5 Results: Investigation of chromosome translocations using custom-made libraries

5.1 Introduction

The chromosome breakpoints of apparently balanced translocations can exhibit additional complexity when studied in detail, in some cases leading to difficulties in generating sequence across the translocation junctions (Puissant et al. 1988; Giacalone and Francke 1992; Borg et al. 2002; Cox et al. 2003; Rodriguez-Perales et al. 2004). Translocation breakpoint mapping by array painting onto microarrays of varying resolution (1Mb, WGTP, fosmid, PCR product, oligonucleotide) followed by long range PCR as described in Chapter 3 has successfully identified the breakpoints for the cases studied in this thesis as these patients carried balanced reciprocal translocations. For cases that are not balanced around the translocation breakpoints, additional methodologies of breakpoint mapping are required. For this reason, we have investigated the generation and screening of custom-made libraries to isolate clones containing the translocation junctions. The sequencing of these clones and the analysis of the sequence generated (approximately 36Kb) will reveal the translocation junction in addition to providing an insight into the complexity of the rearrangement. This chapter assesses 2 methods for creating a custom-made library; one with cosmids (using an in-house protocol) and one with fosmids (using a commercial kit) and the subsequent screening of the library using Alu PCR, PCR product pools and end sequencing.

5.2 Comparison of custom-made library protocols

Two methods for creating custom-made libraries were assessed; the first used enzymes to fragment the DNA which was subsequently ligated into a cosmid vector, and the second used a syringe needle to shear the DNA which was subsequently ligated into a fosmid vector. Both systems have a similar resolution

of approximately 40Kb as they use lambda phage packaging systems. Whole genomic DNA from a patient with a constitutional translocation t(7;13)(q31.3;q21.3) previously characterised by array painting and PCR (Chapter 3) was used to assess these procedures as viable methods of custom-made library production.

5.2.1 Cosmid library

Seven tubes each containing 150ng of DNA were competitively digested using the restriction enzymes *Mbo1* and dam methylase. Different proportions of *Mbo1*: dam methylase were used to obtain the best level of digestion (resulting in fragments around 36Kb in size). Ratios of *Mbo1*: dam methylase used were Tube 1.0 at 1:15,920, Tube 1.1 at 1:7,920, Tube 1.2 at 1:3,120, Tube 1.3 at 1:2,320, Tube B+ at 1:720, Tube B- at 1:720 and Tube dam containing dam methylase only (see Table 2.4 in methods). All tubes (except Tube B-) were treated with calf intestinal alkaline phosphatase to dephosphorylate the ends of the DNA fragments so preventing the formation of concatamers. The results of the digestion, along with the undigested control tubes are shown in Figure 5.1.

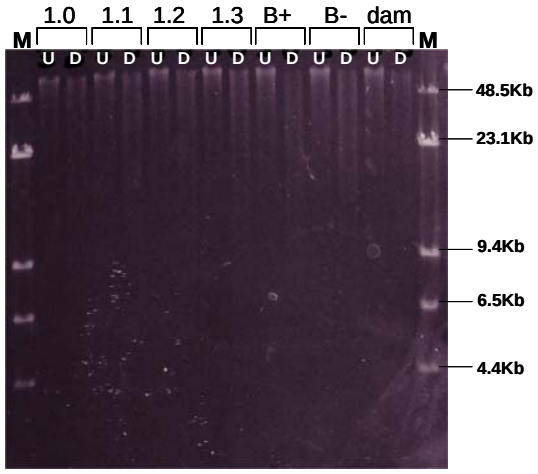

Figure 5.1 Results of competitive digestion using Mbo1 and dam methylase on whole genomic DNA from a translocation patient. U; undigested control, D; digested DNA for each enzyme dilution (1.0, 1.1, 1.2, 1.3, BASIC+, BASIC- and dam) run against M; λ DNA digested with HindIII.

The tube containing the enzyme mix showing the best digestion (in this case, tube 1.0) was selected for ligation to vector arms for production of the library. After ligation, packaging and plating, the library was incubated overnight to allow for colony growth, however, no colonies were observed.

5.2.2 EpiFOS fosmid library

5.2.2.1 Using whole genomic DNA

Whole genomic DNA was sheared by passing through a 21 gauge syringe needle five times and ten times and the fragment size assessed by gel electrophoresis (Figure 5.2).

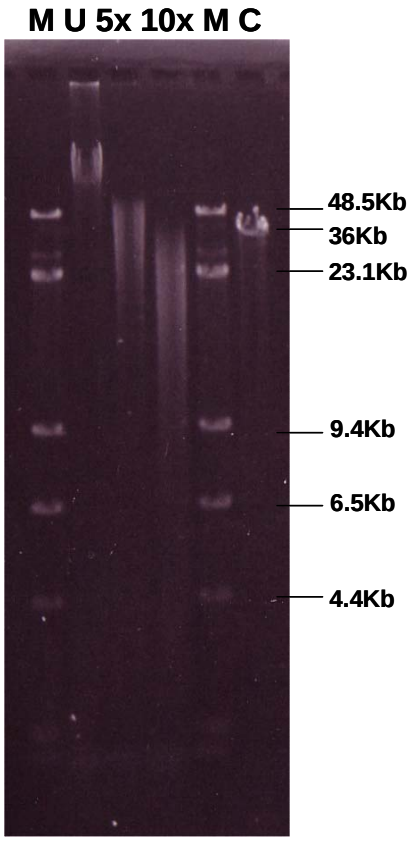

Figure 5.2 Results of shearing of whole genomic DNA from a translocation patient. U; Unsheared DNA, 5x; DNA sheared by passing 5 times through syringe needle, 10x; DNA sheared by passing 10 times through syringe needle run against M; λ DNA digested with HindIII and C; Control DNA of 36Kb.

Fragment sizes of 30 to 44Kb are required for packaging into the phage head. The gel showed that the unsheared genomic DNA was too large, and that the DNA passed 10 times through the needle was too small however the DNA passed 5 times through the needle had an average fragment size of 36Kb, the same size as the fosmid control DNA (DNA fragments at the specified size of 36Kb) which was run on the gel as a control marker. This 5 times sheared DNA was used to create the library and the packaged phage particles were plated onto agar. Clones were selected at random for FISH to assess the successful incorporation of genomic DNA into the vector. Figure 5.3 shows the FISH signal

for one such clone which was seen to hybridise to chromosome 5p15.31-5p15.33.

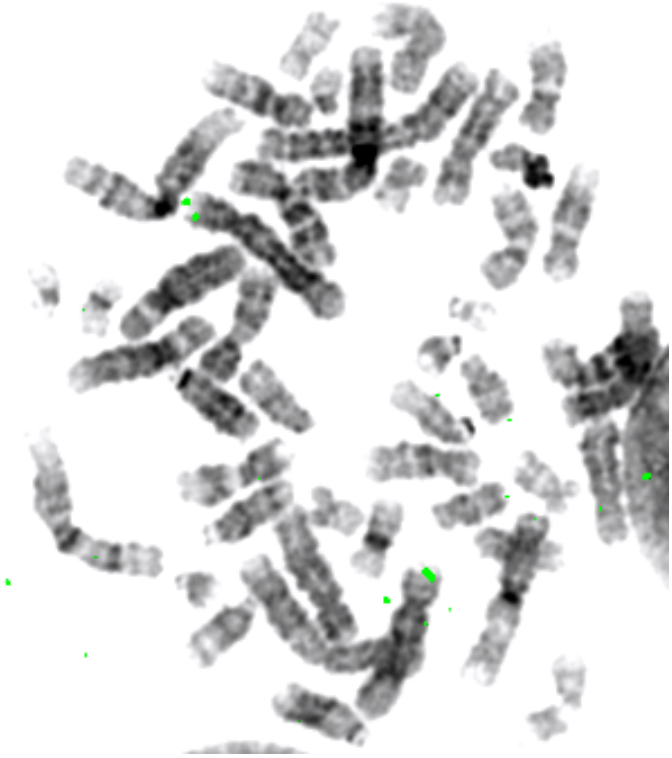

Figure 5.3 Inverted DAPI banding of metaphase chromosomes and FISH signal for a fosmid clone selected at random from the test library created using whole genomic DNA from patient t(7;13)(q31.3;q21.3). The clone hybridised to 5p15.31-5p15.33.

5.2.2.2 Using flow sorted derivative chromosomes

In order to assess the amount of starting material required to make a library suitable for screening, 2,500,000 nuclei were flow sorted and divided into 1000ng, 500ng and 250ng aliquots. Each aliquot was handled in the same way and used to generate individual libraries. Shearing of each aliquot of DNA 3 times through a syringe needle resulted in DNA fragments of approximately the same size regardless of the amount of DNA contained within the sample (Figure 5.4).

Flow sorted derivative chromosomes exhibit a lower molecular weight than isolated genomic DNA, hence reducing the amount of shearing required.

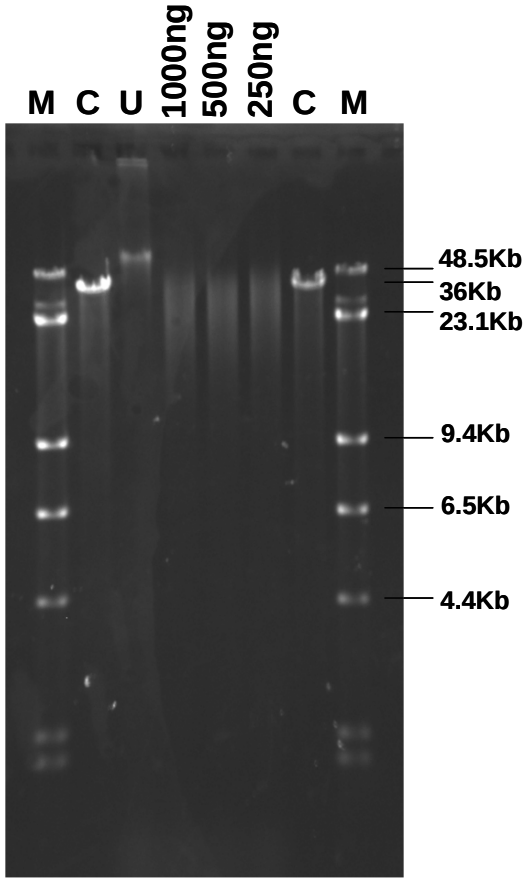

Figure 5.4 Results of shearing using starting amounts of 1000ng, 500ng and 250ng DNA. C; Control DNA of 36Kb, U; Unsheared DNA, 1000ng, 500ng and 250ng aliquots of DNA sheared by passing 3 times through syringe needle run against M; λ DNA digested with HindIII.

Each library was plated out and the number of colonies counted. The libraries generated from 1000ng starting material gave approximately 900 colonies, from 500ng gave approximately 600 colonies and from 250ng gave approximately 200 colonies. These colony numbers were obtained using 1/100th of the library suspension. Extrapolating these figures meant that the 1000ng library would generate 90,000 clones, the 500ng library would generate 60,000 clones and the

250ng library 20,000 clones. This equated to 28x coverage of the genome for the 1000ng library, 19x coverage for the 500ng library and 6x coverage for the 250ng library. Generation of a library from 250ng of starting material was selected as a method that provided reasonable coverage (approximately 3x to 18x coverage depending on the size of the derivative chromosome), whilst minimising the time required for flow sorting of the derivative chromosomes.

5.2.3 Conclusions

Use of the cosmid library protocol required 1050ng of DNA as starting material compared with the 450ng required for the fosmid library approach. This has a major impact on the time needed to flow sort the required number of derivative chromosomes (approximately 2 days for 450ng versus 5 days for 1050ng DNA).

Both library production systems utilised different methods of DNA fragmentation to reduce the molecular weight of the DNA to approximately 36Kb for packaging into the phage head; for the cosmid library, the DNA was fragmented by enzyme digestion and for the fosmid library the DNA was fragmented by shearing through a syringe needle. Size reduction of the DNA by enzyme digestion may produce a bias due to the precise nature of the fragmentation sites, whereas size reduction via shearing should be a random process and not introduce any bias.

For these reasons, and the lack of cosmid clones obtained, the fosmid library procedure was selected over the cosmid library procedure for generation of a custom-made library.

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

A fosmid library was created from a patient carrying a balanced reciprocal t(7;13)(q31.3;q21.3) translocation that had previously been successfully sequenced across the translocation junctions. This library was used to investigate strategies of screening for chimeric clones containing the translocation junctions. Derivative chromosomes 7 and 13 were flow sorted (Figure 5.6) and used to create two separate libraries. 450ng of each derivative chromosome (1,250,000 derivative chromosome 7 and 2,000,000 derivative chromosome 13) was flow sorted, allowing 250ng for generation of the library, and 200ng for gel electrophoresis to assess the DNA fragment size before and after shearing.

Both libraries were generated and grown on round plates as described in the methods. These plated libraries were used to generate two different set of filters for screening. Firstly, a robot was used to pick off colonies into 384 well plates and the clones gridded onto filters and secondly replica round filters were generated from the original round library plates.

The gridded filters were used to assess the methods of screening detailed in section 5.4. The round replica filters representing the whole library were used to assess the library coverage and success in generation of chimeric breakpoint spanning clones.

5.4 Comparison of screening methods using gridded library fosmid clones

A total of 14,208 clones were picked at random from the derivative chromosome 7 library and 13,824 from the derivative chromosome 13 library. These clones represented an estimated 2.8x coverage of the derivative chromosome 7 and

4.5x coverage of the derivative chromosome 13. Clones were picked into 384 well plates and gridded onto filters with a maximum of 16 plates per filter. The derivative chromosome 7 and the derivative chromosome 13 libraries were both gridded onto 3 filters each enabling efficient screening.

5.4.1 Inter Alu PCR screening

Using the translocation spanning BAC clones identified by FISH as template (RP11-384A20 for chromosome 7 and RP11-360I23 for chromosome 13) probes were generated by PCR to amplify the unique sequence present between Alu repeats within the sequence (Figure 5.5).

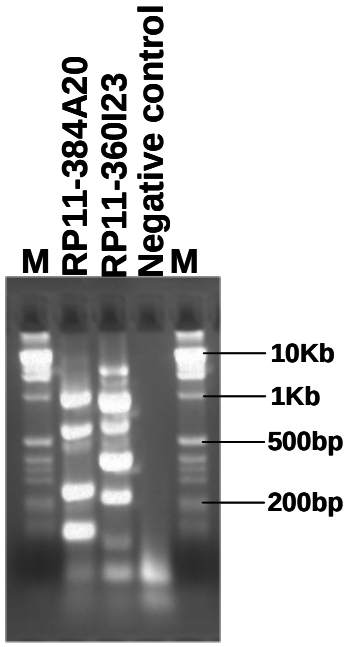

Figure 5.5 Results of Alu PCR amplification assessed by gel electrophoresis with 1Kb marker.

Hybridisation of these probes to the gridded library filters resulted in low level background, but failed to identify specific clones.

5.4.2 PCR product pools

PCR product pools were created containing specific PCR products. PCR primers were designed to amplify one product every 10Kb along the sequence of the spanning BAC clone previously identified by FISH. The chromosome 7 pool consisted of 16 products, and the chromosome 13 pool of 18 products (Appendix A4).

Hybridisation of the chromosome 7 pool of products to the derivative chromosome 7 library identified 12 clones; der7_1C5, der7_1F16, der7_15N14, der7_21H5, der7_23K18, der7_26N14, der7_27P5, der7_35D10, der7_35F10, der7_37B6, der7_37H21 and der7_37N12. Screening the same library with the chromosome 13 pool of products identified 5 clones; der7_2I21, der7_8N13, der7_12B20, der7_28J18 and der7_37N12. One clone; der7_37N12 was identified by both pools implying that the clone contained both chromosome 7 and chromosome 13 DNA and potentially spanned the derivative chromosome breakpoint.

Hybridisation of the chromosome 13 pool of products to the derivative chromosome 13 library identified 2 clones; der13_34D21 and der13_36M16, however, the chromosome 7 pool did not identify any clones. The low numbers of clones identified during the screening suggested that the coverage of this library is poor around the breakpoint region. It was unlikely that a derivative chromosome 13 junction spanning clone is contained within this gridded library as the pools failed to identify a common clone.

5.4.3 End sequencing

A total of 28,032 colonies were picked by the robot from the round agar plates into 384 well plates for gridding and end sequencing; 14,208 from the derivative chromosome 7 library and 13,824 from the derivative chromosome 13 library. Of these, 384 failed to grow; 244 from the derivative chromosome 7 library and 140

from the derivative chromosome 13 library. The remaining clones (13,964 derivative chromosome 7 clones and 13,864 derivative chromosome 13 clones) were end sequenced with a high success rate (Table 5.1) and the resulting data compared with the human reference sequence to investigate the origin of the insert clone DNA with the ultimate aim of finding clones chimeric for chromosome 7 and 13 sequence, and hence containing the translocation breakpoints.

Number of clones	der7	der13
Submitted for end sequencing	13,964	13,684
Passed both ends	12,460 (89.2%)	10,009 (73.1%)
Passed one end only	989 (7.1%)	1,426 (10.4%)
Failed both ends	515 (3.7%)	2,249 (16.5%)

Table 5.1 Summary of end sequencing reads obtained from derivative chromosome 7 and derivative chromosome 13 library clones.

All clones which produced sequence greater than 20bp in length were compared with the human reference sequence to align the clone sequence to it. 88.9% (11,081/12,460) of clones from the derivative chromosome 7 library and 84.2% (8,428/10,009) of clones from the derivative chromosome 13 library were successfully aligned to the human reference sequence (Table 5.2).

der7	der13
6550 (59.1%) on chr7	3600 (42.7%) on chr7
2225 (20.1%) on chr13	4201 (49.8%) on chr13
2022 (18.2%) on chr6	249 (3.0%) on chr14
265 (2.4%) on others	368 (4.4%) on others

Table 5.2 Summary of fosmid insert DNA alignments against the human genome reference sequence for the derivative chromosome 7 and derivative chromosome 13 library clones.

In the derivative chromosome 7 library, 79.2% of clones hit chromosomes 7 or 13 as expected. A further 18.2% aligned to chromosome 6. For the derivative chromosome 13 library, 92.5% of clones aligned to chromosomes 7 or 13 as expected and a further 3.0% aligned to chromosome 14. Analysis of the patient's

flow karyogram revealed that the derivative chromosome 7 lay close to chromosome 6 and the derivative chromosome 13 lay close to chromosome 14 (Figure 5.6) explaining the contamination seen in the libraries. During the flow sorting process, a small number of chromosomes can become fragmented. This DNA, once carried through the library preparation procedure would explain the fosmid clones that show DNA inserts that aligned to other chromosomes.

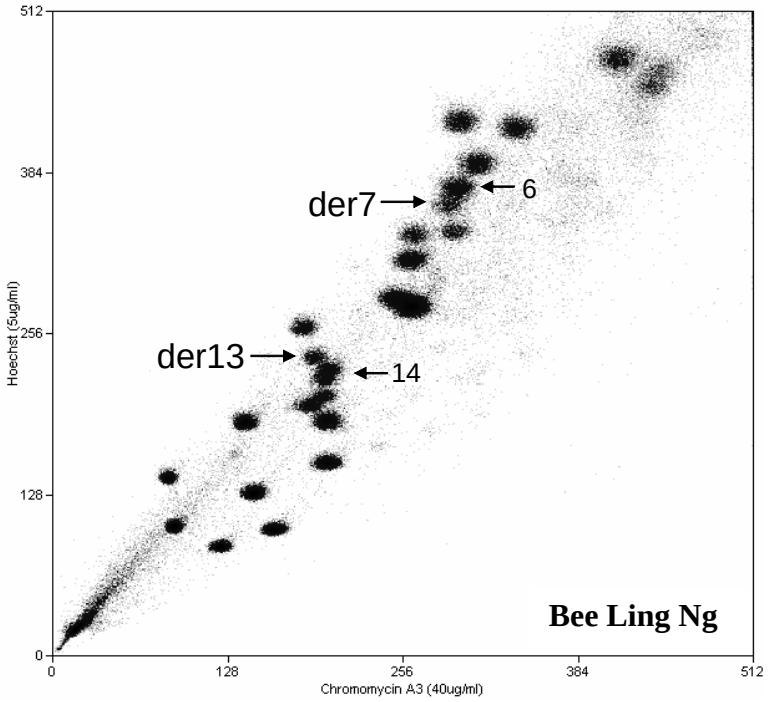

Figure 5.6 Flow karyogram for patient t(7;13)(q31.3;q21.3). The derivative chromosome 7 was close to the normal chromosome 6 and the derivative chromosome 13 close to the normal chromosome 14.

19 clones from the derivative chromosome 7 library and 10 clones from the derivative chromosome 13 library could not be aligned to the human reference sequence because the two ends of each clone aligned to different chromosomes. Further analysis of the end sequences revealed that all 29 of these clones were chimeric for both chromosome 7 and chromosome 13 sequence. One clone

(der7_37N12) was chimeric for chromosome 7 and chromosome 13 sequences at the correct locations, and therefore potentially contained the translocation junction sequence for the derivative chromosome 7. The remaining 28 clones which mapped to unexpected regions of chromosomes 7 and 13 must have been produced as a by-product of the fosmid library construction; incomplete phosphorylation of the ends of the insert DNA could have allowed smaller fragments of DNA to join, resulting in an artificially created fragment of DNA that is the correct size to be packaged into the phage head.

5.4.4 Specific breakpoint PCR product

In order to assess the number of chimeric fosmid clones present in the gridded libraries, two PCR products which spanned the derivative chromosome 7 and 13 translocation junctions were created (Table 5.3).

Primer	Forward primer	NCBI Build 35 position	Reverse primer	NCBI Build 35 position
der7	GTAGTGATTCGGCCTTGCAT	chr7:121245417-121245436	TGGCCATATTTGGCTTTTTG	chr13:71023936-71023955
der13	TCCATTCATGTTGCTGCATT	chr13:71023730-71023749	GGAAGACAGGATGGATTCAAA	chr7:121245470-121245490

Table 5.3 PCR primers used to create derivative chromosome specific products for library screening.

Screening of the derivative 7 library with the derivative 7 specific breakpoint PCR product identified 3 fosmid clones; der7_12B20, der7_28J18 and der7_37N12. Screening of the derivative 13 library with the derivative 13 specific breakpoint PCR product identified a single fosmid clone; der13_10O8. Analysis of the end sequence data previously obtained for these clones showed that der7_37N12 was chimeric for chromosome 7 and 13 sequence and contained the derivative chromosome 7 translocation junction. However, the other three fosmid clones aligned to regions of the genome not believed to be involved in the translocation (Table 5.5). The insert sizes for these clones were 41,206bp for der7_12B20, 36,062bp for der7_28J18 and 34,447bp for der13_10O8. These clones were subsequently analysed by FISH.

Fosmid clone der7_12B20 hybridised mainly to chromosome 7q21.2-7q31.3 with weaker signals on 13q21.31-13q31.2 and 7p12.1-7p14.2. Analysis of the end sequence data showed that the clone localised to 7p12.1 of the human genome reference sequence. These results suggested that the well contained mixed colonies due to incomplete separation of phage clones during the plating out stage of library production.

FISH analysis of der7_28J18 showed a hybridisation signal at 6q21-6q22.33 whilst the end sequence data aligned the clone to 6q22.1. Blast analysis of the sequence for this region showed no homology to the specific breakpoint product used for library screening, suggesting that the well must contain a mixture of 2 clones from poor separation during plating.

Analysis of der13_10O8 showed alignment to 13q13.1 by end sequence analysis and hybridisation to 13q12.3-13q14.13 and 13q14.3-13q21.3 by FISH. This FISH data showed that the clone lay close to or at the translocation breakpoint. This region may have been sufficient for the breakpoint specific PCR product used during radioactive screening to hybridise to, producing a positive result. The more proximal FISH signal may have been the result of a mixed well.

In order to confirm that these 3 clones were from mixed colonies, each well would need to be streaked to single colony and re-analysed.

5.4.5 Summary of fosmid clones identified by the different screening methods

A summary of the clones identified by the different screening strategies is depicted in Table 5.4.

		derivative chromsome 7 library	derivative chromsome 13 library
PCR pool screening	chr7 pool	der7_1C5, der7_1F16, der7_15N14, der7_21H5, der7_23K18, der7_26N14, der7_27P5, der7_35D10, der7_37B6, der7_37H21, der7_37N12	None
	chr13 pool	der7_2I21, der7_8N13, der7_12B20, der7_28J18, der7_37N12	der13_34D21, der13_36M16
Specific breakpoint product		der7_12B20, der7_28J18, der7_37N12	der13_10O8

Table 5.4 Summary of fosmid clones identified by PCR pool screening and screening with specific breakpoint products. der7_37N12 was identified by both approaches as containing the derivative chromosome 7 translocation junction.

For the derivative chromosome 7 library, screening by PCR product pools and using a specific breakpoint product identified clone der7_37N12 which had previously been identified by end sequencing to contain the derivative chromosome 7 translocation junction sequence.

None of the screening strategies identified a chimeric clone in the derivative chromosome 13 library. A likely explanation for this is that the gridded library did not actually contain a derivative chromosome 13 spanning fosmid.

The end sequence data for all clones detailed in Table 5.4 was retrieved and summarised in Table 5.5. All clones identified by PCR product pools were seen to align close to the translocation breakpoints. These clones were identified because the PCR products used to screen the library were designed at a 10Kb resolution across the whole length of the spanning BAC clone.

Clonename	Primer	BLAT hit (May 2004)		
		chr	start	end
der7_1C5	Forward	no data		
	Reverse	no data		
der7_1F16	Forward	7	121,159,146	121,159,888
	Reverse	7	121,195,221	121,194,772
der7_2I21	Forward	13	70,922,368	70,923,133
	Reverse	13	70,958,461	70,957,741
der7_8N13	Forward	13	71,085,684	71,086,475
	Reverse	13	71,123,862	71,123,035
der7_12B20	Forward	7	51,986,318	51,986,666
	Reverse	7	51,944,751	51,945,460
der7_15N14	Forward	no data		
	Reverse	no data		
der7_21H5	Forward	7	121,199,488	121,200,334
	Reverse	7	121,236,279	121,235,430
der7_23K18	Forward	7	121,195,574	121,196,211
	Reverse	7	121,229,567	121,228,758
der7_26N14	Forward	no data		
	Reverse	7	121,154,377	121,155,087
der7_27P5	Forward	7	121,118,879	121,185,690
	Reverse	7	121,221,332	121,220,529
der7_28J18	Forward	6	117,769,406	117,770,080
	Reverse	6	117,805,731	117,806,142
der7_35D10	Forward	7	121,183,478	121,184,256
	Reverse	7	121,223,748	121,223,501
der7_35F10	Forward	7	121,129,871	121,130,669
	Reverse	7	121,165,835	121,165,011
der7_37B6	Forward	7	121,160,129	121,160,933
	Reverse	7	121,198,097	121,197,251
der7_37H21	Forward	7	121,193,308	121,193,601
	Reverse	no data		
der7_37N12	Forward	13	71,035,495	71,034,670
	Reverse	7	121,226,143	121,226,939
der13_10O8	Forward	13	32,131,659	32,132,406
	Reverse	13	32,166,106	32,165,301
der13_34D21	Forward	13	70,922,561	70,923,411
	Reverse	13	70,955,023	70,954,228
der13_36M16	Forward	13	70,959,057	70,959,865
	Reverse	13	70,994,435	70,993,649

Table 5.5 End sequence data available for all clones identified by alternative screening strategies.

5.4.6 Conclusions

In total, 14,208 colonies were picked from the derivative chromosome 7 agar plates, resulting in an estimated 2.8x coverage along the chromosome. For derivative chromosome 13, 13,824 colonies were picked resulting in an estimated 4.5x coverage. These gridded clones were used to generate filters used to establish a method of screening custom-made translocation libraries.

End sequencing of every clone within the gridded libraries identified a potential spanning clone for the derivative chromosome 7 library, but not the derivative chromosome 13 library. However, it was possible to use the data obtained from the end sequencing of all clones to refine the mapping of the translocation breakpoint region on chromosome 13 where a spanning clone was not identified. Analysis of the alignment of the clones from both libraries to the human reference sequence along chromosome 13 revealed a switch from one chromosome to the other (Figure 5.7) so indicating the breakpoint region. The different depths of coverage for each library (2.8x for the derivative chromosome 7 library and 4.5x for the derivative chromosome 13 library) are also illustrated in Figure 5.7.

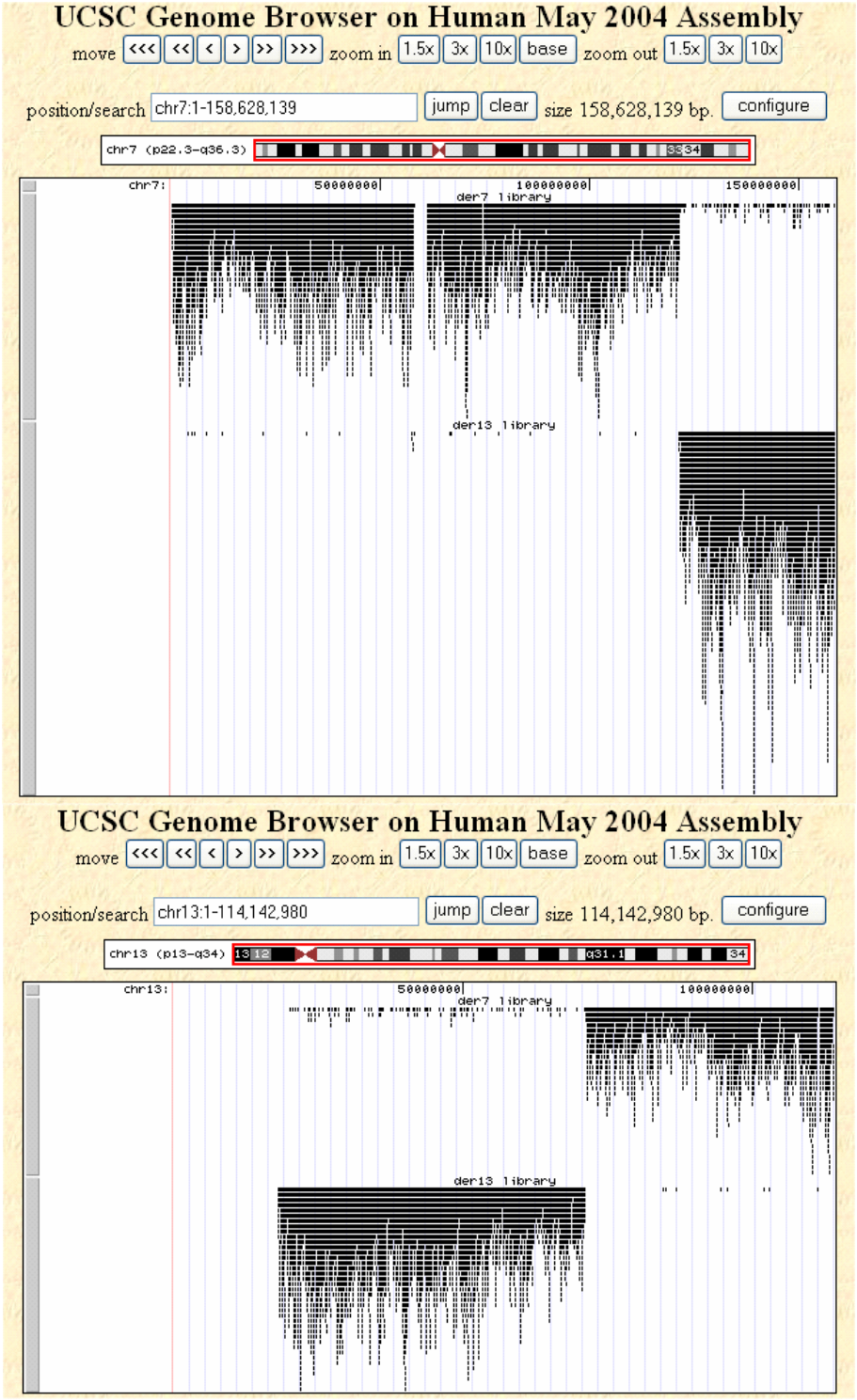

Figure 5.7 UCSC download showing clone coverage across a; chromosome 7 for both derivative chromosome 7 and 13 libraries and b; chromosome 13 for both derivative chromosome 7 and 13 libraries.

In cases where no spanning clone was identified, as in the case of the derivative chromosome 13, the translocation junction could be mapped between the clones mapping to chromosome 13 in the derivative chromosome 13 library and the clones mapping to chromosome 13 in the derivative chromosome 7 library (Figure 5.8). The breakpoint was mapped between clones 3615_13_36m16 from the derivative chromosome 13 library and Der7_8n13 from the derivative chromosome 7 library. The chromosome 13 breakpoint region that fell between these clones was from 70,966,023 to 71,085,684bp, a gap of 130,611bp. Further mapping by STS PCR would have been required in this situation to refine the breakpoint region prior to LR PCR.

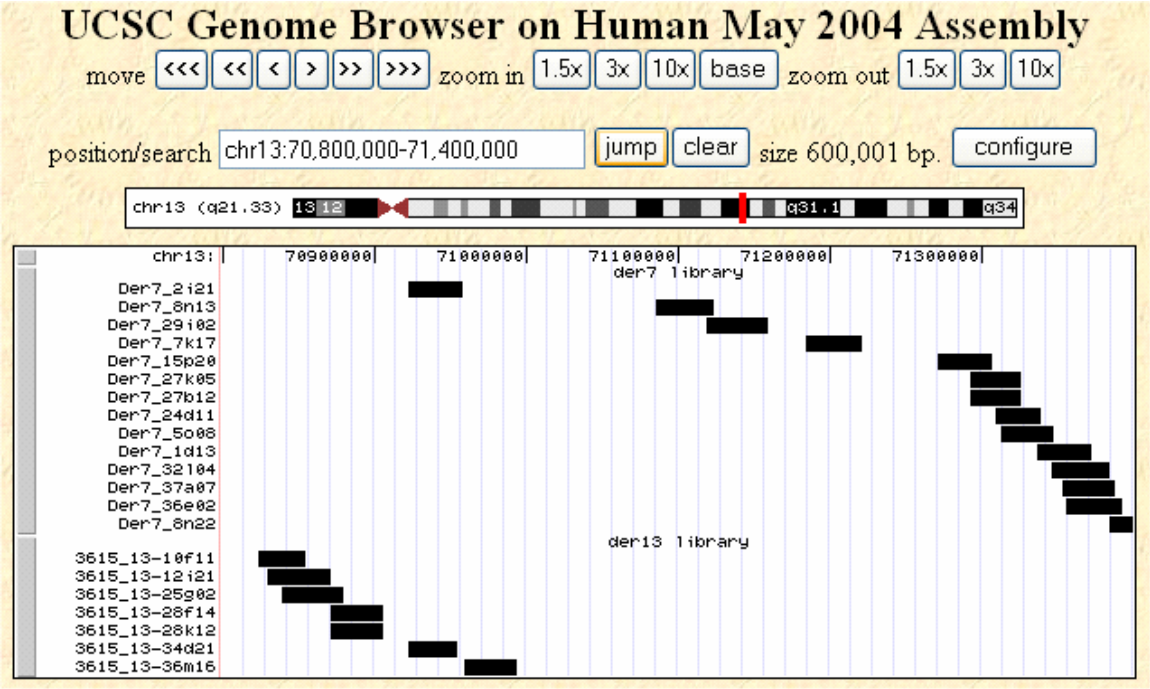

Figure 5.8 Fosmid clone coverage across the breakpoint region of chromosome 13 for the derivative chromosome 7 and 13 libraries. No spanning clone was identified by end sequencing across this breakpoint. Clone der7_2i21 from the derivative chromosome 7 library aligned proximal to the chromosome 13 breakpoint. This clone must have been generated from a chromosome 13 flow sorting artefact as the derivative chromosome 13 and normal chromosome 13 peaks were close to each other on the flow karyogram (Figure 5.6).

Screening the libraries by end sequencing and PCR product pools both identified the derivative chromosome 7 breakpoint spanning clone. The data generated by end sequencing provided a valuable resource for assessing and verifying results obtained by screening with PCR product pools. In this instance the end sequence data confirmed the simplicity of the rearrangement but in patients with complex rearrangements the end sequence data may help to resolve the rearrangement structure. However, end sequencing of an entire library remains an expensive process, so may not be a viable option for routine breakpoint mapping. Screening by PCR product pools involves the hazards associated with radiation but remains a faster and less costly method than screening by end sequencing.

5.4.7 Screening of round library plates

Screening of the derivative chromosome 7 gridded library with 2.8x coverage by end sequencing found 1 translocation junction spanning clone but the derivative chromosome 13 library with 4.5x coverage did not reveal a translocation spanning clone.

The clones in the gridded version of the library were picked by robots which could only select clones that were discreet, circular and larger than 1.5mm in diameter. For this reason, the gridded library did not contain all the clones originally created and plated out. To investigate whether translocation spanning clones were present on the round library plates, but had failed to be selected by the robots, lift filters were created from the original plated libraries for screening.

The same breakpoint specific products used to screen the gridded filters were used to screen the round filters. Due to the construction of the robot, the plates were not picked in a sterile environment, and unfortunately some plates were lost due to infection. The derivative chromosome 7 library was originally plated onto 115 plates, and subsequently 39 were unusable by the time it came to make the

filters. For the derivative chromosome 13 library, 31 of 87 plates were not available to be used for lift filter generation.

Screening of the derivative chromosome 7 library plates identified 2 clones, der7_round76 and der7_round78. Screening of the derivative chromosome 13 library plates identified 3 clones, der13_round14, der13_round60 and der13_round80. To verify that the identified clones contained the translocation junction, the clones were streaked to single colony, 10 colonies were selected and colony PCR using primers designed to amplify across the translocation junctions was used to verify that the translocation junction was present (Table 5.3). For the derivative chromosome 7 clones, der7_round76 could not be confirmed as containing the translocation junction, but der7_round78 was positive by colony PCR. For the 3 derivative chromosome 13 clones identified; der13_round 14 failed to grow, der13_round60 was confirmed by colony PCR and the plate containing der13_round80 exhibited infection, so could not be accessed for confirmation. Fosmid der13_round60 was subsequently sequenced in full and analysis of the sequence revealed that the clone sequence was identical across the derivative chromosome 13 translocation junction as had previously been identified by array painting and PCR (see Figure 3.29).

The screening of the round library filters showed that while the DNA fragments containing the derivative chromosome junction sequences were cloned during the library process, they were not necessarily selected by the robots for gridding and that as a result, the coverage of the gridded library was lower than the coverage actually generated.

Generation of the gridded library filters was a less time consuming process than generation of the round lift filters and screening of those gridded libraries also greatly reduced the amount of radioactive isotope required during the screening process. However, not all clones created in the library were selected for gridding,

emphasising the need to retain the original round library plates as a back-up if screening of the gridded library failed to identify chimeric clones.

5.5 Mapping translocation breakpoints using a custom-made fosmid library from a patient with a t(2;6)(q21.1;q25.1) translocation

To test the method of translocation breakpoint mapping using custom-made fosmid libraries, a patient with an uncharacterised reciprocal translocation was selected. This patient had not undergone any mapping other than routine G-banding which identified a t(2;6)(q21.1;q25.1) translocation (Figure 5.9). To map the translocation breakpoints prior to generation of a PCR product pool for screening of the library, flow sorted derivative chromosomes were array painted onto a WGTP microarray.

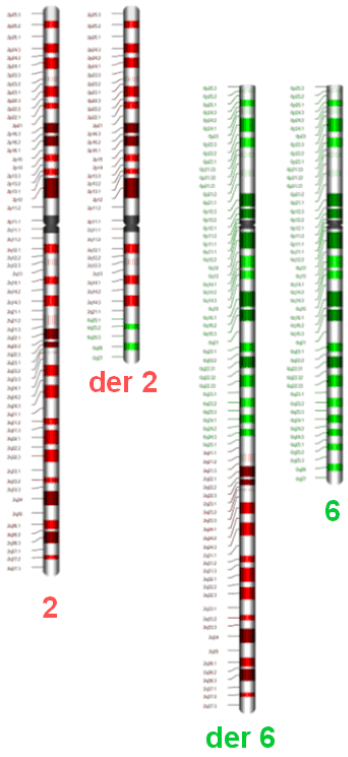

Figure 5.9 Ideogram images depicting normal chromosomes 2 and 6 and derivative chromosomes 2 (der 2) and 6 (der 6) in a patient with a t(2;6)(q21.1;q25.1) translocation.

5.5.1 Isolation of derivative chromosomes by flow sorting

Flow sorted derivative chromosomes were required to map the translocation breakpoints by array painting and for generation of the custom-made fosmid library. Assessment of the patient's karyotype (Figure 5.10) revealed that both derivative chromosomes were different from the normal homologues in size and base composition and therefore could be segregated by flow sorting. The expected size of the derivative chromosome 2 was approximately 160Mb and the derivative chromosome 6 approximately 287Mb.

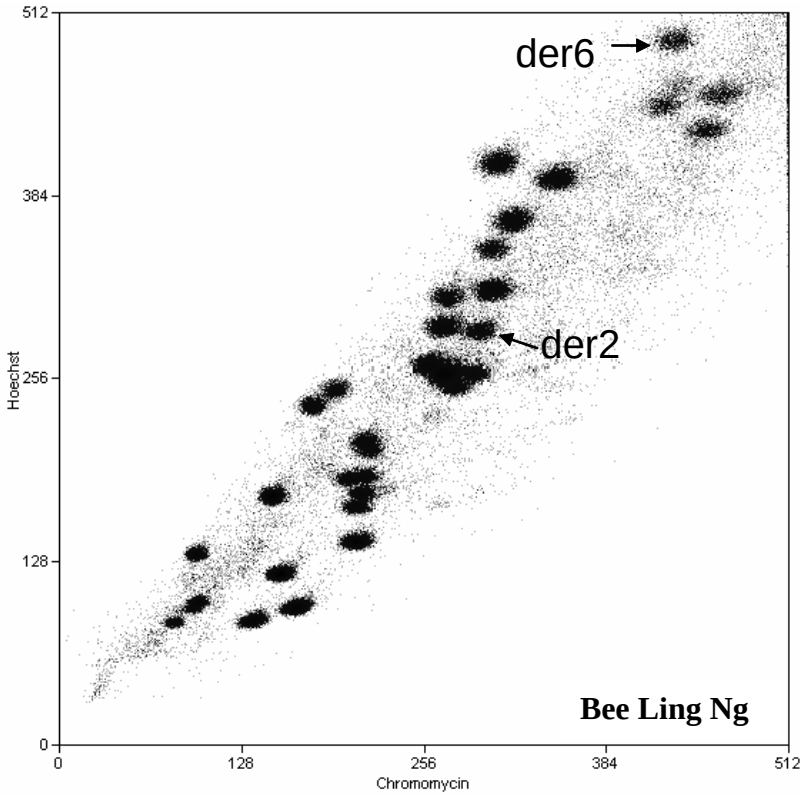

Figure 5.10 Flow karyogram for patient t(2;6)(q21.1;q25.1).

5.5.2 Verification of flow sorted derivative chromosomes by reverse chromosome painting.

The flow sorted derivative chromosomes were reverse painted onto metaphase chromosomes from a normal cell line as a method of verification and to assess the amount of contamination from other chromosomes (Figure 5.11).

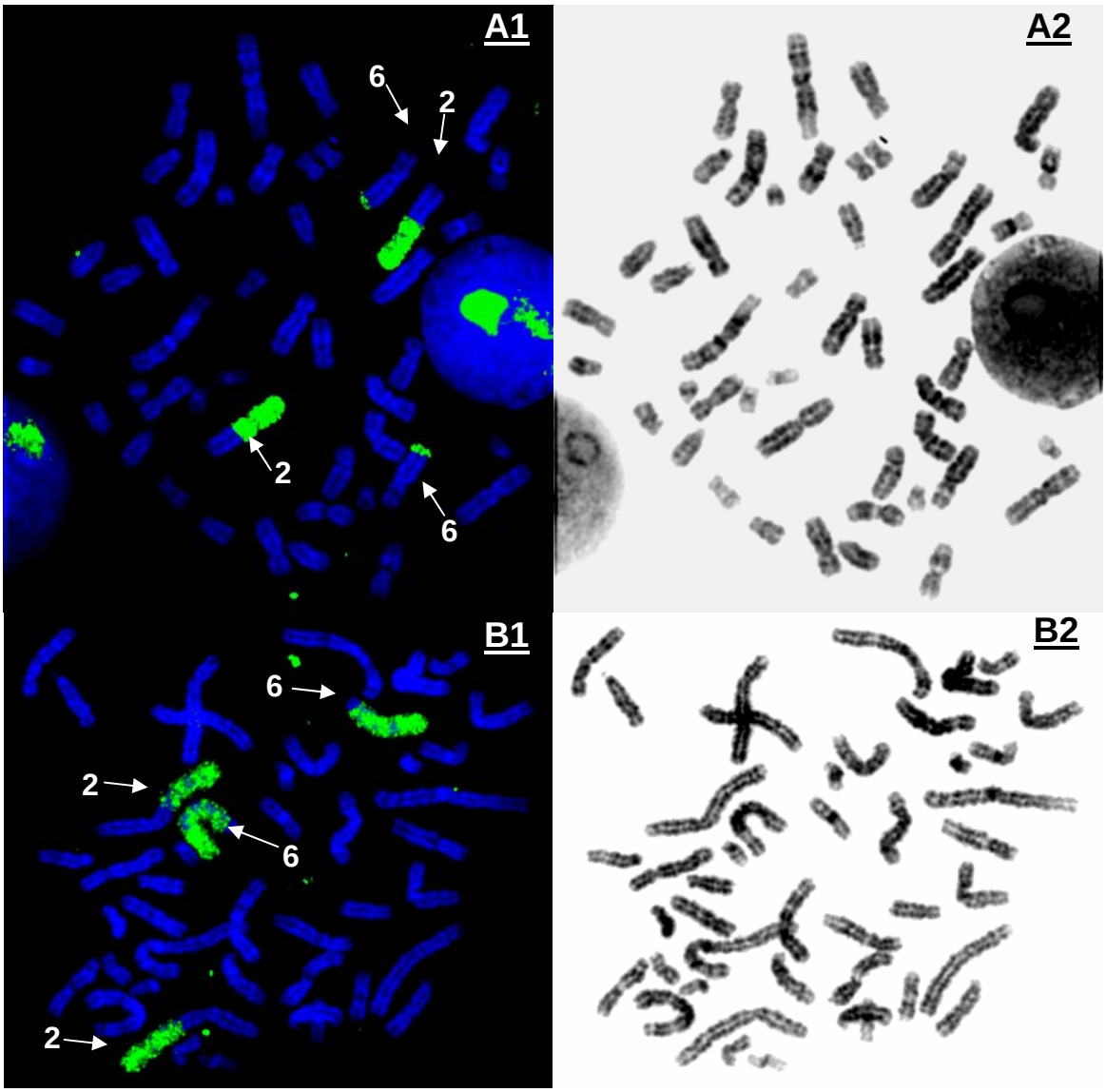

Figure 5.11 Reverse chromosome painting results for A flow sorted derivative chromosome 2 and B derivative chromosome 6. 1 metaphase FISH image and 2 inverted DAPI banding of metaphase chromosomes.

Reverse chromosome painting showed that the flow sorted derivative chromosome 2 hybridised to 2p and part of 2q and also the telomeric portion of chromosome 6q. The flow sorted derivative chromosome 6 hybridised to 6p and the majority of 6q and also the distal two thirds of 2q. Neither hybridisation detected any signal on chromosomes other than 2 and 6, suggesting that at the resolution afforded by metaphase chromosomes, the rearrangement was a simple reciprocal translocation, that no other chromosomes were involved and that the sort was not significantly contaminated.

5.5.3 Mapping translocation breakpoints by array painting

To refine the breakpoint regions from the cytogenetic bands assigned by G-banding analysis, the derivative chromosomes were differentially labelled and hybridised in an array painting experiment to a the Sanger 30K Whole Genome Tiling Path microarray (Figure 5.12).

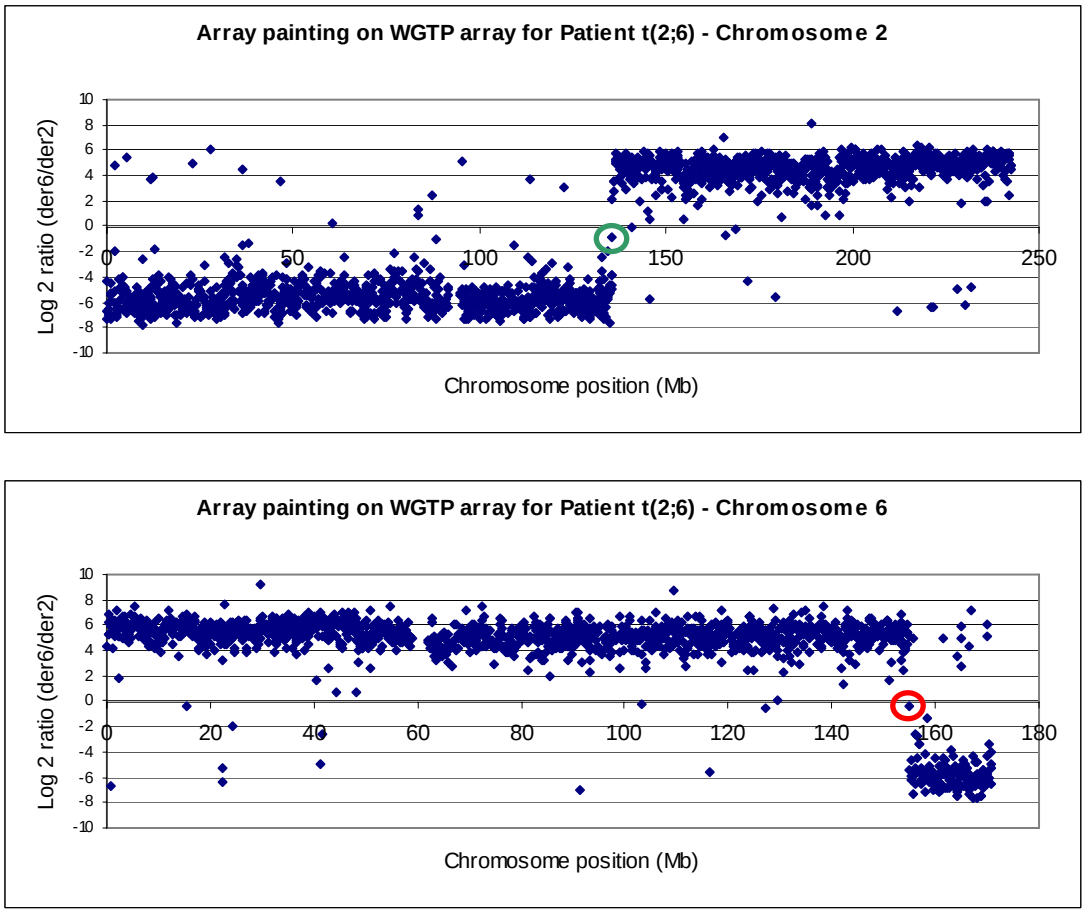

Figure 5.12 Array painting results for patient t(2;6)(q21.1;q25.1) on the whole genome tilepath microarray. Spanning clones; RP11-341H1 and RP11-486M3.

Analysis of the clone ratios in the array painting experiment revealed a single spanning clone on chromosome 2; RP11-341H1 and a single spanning clone on chromosome 6; RP11-486M3. Both spanning clones exhibited ratios close to zero implying that the translocation breakpoints lay towards the centre of the genomic DNA insert contained within the clone. Identification of the breakpoint regions at this resolution enabled the creation of PCR products pools for the screening of a custom-made library.

5.5.4 Isolation of breakpoint spanning fosmid clones using a custom-made library

5.5.4.1 Production of the t(2;6)(q21.1;q25.1) fosmid library

Approximately 450ng of both derivative chromosomes was flow sorted (corresponding to 1,750,000 derivative 2 chromosomes and 1,000,000 derivative 6 chromosomes) for generation of the derivative chromosome fosmid libraries. 450ng of chromosomes from the 9-12 peak was also sorted to be used as a control. The DNA was sheared by passing 3 times through a syringe needle (Figure 5.13).

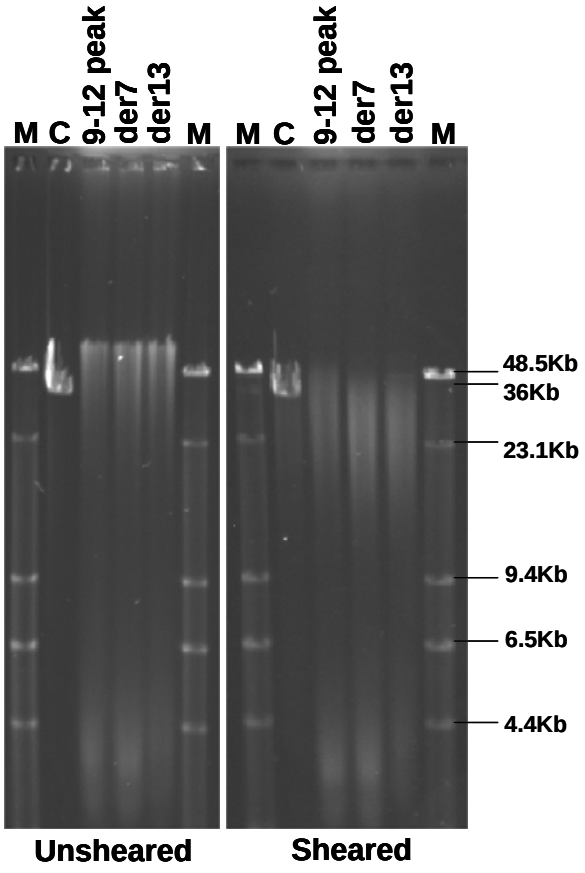

Figure 5.13 Results of DNA shearing of flow sorted chromosomes 9-12, derivative chromosome 2 and derivative chromosome 6 DNA for patient t(2;6)(q21.1;q25.1) run against M; λ DNA digested with HindIII and C; control DNA of 36Kb.

5.5.4.2 Screening of the t(2;6)(q21.1;q25.1) library using PCR product pools.

The sequence associated with both spanning BAC clones identified by the WGTP array paint experiment (Section 5.5.3) was used to design the pools of PCR products for screening of the custom-made fosmid library. The chromosome 2 spanning BAC, RP11-341H1 (Accession AC016725.5) had 162,180bp of sequence associated with it from 135,311,916 to 135,474,095bp. The chromosome 6 spanning BAC, RP11-486M3 (Accession AL357075.17) was 178,461bp in length from 154,757,465 to 154,935,925bp. The extended regions selected for primer design were from 135,311,000 to 135,475,000bp on chromosome 2 and 154,756,000 to 154,937,000bp on chromosome 6 which allowed for the possibility that the full insert of the clone had not been sequenced.

A total of 19,200 clones (4.3x coverage) of the derivative chromosome 2 library and 19,968 clones (2.5x coverage) of the derivative chromosome 6 library were picked by robots and gridded onto filters for screening.

Primer pairs were designed to generate 1 product every 10Kb along the selected regions, resulting in a pool of 17 products covering the chromosome 2 sequence and 19 products covering the chromosome 6 sequence (Appendix A5). Use of these primer pools to screen both the derivative chromosome 2 and 6 libraries revealed 3 fosmid clones identified by both pools for the derivative chromosome 2 library; der2_16D21, der2_33M22 and der2_42P23 and 2 fosmid clones identified by both pools in the derivative chromosome 6 library; der6_9G19 and der6_35N15. All 5 fosmid clones were end sequenced to confirm that they were chimeric for the translocation (Figure 5.14).

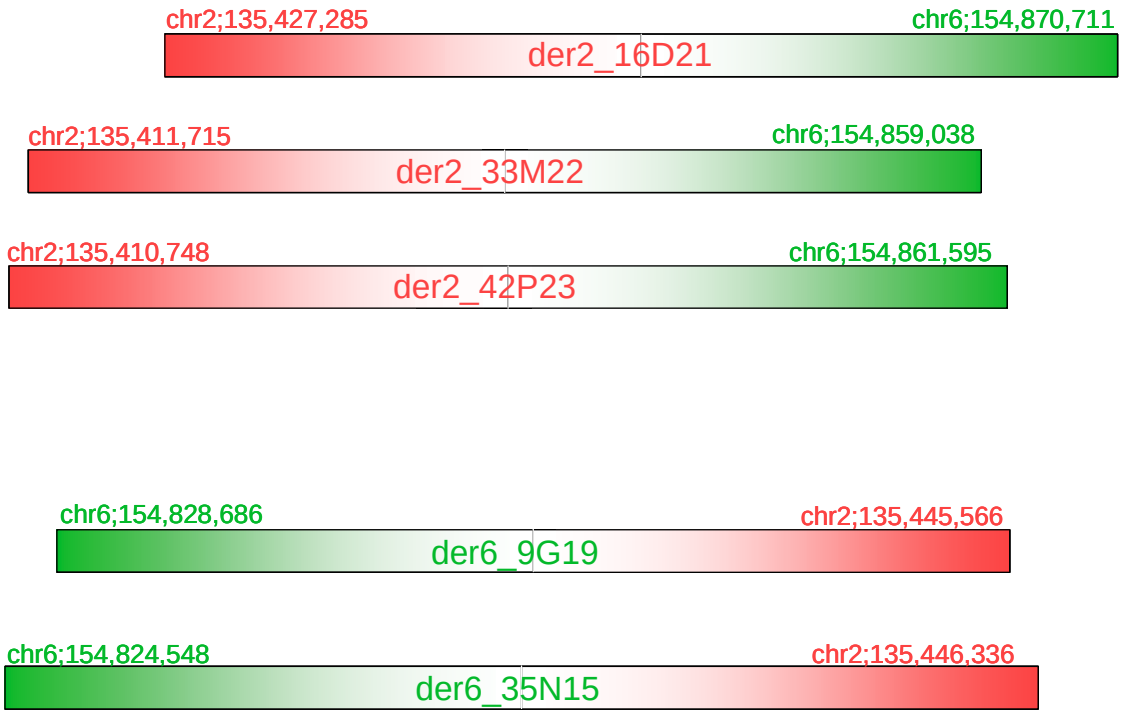

Figure 5.14 Alignment of chimeric fosmid clones by end sequence data.

On the basis of the end sequence data obtained from all 5 chimeric fosmid clones, der2_42P23 and der6_35N15 were selected for full sequencing as they appeared to contain the translocation junctions towards the middle of the clone DNA.

5.5.5 Identification of translocation breakpoints by sequence analysis

Analysis of the sequence from clone der2_42P23 showed that the insert DNA aligned to chromosome 2 from 135,410,748 to 135,432,727bp and chromosome 6 from 154,849,221 to 154,861,595bp and that clone der6_35N15 aligned from 154,824,548 to 154,849,219bp of chromosome 6 and 135,432,735 to 135,446,336bp of chromosome 2. The chromosome breakpoints were therefore deemed to be at 135,432,727bp on chromosome 2 and 154,849,219bp on chromosome 6.

Analysis of the sequence around the breakpoints revealed a 7bp deletion from chromosome 2 and a 3bp deletion from chromosome 6 (Figure 5.15).

```
t(2;6)(q21.1;q25.1)
chr2; GCTTTTGTTAGTTTCCCTGGAGCATCTATTTTTGTAATACATTTTGCCAGTGGGACTGAAAAGCT
der2; GCTTTTGTTAGTTTCCCTGGA          cagtgatttctcctgcctcggcctcctgggtagc
der6; cggctcatggcagcctccgcctcctggcTTTTTGTAATACATTTTGCCAGTGGGACTGAAAAGCT
chr6; cggctcatggcagcctccgcctcctggcttccagtgatttctcctgcctcggcctcctgggtagc
```

Figure 5.15 Alignment of sequence across the translocation junctions to the human genome reference sequence for patient t(2;6)(q21.1;q25.1). Deleted bases are shown in red.

Analysis of the full sequence generated from both der2_42P23 and der6_35N15 showed that the sequence aligned to the human genome reference sequence with 100% homology, indicating that there was no additional complexity.

5.6 Conclusions

Studies have shown that apparently balanced reciprocal translocations sometimes exhibit more complexity than originally observed through G-banding. If this additional complexity is around the translocation breakpoint, mapping by array painting and PCR may not be sufficient to fully resolve the translocation junctions. Conventional methods of breakpoint mapping such as FISH or generation of somatic cell hybrids followed by Southern blotting can be laborious, enhancing the need for a robust method of mapping these more complicated breakpoints. This chapter describes how the generation of a custom-made fosmid library and its screening using PCR product pools is a viable method for the identification of clones containing the translocation junctions. Once a spanning BAC clone has been identified by array painting or FISH, generation and screening of the custom-made library is sufficient for the isolation of the breakpoint sequence. Sequencing of the identified spanning fosmid clones generates approximately 36Kb of sequence around the translocation, providing information about the sequence surrounding the breakpoints. This analysis would

reveal any additional complexity around the translocation breakpoints which may have proven difficult to map by array painting and PCR.

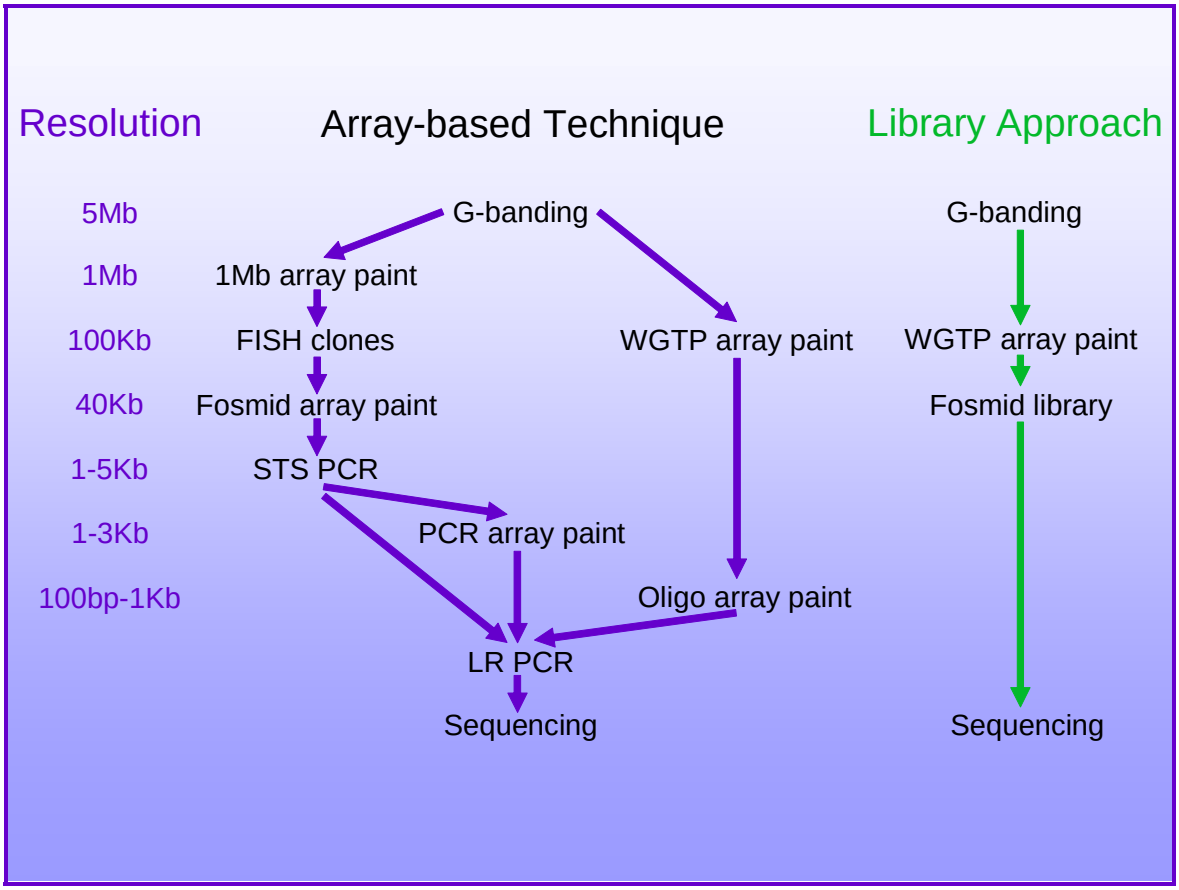

Figure 5.16 Comparison between array-based methodologies and custom-made libraries for breakpoint mapping.

A comparison between breakpoint mapping by array painting and PCR and custom-made fosmid libraries is shown in Figure 5.16. Use of custom-made fosmid libraries for breakpoint mapping does not rely on the assumption that a reciprocal translocation is balanced, and generates sequence data around the breakpoint which can reveal any cryptic rearrangements that may be present.