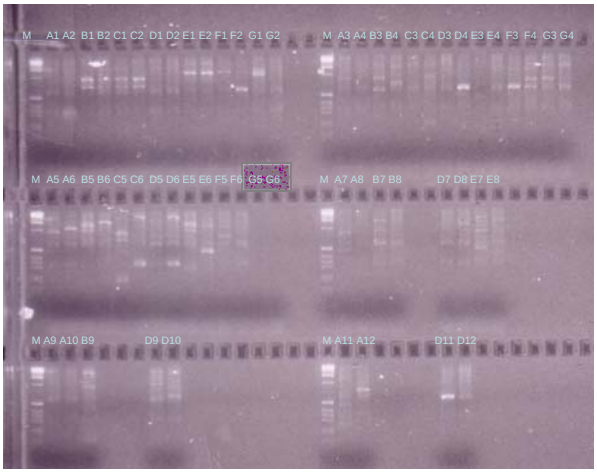
4.3: Preparation and initial verification of the 22q tile path array.

A 22q tile path array was constructed as described in 2.1. This comprised 470 clones, including cosmids, fosmids, PACS and BACS and covered the whole of the q arm of chromosome 22. 95 Chromosome X clones were also spotted onto the array. These were used as an intrinsic control to measure single copy number changes in male versus female DNA hybridisations.

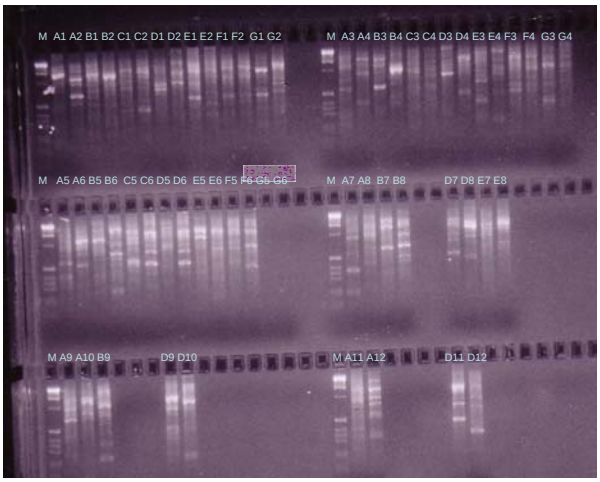
4.3.1: Amplification of chromosome 22 tile path clones.

The clones were first amplified by degenerate oligonucleotide primed PCR (DOP-PCR) using three different primers as described in 2.1.2.1. To ensure the PCRs had been successful and that no contamination had taken place, 5µl of the PCR product was assessed by gel electrophoresis as shown in Figure 4.3.

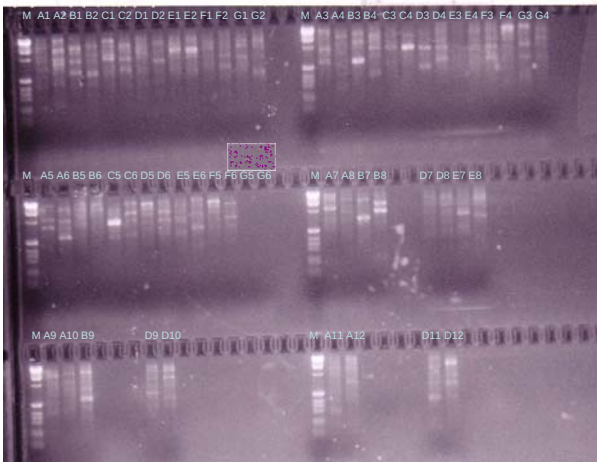
Clones were then amplified by a second round amino-linking PCR as described in 2.1.2.2 (see Figure 4.4).



DOP 1



DOP 2

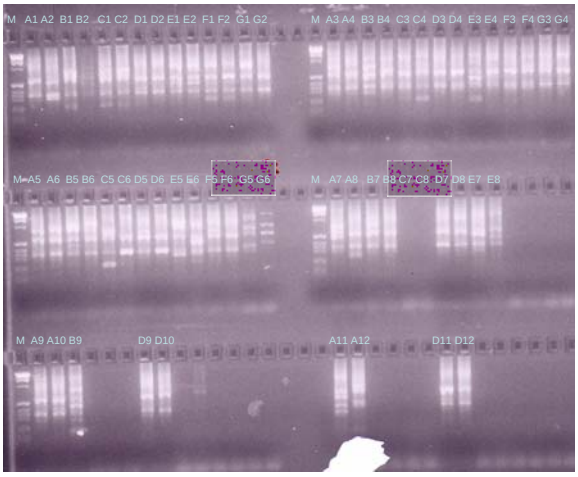


DOP 3

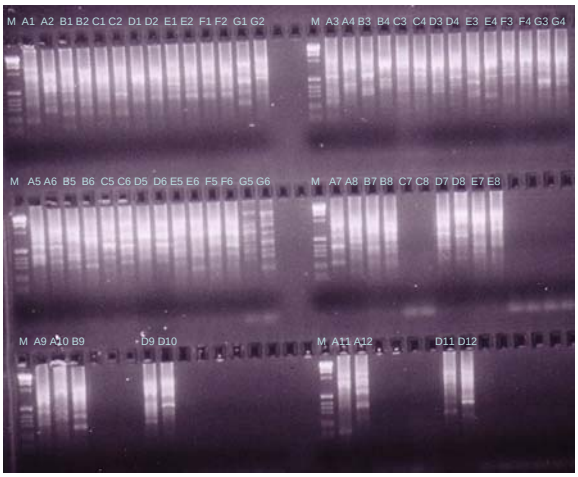
Key:

A1	p393	C1	100h	E4	bK217C2
A2	dJ1172A22	C2	91c	E5	bA140I15
A3	dJ127B20	C3	8C	E6	bK358H9
A4	dJ366L4	C4	cN81C12	E7	bK1109B5
A5	dJ403E2	C5	cN74G7	E8	kB63E7
A6	dJ20O8	C6	cB23F1	F1	cN61N10
A7	dJ101G11	D1	b445	F2	444P20
A8	dJ697G8	D2	b9j16	F3	174C04
A9	dJ437O22	D3	kB1027C1	F4	188L13
A10	dJ1014D13	D4	kB1269D1	F5	61P17
A11	dJ319F24	D5	kB1572G7	F6	73P10
A12	dJ340K22	D6	kB113H7	G1	bA9F11
B1	dJ796I17	D7	kB1839H6	G2	bA354I12
B2	dJ549K18	D8	kB1125A3	G3	bA255N20
B3	dJ477H23	D9	bK322B1	G4	fF96H12
B4	dJ162H14	D10	kB1561E1	G5	negative ctl
B5	dJ37M3	D11	kB1674E1	G6	negative ctl
B6	dJ477J10	D12	kB1896H1		
B7	dJ100G10	E1	kB1195A5		
B8	dJ34P24	E2	bK437G10		
B9	dJ566L20	E3	bK414D7		

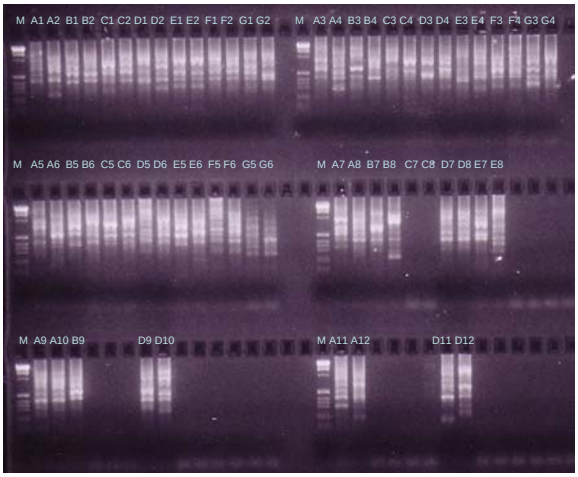
Figure 4.3: DOP-PCR amplification of a selection of chromosome 22 tile path clones, as indicated in the key.



Amino 1



Amino 2



Amino 3

Key:

A1	p393	C1	100h	E2	bK437G10
A2	dJ1172A22	C2	91c	E3	bK414D7
A3	dJ127B20	C3	8C	E4	bK217C2
A4	dJ366L4	C4	cN81C12	E5	bA140I15
A5	dJ403E2	C5	cN74G7	E6	bK358H9
A6	dJ20O8	C6	cB23F1	E7	bK1109B5
A7	dJ101G11	C7	negative ctl	E8	kB63E7
A8	dJ697G8	C8	negative ctl	F1	cN61N10
A9	dJ437O22	D1	b445	F2	444P20
A10	dJ1014D13	D2	b9j16	F3	174C04
A11	dJ319F24	D3	kB1027C1	F4	188L13
A12	dJ340K22	D4	kB1269D1	F5	61P17
B1	dJ796I17	D5	kB1572G7	F6	73P10
B2	dJ549K18	D6	kB113H7	G1	bA9F11
B3	dJ477H23	D7	kB1839H6	G2	bA354I12
B4	dJ162H14	D8	kB1125A3	G3	bA255N20
B5	dJ37M3	D9	bK322B1	G4	fF96H12
B6	dJ477J10	D10	kB1561E1	G5	negative ctl
B7	dJ100G10	D11	kB1674E1	from DOP	
B8	dJ34P24	D12	kB1896H1	G6	negative ctl
B9	dJ566L20	E1	kB1195A5	from DOP	

Figure 4.4: Amino-linking PCR amplification of a selection of Chromosome 22 tile path clones as indicated in the key.

4.3.2: Male:male hybridisation onto the chromosome 22 tile path array.

A male self:self hybridisation was carried out using DNA extracted from HRC 575 lymphoblastoid cell line to assess the background variation in measurements.

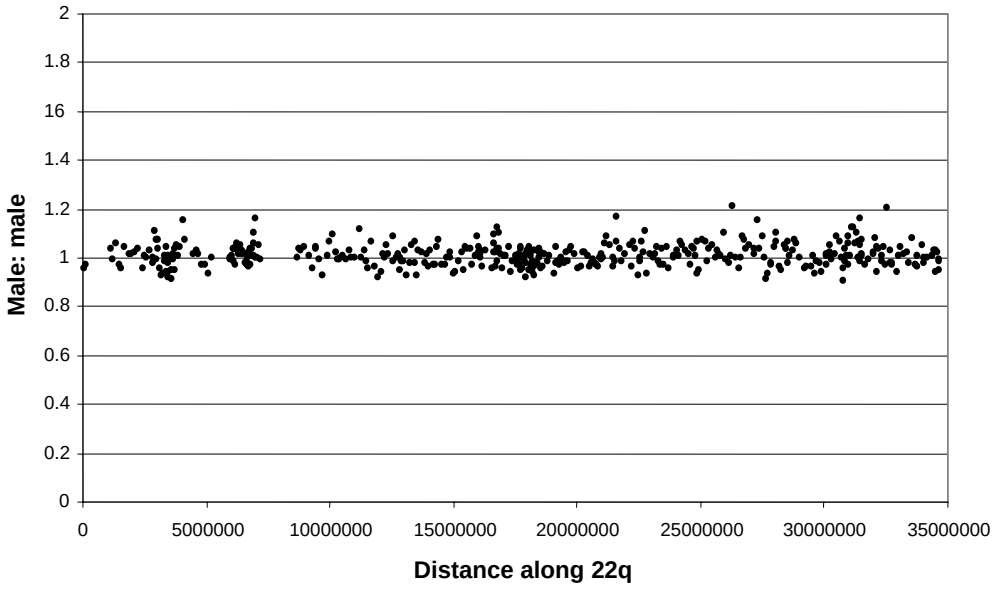


Figure 4.5: Male self:self hybridisation on the chromosome 22q array

The self:self hybrisiation was performed in duplicate. Four of the 470 clones represented on this array were excluded by the analysis program, because the clone intensities were not sufficiently above the *Drosophila* BAC clone background level, or because the triplicate spots were not all within 5% of the mean for that triplicate (for details of the analysis see section 2.5.2.). The average ratio of all the chromosome 22 clones was 1.00 with a standard deviation of 0.04.

4.3.3: Male:female hybridisation onto the array.

A male:female hybridisation was carried out using differentially labelled DNA extracted from a male lymphoblastoid cell line (HRC 575) and a female lymphoblastoid cell line (HRC 160). The aim of this experiment was to verify that X clones on the array accurately reported a single copy number difference between the male and female DNA (i.e. a ratio of 0.5).

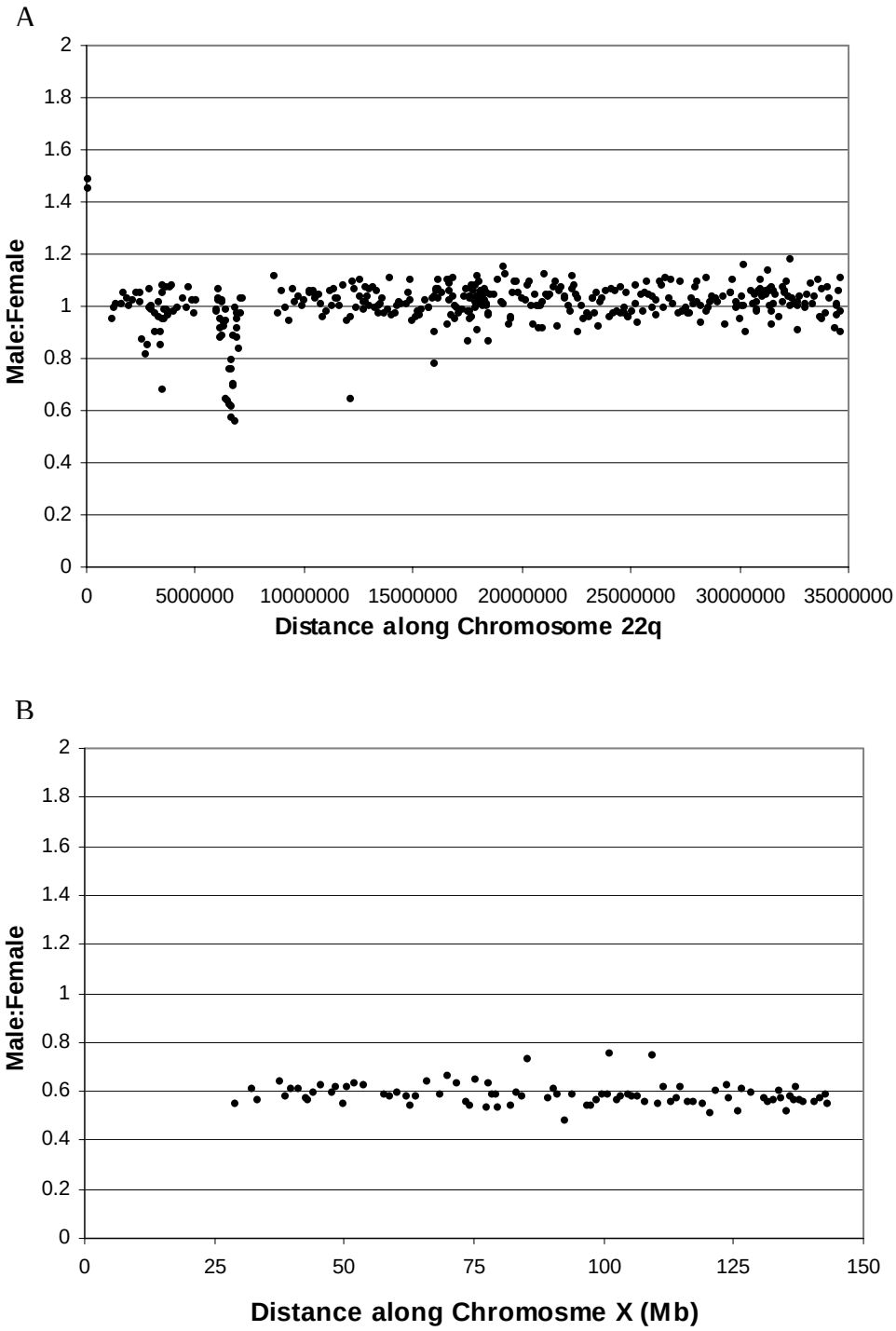


Figure 4.6: Male:female hybridisation on the constructed 22q array. A: Male:Female ratio on the 22q clones. B: Male:Female ratio on the X clones on the 22 array.

Fifty of the 470 clones were rejected during the analysis stage. The average male:female ratio on the chromosome 22 clones was 1.00 with a standard deviation of 0.09 (Figure 4.6). The region 6433945 – 6823353bp along the q arm of chromosome 22 shows clone ratios that could be interpreted as either a single copy

deletion within the male cell line, or a gain in the female cell line. This was further investigated (detailed in section 7.3) and revealed a deletion in the male cell line. Omitting this region from the statistical analysis, the standard deviation of the data points reduces to 0.07. Other clones on the chromosome 22q tile path array also reported unexpected ratios and are summarised in Table 4.1.

Table 4.1: Clones showing unexpected ratios in a male:female hybridisation

Clone	**Clone position (bp)**	**Possible reason for aberrant ratio**
cN14H11	99514	Centromeric clone
cN64E9	114958	Centromeric clone
59f	3467897	Rich in low copy repeats
995o6	2710127	Rich in low copy repeats
699j1	2822641	Rich in low copy repeats
519d21	2577096	Rich in low copy repeats
83e8	3443824	Rich in low copy repeats
dJ477H23	12110340	Clone not verified
cN113A11	16010331	Clone not verified

Seven of the 93 chromosome X clones were rejected by the analysis criteria described in section 2.5.2. The average male:female ratio on the chromosome X clones was 0.58 with a standard deviation of 0.04.

4.3.4: G1 self:self phase DNA Hybridisation onto the 22q tile path array.

A G1 self:self hybridisation was carried out to assess whether extraction from cell sorted nuclei affected ratio measurement variance. DNA was obtained from the G1 phase of the cell cycle as described in 2.3.2, differentially labelled and hybridised to the 22q array.

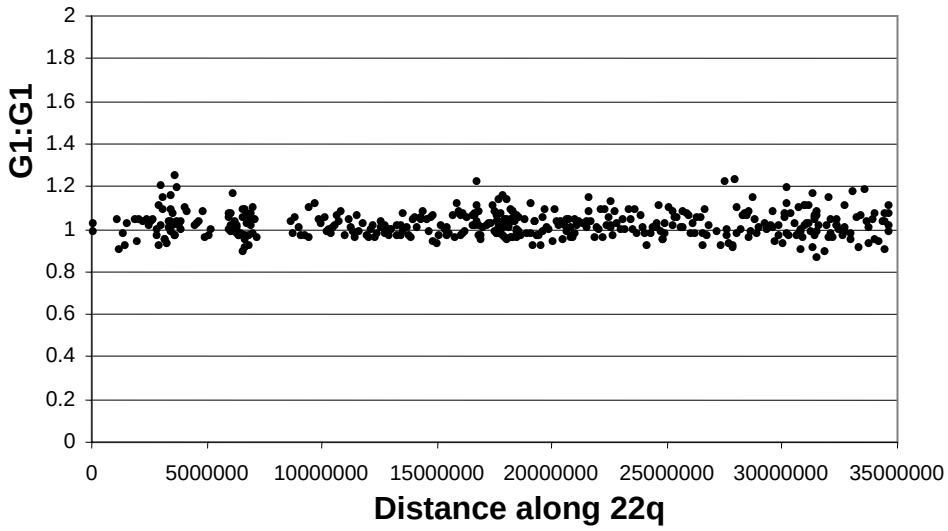


Figure 4.7: G1:G1 hybridisation on the 22q array.

Eighteen of the 444 clones were excluded from the analysis according to the criteria described in 2.5.2. The mean ratio reported was 1.00. The standard deviation of the ratios was 0.06.

4.4: Control Hybridisations on the 1Mb array

Similar verification experiments as detailed above were performed on an array sampling the whole genome at a 1Mb resolution.

4.4.1: Male:male hybridisation on the 1Mb array

A single male:male hybridisation was carried out. Of the 3126 clones on the array 82 were excluded at the analysis stage. The mean ratio reported by the remaining clones was 1.00 with a standard deviation of 0.039. The ratio profiles for all chromosomes can be seen in Appendix 4.

4.4.2: Male:female hybridisation on the 1Mb array

A male:female hybridisation was carried out on the 1Mb array. Of the 2955 autosomal clones on the 1Mb array, 256 were excluded at the analysis stage. The average ratio reported was 1.00 with a standard deviation was 0.10. Of the 150 Chromosome X clones on the array, 17 were excluded at the analysis stage. The average ratio reported was 0.75 and the standard deviation was 0.051. The chromosome profiles for all chromosomes can also be seen in Appendix 5.

4.5: Production of a high resolution array from PCR products.

A high resolution array was constructed sampling a 4.5Mb region of chromosome 22, 15398721 – 19982021bp along the q arm at a resolution of one approximately 500bp PCR product every 10Kb. In addition, overlapping 500bp products were designed to cover the region 16495000-16695000bp along the q arm of chromosome 22. The design of primers is described in section 2.2.1.

The first round of amplification was performed using clone DNA as template. Products from each PCR plate were analysed by gel electrophoresis using a 2.5% agarose gel. A successful PCR was indicated by a single band with a product size of approximately 500bp (Figure 4.8)

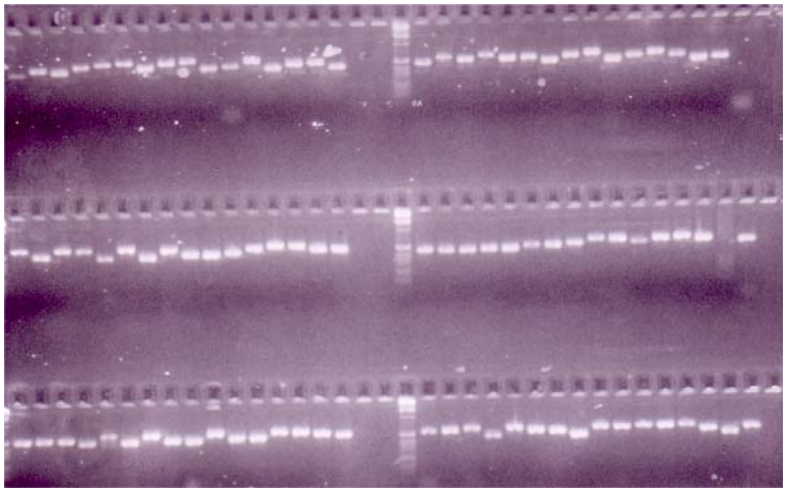


Figure 4.8: PCR products obtained from the amplification of primers STSG 495474 – STSG 495569 in a 96 well format as detailed in Appendix 2b.

A strong, clean amplification product was observed for 599 of the 714 primer pairs tested. A further 16 primer pairs produced a weak product whilst 99 produced no product. Primers producing a weak product or no product were re-amplified using genomic DNA as a template. Of the 115 primer pairs re-amplified, 68 gave a strong product, 16 gave a weak product, 30 produced no product and 1 generated a double band, suggesting amplification of more than one region of the genome, although this was not confirmed.

Each PCR product was spotted on the array in quadruplicate. As with the tile path arrays self:self and male:female hybridisations were used for array verification as shown in Figure 4.9 and Figure 4.10.

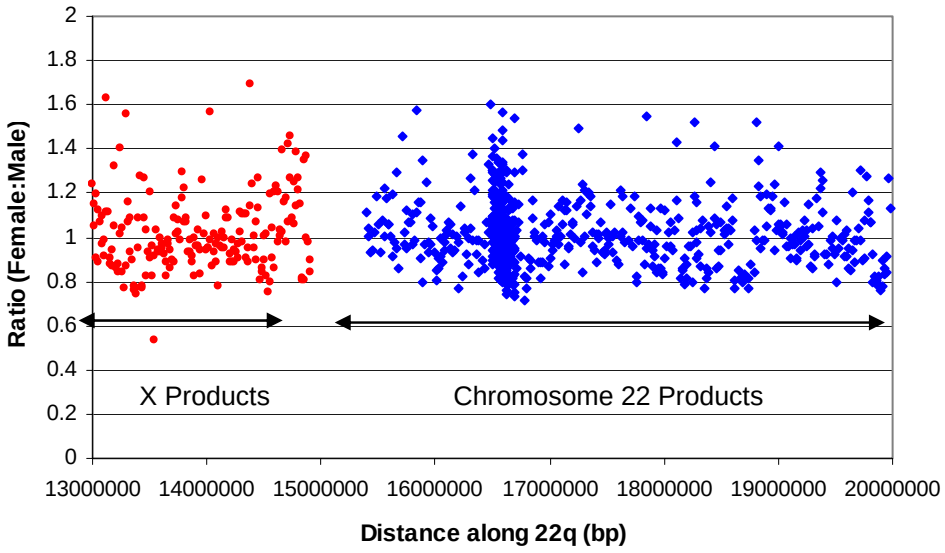


Figure 4.9: A G1:G1 hybridisation on the high resolution PCR product array.

For the G1:G1 hybridisation, 50 of the 714 Chromosome 22 PCR products failed the analysis criteria due to the criteria given in 2.5.2. Of the remaining 664 clones, the mean ratio reported was 1.00 and the standard deviation was 0.15.

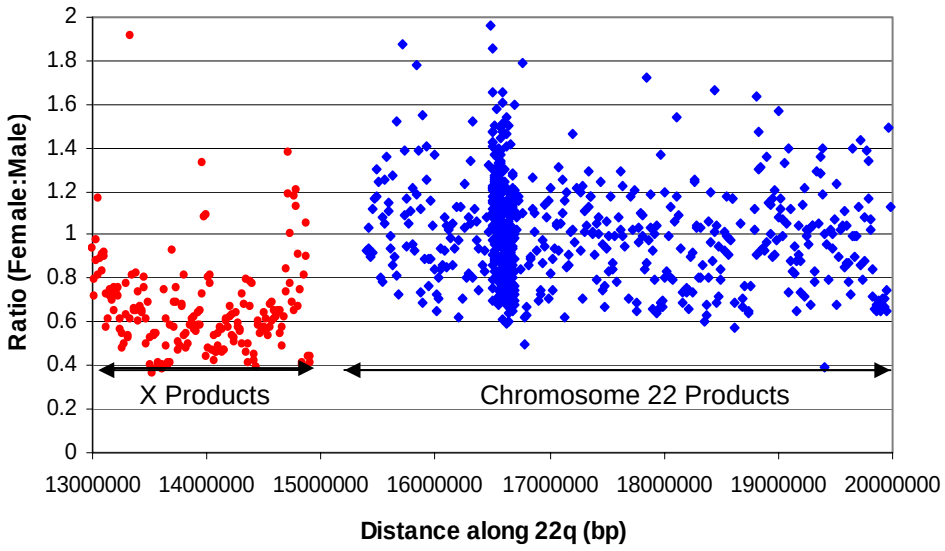


Figure 4.10: A male:female hybridisation on the high resolution PCR product array.

Analysis of the male:female hybridisation revealed forty of the 714 chromosome 22 PCR products on the array failed the analysis criteria described in 2.5.2. The standard deviation of the remaining loci was 0.23. The chromosome X PCR products were analysed, and gave an average male: female ratio of 0.67. The average standard deviation of the chromosome X loci on the array was 0.26.

4.6: Detection of chromosome 22 copy number changes on clone arrays

4.6.1: Detection of chromosome 22 copy number change on the 1 Mb tile path array.

The reporting of a copy number change by a clone, in response to a chromosome 22 sequence in the hybridisation mix, was assessed by the addition of flow sorted chromosome 22 DNA to a self:self hybridisation utilising genomic DNA. This is described in section 2.5.2.3. Results for the 1Mb resolution genomic array are shown in Figure 4.11.

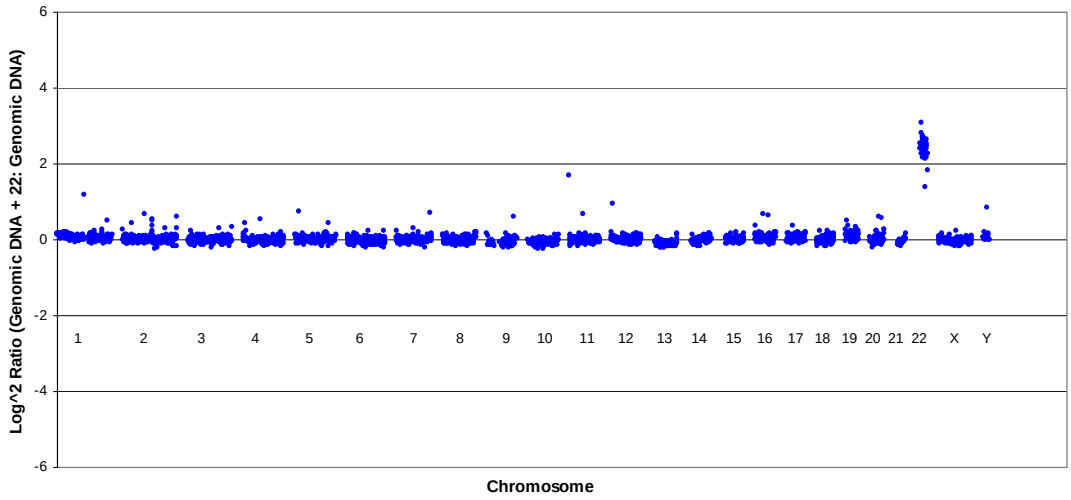


Figure 4.11: A genomic DNA + Chr 22:genomic DNA hybridisation on the 1 Mb array.

All chromosome 22 clones on the 1Mb array responded to the addition of five copies of chromosome 22 into the hybridisation mix by showing a copy number gain. However some clones on other chromosomes also report a copy number gain.

Examining the chromosome 22 clones in detail, the average ratio reported was 5.57 with a standard deviation of 0.94 (see Figure 4.12).

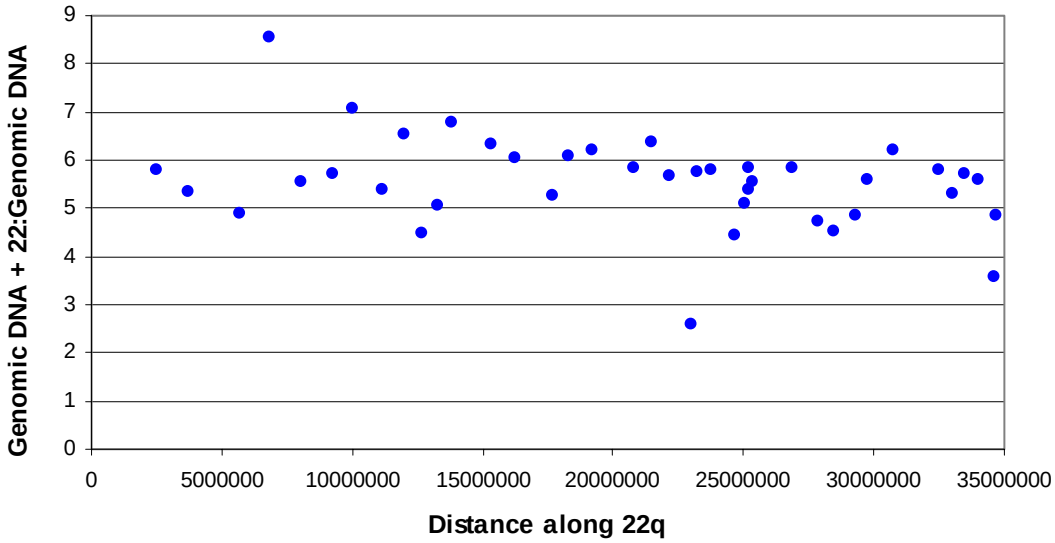


Figure 4.12: Response of the chromosome 22 clones to a chromosome 22 add-in experiment.

One clone, RP11-50L23 located 6.8 Mb along the q arm of chromosome 22, reported a particularly high ratio of 8.41.

The chromosome 22 clone reporting the lowest ratio (2.66) was CTA-150C2. However this ratio is still significantly above all the ratios reported on clones from other chromosomes, except the chromosome 11 clone CTC-908H22 (discussed below) and so this clone still reports a change in chromosome 22 copy number.

Several clones in the rest of the genome reported high ratios indicating that they too report a response to the increased amount of chromosome 22 in the hybridisation mix. Clones that reported a ratio above the 99% confidence interval for the mean ratio of modal clones are detailed in Table 4.2.

The clone showing the largest response to the chromosome 22 DNA is a clone located at the 11p telomere (CTC-908H22). This is illustrated in Figure 4.13.

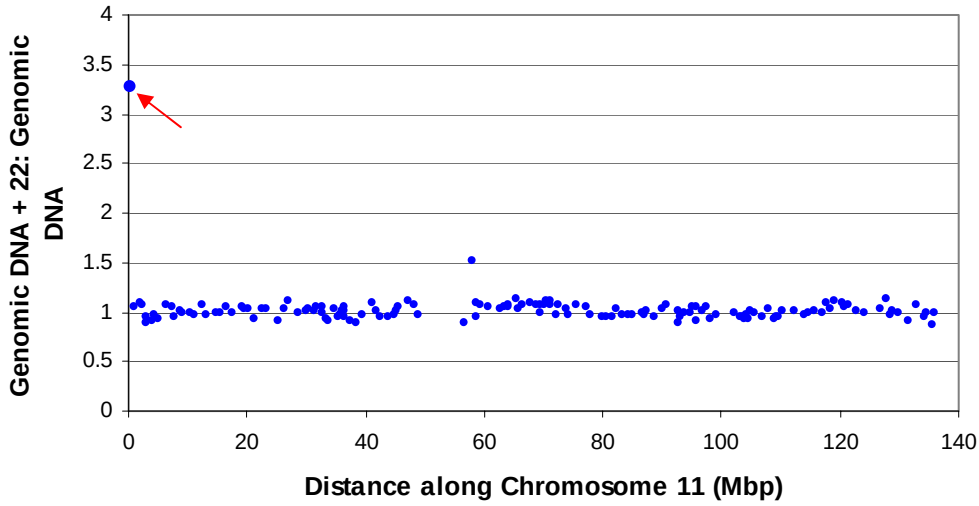


Figure 4.13: Hybridisation ratios reported by chromosome 11 clones after a genomic DNA + 22: genomic DNA hybridisation

4.6.2: Detection of chromosome copy number changes on the 22 tile path array.

A similar experiment was performed on the 22 tile path arrays to test the responsiveness of array loci to chromosome 22 copy number change. Arrays were performed with an estimated 1 additional, 2 additional and four additional copies of

chromosome 22 in the hybridisation mix. A G1 self:self and a G2:G1 hybridisation was also performed within the same batch of arrays. Arrays were normalised against the chromosome X clones, which should report no copy number change. The mean copy number change for the 22 clone was calculated.

The response of clones to copy number changes are shown in Figure 4.14 where the ratio is plotted against the approximate number of extra copies of chromosome 22 added to the hybridisation mix. Hyper-responsive clones plotted on Figure 4.14 are p87O8, pac699j1, dJ293L6, and cN69E4. Clones under reporting copy number change are, b444p24, cN61D6, cN20A6 and cN21F1. Four clones reporting a correct response were also included for comparison. These clones, chosen at random were not located within the first 9Mb of the q arm, known to contain a considerable segmental duplication. These clones are dJ127L4, bK282F2, fF4G12 and bK126B4.

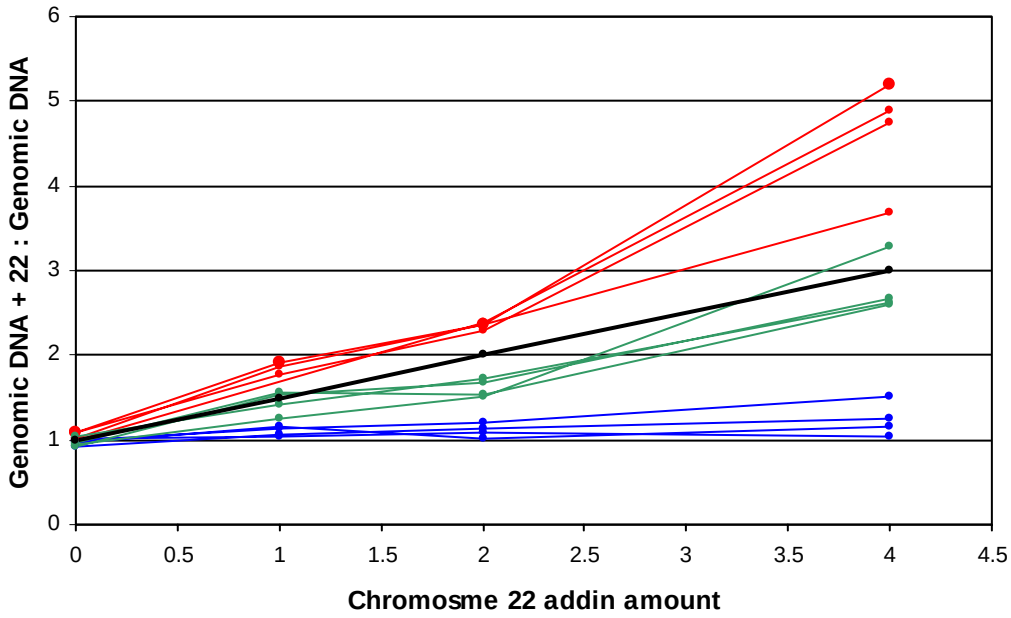


Figure 4.14: Ratios reported when different amounts of chromosome 22 are added into the hybridisation mix. Red: Clones that are hyper-responsive to addition of chromosome 22. Blue: Clones that are not responsive to the addition of chromosome 22. Green: Clones that report a normal response to chromosome 22. Black: Ideal copy number change reported.

Of the 470 clones on the chromosome 22 array, only twenty clones over or under-reported the response to chromosome 22 DNA within the hybridisation mix. This

indicates that 96% of the chromosome 22 clones report copy number changes accurately.

4.7: Discussion

4.7.1: Control hybridisations performed on the clone arrays.

A control self:self hybridisation was performed on the clone DNA arrays and average expected 1:1 ratios were reported by both the 1Mb and 22 tile path arrays. The standard deviations reported by the 22q tile path array and the 1Mb array were comparable, showing the reproducibility of the method when constructing an array from large insert clone DNA using DOP PCR.

Clones from chromosome X were included on each type of array constructed. These clones provide an intrinsic control to allow simple verification of copy number changes using male:female hybridisations. A male:female hybridisation should report a ratio of 1:1 on loci derived from autosome sequence but a ratio of 0.5:1 on loci representing chromosome X due to the X chromosome copy number difference between males and females. A ratio 0.5:1 was not reported on by any of the X loci represented on the arrays. The lowest ratio reported was 0.58:1 on the 22q tile path array and the highest ratio reported was 0.75:1 on the 1Mb array. This could be due the representation of different X clones on the two arrays. There are 46 more X clones on the 1Mb resolution array than there are on the chromosome 22q array. This under reporting of the copy number difference on chromosome X has been reported previously (Pinkel, Segraves et al. 1998; Fiegler, Gribble et al. 2003). Possible reasons for this underestimate could be an under-representation of chromosome X sequences in Cot 1 DNA leading to incomplete suppression of repeats on the chromosome X loci and cross-hybridisation of other regions of the genome with high sequence homology. Chromosome X has been identified as being paticulary rich in LINE repeats (IHGSC 2001). However, assessment of segmental duplications throughout the whole genome (Bailey, Gu et al. 2002) shows that chromosome X is relatively sparse in interchromsomal repeats.

Another possibility for the underestimate of the copy number change reported by the chromosome X clones is that, unlike autosomes, the two copies of chromosome X in the female DNA are not the same but differ epigenetically. In females one X chromosome is epigenetically silenced, rendering it transcriptionally inactive. This is to ensure that there is the same dosage of genes encoded on chromosome X in males and females (for review see (Avner and Heard 2001)). This epigenetic silencing involves the tight condensation of the chromatin into an inactive barr body. Inactivation makes the DNA within the inactive X chromosome very inaccessible which may affect DNA labelling such that Cy dye is not incorporated into the inactive chromosome X with the same efficiency as it is into active chromosomes. This means that after a male:female hybridisation a full 1:2 ratio would not be reported on the X clones.

4.7.2: Verification of the 1 Mb resolution and chromosome 22 Tile path arrays.

Further verification on the chromosome 22 and 1Mb arrays were performed with a series of experiments utilising different amounts of additional chromosome 22 DNA in the hybridisation mix. On the 1Mb array, one clone from chromosome 22 (RP11-50L23) can be seen to be hyper-responsive to the chromosome 22 DNA. This clone is located 6.8 Mb along the q arm of chromosome 22 within the locus encoding the immunoglobin light chain λ region. During lymphoblastoid development this region undergoes rearrangement and deletion. The control cell line, from which the genomic DNA was extracted (HRC 575), has been shown to have a deletion in this region (see 4.3.3 and 7.3). It is therefore likely that the hyper-sensitivity of this clone is due to the presence of only one copy of chromosome 22 in the cell line the genomic DNA was extracted from. Calculations reported in section 4.6.2 assumed two copies of chromosome 22 in the genomic DNA.

Chromosome add-in experiments on the 22q tile path array showed over 96% of loci reported the correct response to increased dosage of chromosome 22. The linear response reported by representative clones of this majority group (see Figure 4.14) confirm that the clones responded appropriately to the extra copies of chromosome 22 added.

Several clones did not report the correct response to additional copies of chromosome 22. On the 1Mb array the chromosome 22 clones adjacent to the telomere reported a depressed response to the addition of chromosome 22 to the hybridisation mix. This is unsurprising as the telomeric region contains a large amount of genome repeats (see section 1.3.1). Therefore it is likely that these clones will cross hybridise with other regions of the genome. Table 4.2 details clones not mapping to chromosome 22 represented on the 1Mb array that responded to the additional copies of chromosome 22 in the hybridisation mix.

Table 4.2: Clone not mapped to chromosome 22 that responded to extra chromosome 22 in the hybridisation mix

Clone	Chr.	position	Ratio	End sequence	Segmental Duplications
RP11-114F20	3	197298109	1.25	Match	None
RP5-1107C24	20	60245780.5	1.26	Match	None
CTD-3113P16	19	244656.5	1.28	multiple, none on 22	19,21,4,5,8,6,22
RP11-260J21	2	236147059	1.30	Match	None
CTD-2547N9	19	9002070.5	1.31	Match	None
RP11-278G12	2	38037637	1.33	Match	None
RP11-260A9	17	27226356.5	1.36	Match	None
RP11-565I3	4	7435638.5	1.38	Maps to Chr 14	None for 4
RP11-1E1	4	78240481	1.38	Match	18,9
RP11-24O13	2	130447729	1.38	multiple, On 22	None
RP11-276J4	1	223457232.5	1.44	multiple, none on 22	9,13,10,1,5
RP11-30F17	19	6552559	1.44	No end sequence	19
RP11-205K6	9	120945187	1.47	Match	9
RP11-100N3	11	58074075	1.52	Match	None
RP4-724E16	20	51861416.5	1.53	Match	None
GS1-172I13	2	241706787	1.53	No end sequence	1,2,21
RP11-209H16	2	129390774.5	1.54	multiple, On 22	2
RP11-71B7	2	93952867	1.57	multiple, none on 22	None
RP11-122O1	20	37014297	1.59	No end sequence	4,1,7,11,14,3,12,9,X
RP1-29O12	5	14786570.5	1.59	No end sequence	5
RP11-408D2	16	35065398.5	1.61	multiple, On 22	16,6
RP11-165M2	16	55986775	1.66	Match	16
RP11-208G20	7	150257256.5	1.70	multiple, none on 22	7
RP11-434F12	24	18960023	1.78	multiple, none on 22	Y,12,3,UL
RP3-467F14	12	6148320.5	1.97	multiple, On 22	15,4
CTC-908H22	11	175000	3.26	multiple, On 22	1,4,11

End sequence match= sequence from end sequencing of the chromosome matched their location in Ensembl. UL= Unlocated, contig not mapped to any chromosome

The 26 1Mb array clones that cross hybridised with chromosome 22 were analysed in two different ways to see if the cross hybridisation could be explained. All the clones

in the 1Mb clone set had been end sequenced and compared to the genome sequence to verify position, and locate other regions of similarity. The study of this database (http://intweb.sanger.ac.uk/cgi-bin/humace/1mbsetends.cgi) showed that five clones had end sequences that contained a significant amount of homology to chromosome 22 sequence. This could either indicate a mixed well when the clone was picked or a segmental duplication within the DNA that was end sequenced. A mixed well would lead to representation of more than one region of the genome on the array such that the reporting of copy number changes at this locus would be inaccurate. The presence of segmental duplications within the clone results in cross hybridisation of other regions of the genome. The end sequences of one clone, RP11-565I3, mapped to chromosome 14, not chromosome 4 as previously thought.

Clones were also analysed using the segmental duplication track on the UCSC genome sequencing database (http://humanparalogy.gene.cwru.edu). The segmental duplications were identified as described by Bailey et al (Bailey, Gu et al. 2002). This analysis revealed a further clone with homology to chromosome 22. However it should be noted that not all clones showing a homology to chromosome 22 by their end sequence are detected on this database. This confirms the incomplete status of this database and the human genome sequence at the time of analysis (Bailey, Gu et al. 2002), (IHGSC 2001).

A further 11 clones had end sequences that mapped to more than one chromosome, or segmental duplications involving chromosomes other than 22. Although this does not explain the cross hybridisation with sequences from chromosome 22, it does indicate that these clones contain repetitive DNA. Inefficient blocking by Cot 1 DNA, or the presence of chromosome 22 segmental duplications that were not identified by Bailey *et al* (Bailey, Yavor et al. 2002) may explain the cross hybridisation with chromosome 22.

The remaining eight clones had end sequences that match their positions assigned on the 1Mb profiles (Appendix 4) and no duplications within chromosome 22. However, most of these clones have ratios toward the lower end of those identified in Table 3.2. The statistical analysis used to identify clones with a significant response to the additional chromosome 22 DNA uses the 99% confidence level of modal

values. On a purely statistical basis, on an array containing 3,500 clones, 35 clones would be expected to report a ratio over the 1.24 cut-off identified.

To be classified as an atypical reporting clone the clone had to report a copy number change with a standard deviation outside the 99% confidence intervals in two of the three arrays. These clones are summarised in Table 4.3.

Table 4.3: Clones not responding with the correct copy number change when chromosome add-in experiments were performed on the tiling path arrays.

Clone	Accession No.	Position	No. of arrays	Comments	Segmental duplication*
Clones under reporting copy no. changes					
cN64E9	AP000526	114958	2	centromeric	1,2,9,10,14,16,22,UL
p87o8	AC007064	1248002	3	Seg dup	None
pac699j1	AC008103	2822641	3	Seg dup	1,4,5,6,13,20,22,UL
56c	AC000080	3878158	2	Seg dup	None
2H8	D87003	6336208	3	Seg dup	1,2,4,15,16,22,UL
bA541J16	AL080241	12558437	2		None
bA329J7	AL118497	12578094	2		None
cE78G1	Z70288	17685472	3		None
dJ293L6	AL049749	20707056	3		None
dJ591N18	AL031594	24555326	2		None
dJ408N23	Z98048	24844579	2		None
cN69F4	Z72006	32361681	3	telomeric	22
n1g3	AC002055	34687355	3	telomeric	2,22
Clones over reporting copy no. changes					
b444p24	AC007663	4165628	2	seg dup	22, UL
cN61D6	D87012	5997695	3	VJ region	None
cN75C12	D87017	6963826	2		22
cN20A6	Z69713	17796756	2		None
bK299D3	Z84468	32481166	2	-	None
cN21F1	Z94162	33107523	3	-	None
66C4	AC000050	34627340	2	telomeric	None

* as reported by CWRU browser: Segmental duplication database on the UCSC website. Seg dup = clone contains a segmental duplication, UL= Unlocated, Contig not mapped to any chromosome

Several of the clones under-reporting copy number changes contained segmental duplications. Regions with homology on other chromosomes will cross hybridise with DNA from other chromosomes. This cross hybridisation will depress the ratios reported. For example, if all the sequence within a clone is duplicated on another

chromosome there will be four copies present within the genomic DNA. Addition of an extra copy of chromosome 22 into the hybridisation mix will result in an 5:4 ratio as compared to the 3:2 ratio if the clone contained unique sequence. In this way copy number changes will be underestimated for regions of segmental duplication involving other chromosomes. Other clones under reporting copy number change were located adjacent to the centromere or telomere and contain an abnormal amount of common repeat elements. The incomplete suppression of common repeat elements by Cot 1 may lead to the under or over reporting of chromosome 22 DNA copy number.

One clone that was hypersensitive to the chromosome 22 DNA was the clone cN61D6 (Accession no. D87012). This is located in the region encoding the immunoglobulin light chain λ. This has been shown to be deleted in some lymphoblastoid cell lines (4.3.3 and 7.3). The increased ratio reported in response to the additional copies of chromosome 22 may therefore be due to the single copy of chromosome 22 for this region within the genomic DNA used for the hybridisation.

This is an intrinsic problem with an array containing DNA representing an entire chromosome. More detail about regions of the array that contain a significant amount of segmental duplication are detailed in section 7.4.1. Clones with a high quantity of common repeats may not be fully blocked by the inclusion of Cot 1 in the hybridisation mix. Again this would result in cross hybridisation with other regions of the genome, under reporting the response to chromosome 22 DNA. This should be taken into account when reporting data from these clones.

4.7.3: Control hybridisations on the 500bp PCR product array

The standard deviation reported by a self:self hybridisation on the high resolution PCR product array was 2.5 times the standard deviation of the 22 tile path array and reflects the excessive noise shown around the 1:1 ratio. The standard deviation reported by a male:female hybridisation was also 2.5 times the standard deviation observed for the chromosome 22 array. The ratio reported by the chromosome X clones on the array did not represent the expected ratio for a full single copy number loss. The average standard deviation reported by the chromosome X clones was also

much greater than the standard deviation of the chromosome X clones on the chromosome 22 tiling path array.

It was noted that the intensities of the Cy3 and Cy5 signals from the PCR product arrays were considerably reduced compared to the large insert arrays. On average the signal intensities were 100 times less than those reported for the arrays spotted from DOP amplified clone products. The signal:background ratio of each spot on the array was therefore higher than for the clone-based arrays reducing the sensitivity and reproducibility of the PCR-based arrays The reduced intensities could be due to the PCR products spotted onto the array being smaller than those used for the clone arrays, or because they undergo one less round of amplification before spotting onto the array so that the final concentration of DNA in the spotting buffer is decreased. One way to increase the DNA concentration may be to include an extra round PCR amplification, prior to spotting of the products onto the array.

The self:self hybridisation on the PCR product array also revealed regions which show reduced ratios. This can be seen on Figure 4.9 located 18.62-18.76 and 19.82-19.90 Mb along 22q. This could be due to a labelling bias, where Cy3 and Cy5 are incorporated with different efficiency into GC or AT rich DNA. To test this hypothesis, the correlation between the GC content of the PCR product and the ratio reported by the self: self hybridisation, for a random selection of loci, was plotted. As seen on Figure 4.15 there is no correlation between GC content of sequence and ratio reported, so it is unlikely a labelling bias is responsible for the high standard deviations of the ratios observed.

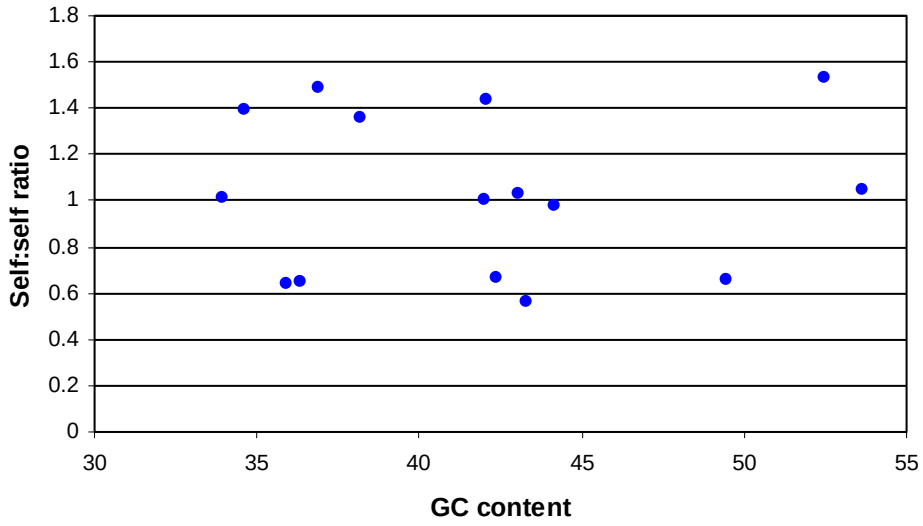


Figure 4.15: Correlation between GC content of PCR product and the ratio reported by a self:self hybridisation.

The average male:female hybridisation ratios reported on the X loci was 0.67. Again the standard deviation of ratios were larger than those seen for the clone arrays, but the average ratio reported for the X clones was comparable.

4.7.4: Summary

In summary, the clone array verification experiments showed that the reporting of copy number change by the constructed 22q tile path array was accurate. The chromosome add-in experiments showed that a vast majority of the clones represented on the array reported the expected response to additional copies of chromosome 22. It was concluded that the arrays were suitable for detecting the small copy number changes necessary for the assay of replication timing.

The verification experiments performed on the PCR product array showed wide variation in the ratios reported by control experiments indicating that these arrays would be less sensitive to replication timing differences compared to clone based arrays.