Chapter 4

Construction of a sequence-ready bacterial clone contig of 1pcen – 1p13

4.1 Introduction

Alterations of human chromosome 1 are among the most common form of chromosomal abnormality detected in human disease. Chromosomal aberrations localised to chromosome 1pcen-1p13 have been associated with human neoplasia such as acute megakaryoblastic leukaemia (Mercher *et al*., 2001), radiation induced meningioma (Zattara-Cannoni *et al*., 2001), colorectal Cancer (Nitta *et al*., 1987), as well as other diseases such as non-goitrous hypothyroidism (Dracopoli *et al*., 1986) and autosomal recessive ventricular tachycardia (Lahat *et al*., 2001).

Previously, two attempts have been made to construct physical maps within 1pcen-1p13 (Carrier *et al*., 1996, Brintnell *et al*., 1997). These maps, primarily consisting of YACs, were constructed for the purposes of characterising the genic environment of the nerve growth factor gene (NGFB) (Carrier *et al*., 1996) and for the elucidation of putative candidate genes within a smallest region of overlapping loss of heterozygosity (LOH) in breast cancer studies (Brintnell *et al*., 1997). Though the YACs used in the construction of these maps provided a means of generating coverage of large genomic regions with relatively few clones, the comparative difficulty of constructing the libraries and of analysing the cloned DNA (including shotgun sequencing) means that YACs as the primary resource are not well-suited for the construction of sequence ready maps. Instead, YACs have provided a means of linking bacterial clone contigs where genomic sequence is not represented within the bacterial clone libraries.

In contrast to YACs, bacterial clone libraries are easier to make and the cloned DNA is more easily manipulated. The use of bacterial clones for construction of a high resolution sequence ready map within 1pcen-1p13 would not only facilitate the identification of disease genes but would also provide the basis for a detailed characterisation of the genomic landscape of the interval. This chapter describes the construction of a sequence ready bacterial clone map in 1pcen-1p13 and compares previously published genetic, radiation hybrid and physical maps within the interval.

4.2 Construction of sequence-ready map of 1pcen – 1p13

A hierarchical strategy was used to construct the map as follows (see figure 4.1). At the start of the project, 3642 markers were publicly available which had been localised to the region by genetic or RH mapping. To provide additional markers for map construction a small insert library (SIL) derived from flow sorted chromosomes was constructed (see section 4.2.1). SIL clones were picked at random and sequenced to generate novel STSs. A subset of these novel STSs were placed in the region by RH mapping. The combined set of markers was used to initiate contig coverage by screening large-insert bacterial clones. Bacterial clone coverage was supplemented with contigs generated by McPherson *et al.,* (2000) from a whole genome fingerprint database. Contigs from this database were selected for inclusion in the 1pcen-1p13 map if they overlapped with existing clone contigs on the basis of shared fingerprints and/or marker content. The map was completed by iterative rounds of walking, using sequences at or near the end of each contig to re-screen the available libraries for clones.

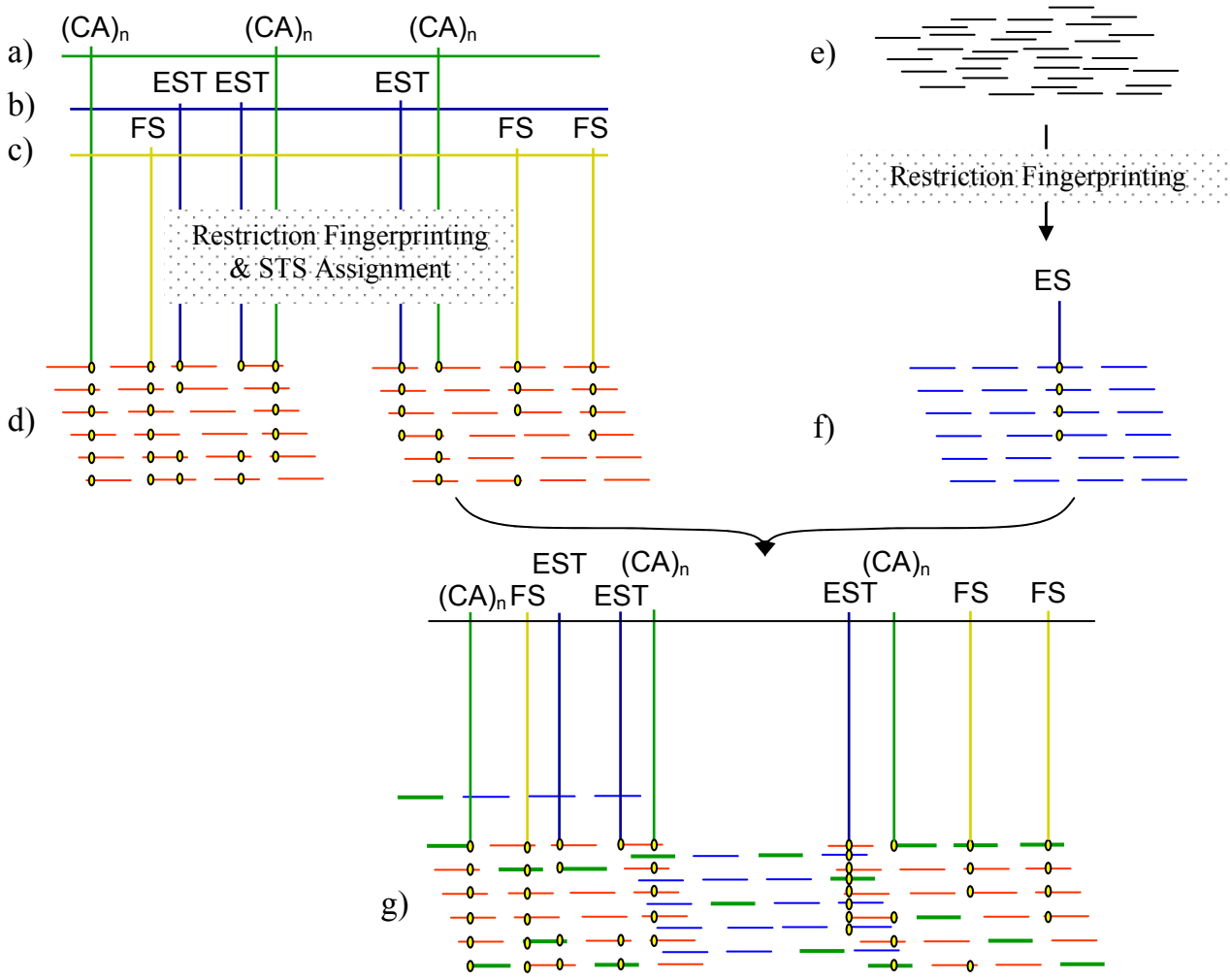

Figure 4.1: A representation of the two strategies used to construct a sequence-ready bacterial clone map of 1pc – 1p13. The hierarchical method utilises polymorphic genetic markers CA)n a), expressed sequence tags (EST) b) and markers generated by sequencing small insert libraries (FS) c) which are radiation hybrid mapped and screened across genomic libraries. Restriction fingerprinting and STS assignment is performed in parallel prior to data assimilation in FPC (orange lines, d). The whole genome fingerprinting approach utilises restriction digest fingerprinting of a 15-fold redundant genomic library e), including some marker data, to establish bacterial clone coverage f). Data from both these techniques is combined to generate contiguous map coverage g) proving a resource for the selection of a minimum tile path clones, bold green lines.

4.2.1 Small insert library construction

A bivariate flow karyotype of human DNA (figure 4.2a) was generated to facilitate purification of chromosome 1 DNA from other chromosomes (flow sorting was performed by Nigel Carter). Purified DNA was then completely digested with *Hin*d III prior to cloning into pBluescriptII vector and electroporated into *E. coli* XL1 blue electrocompetent cells (figure 4.2b). Two hundred test recombinants were picked from Xgal indicator plates, miniprepped (Birnboim *et al*., 1979) and DNA separated by electrophoresis on a 1% agarose gel as undigested and digested to ascertain average insert sizes (figure 4.2c). Of the two hundred test recombinants, 96% contained inserts and the average size was estimated to be ~5.6 kb. Successful generation of sequence data from a high percentage of 400 test SIL clones (data not shown) prompted a further 7296 SIL clones to be picked and sequenced (by others). A 36% failure rate (2611) at the prepping stage reduced the number of sequence templates to 4685.

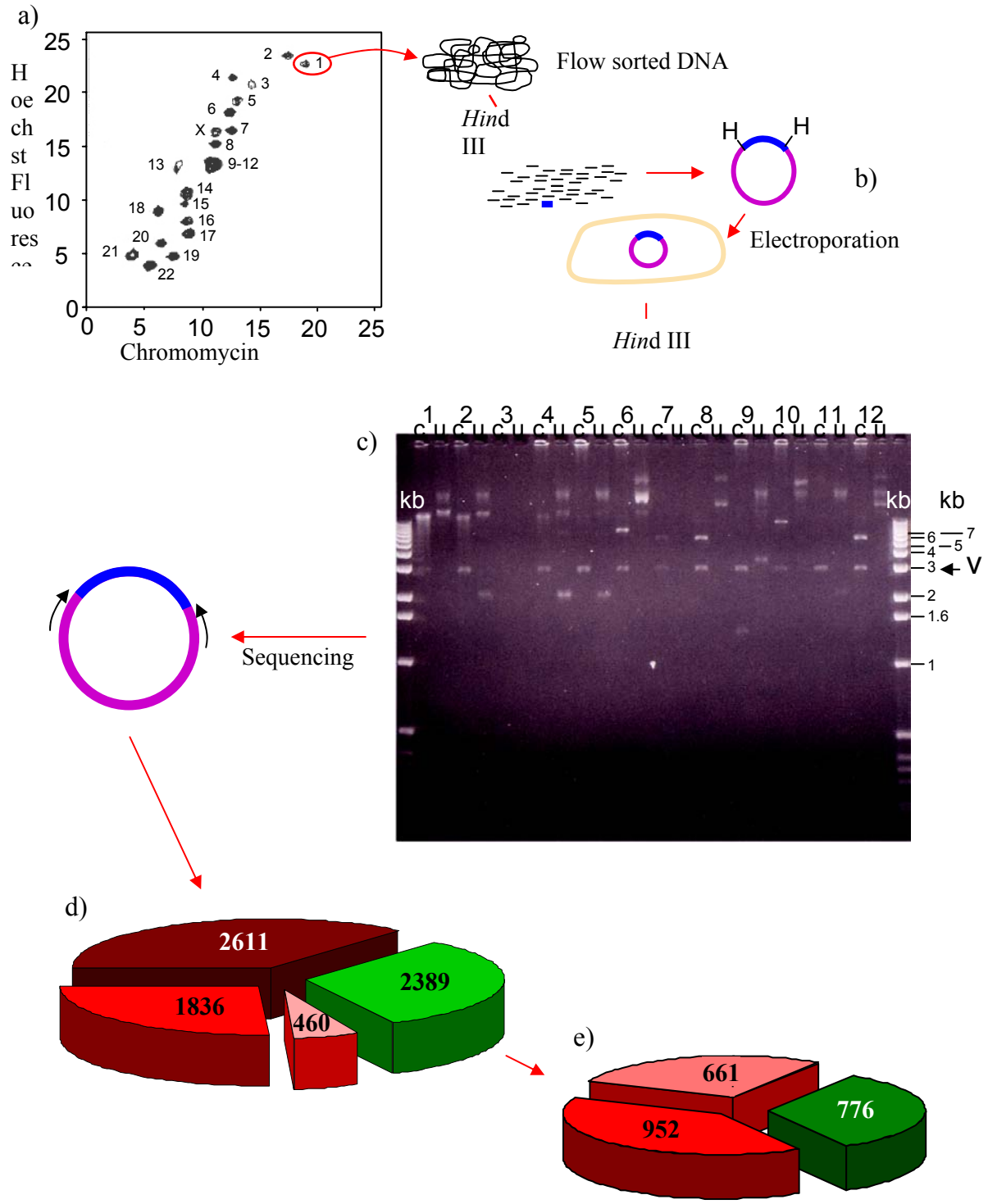
a)
Chromomycin
Flow sorted DNA
Hind III
b)
Electroporation
Hind III
c)
kb
V
Sequencing
d)
2611
2389
1836
460
e)
661
776
952

Figure 4.2: The construction of a chromosome 1 specific small insert library. a) A bivariate flow karyotype was generated to flow sort chromosome 1 DNA. b) DNA was completely digested with *Hin*dIII, cloned into pBluescriptII vector and electroporated into *E. coli* XL1 blue electrocompetent cells. c) Subclones were separated by electrophoresis cut (c) and uncut (u) on an agarose gel and insert sizes determined by use of 1 kb ladder (kb). The vector band (v) was used to correlate the insert size with the undigested band. d) Prior to primer design there were small insert library clone failures at prepping (dark red), sequencing (including sequence of < 80 bp (red)) and STS design stages (pink). Further failures e) were due to repeats contained within either primer (dark red) or during experimentation (pink). The total number of RH mapped flow sorted markers with unambiguous placement on the physical map is represented in figure 4.2e) (dark green).

Figure 4.3 illustrates the size distribution of the 4685 SIL clones that were sequenced. Of these clones, 49% had insert sequences that were inappropriate for primer design. This set comprised 8% that failed to produce sequence; 31% with sequences <80 bp; and 10% with sequence that did not satisfy primer design parameters (e.g. unequal or suboptimal GC content preventing primer design, PCR product size shorter than an acceptable minimum length). The remaining 2389 (51%) small insert library SIL clones sequences (31% of the original number of picked SIL clones) produced sequence from which primer pairs were could be designed (figure 4.2d).

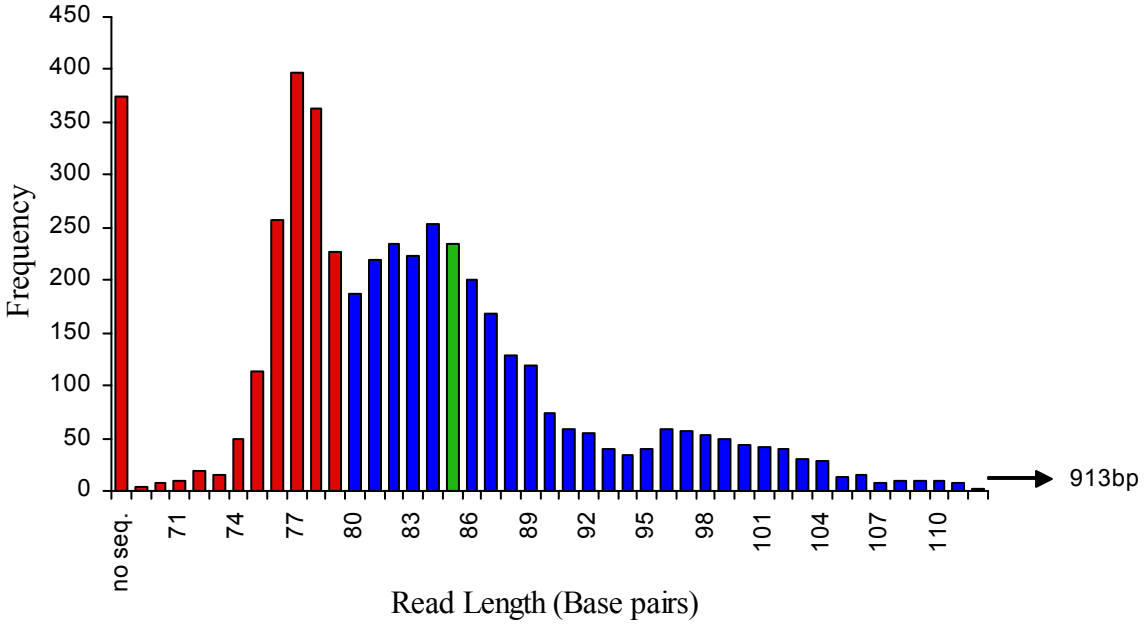

Figure 4.3: Sequence length and frequency of SILs passing STS design stage. Sequences too short for primer design (red), statistically significant proportion (91%) of STSs that generated sequence that were suitable for primer design (blue) and the mean read length, 85 base pairs (green).

4.2.2 Radiation Hybrid Mapping of SIL markers

All 2389 primer pairs were synthesised and tested by RH mapping (by others). Of these, 952 were rejected as at least one of a primer pair wholly or partly contained repeat sequence and a further 439 were experimental failures. These failures included mapping of a marker to multiple chromosomal locations, failure to map to a human chromosome and poor specificity (indicated by production of multiple PCR products generated when different annealing temperatures were used during PCR). The 776 flow sorted STS markers that were successfully

radiation hybrid mapped and placed uniquely on the physical map represent 11% of the flow sorted small insert library clones originally picked for sequencing. As a result of this work, the total number of unique markers in the chromosome 1 RH map was increased from 3642 to 4418.

4.2.3 Bacterial clone contig construction

The combined set of 204 markers localised to the 1pcen-1p13 interval by RH mapping were used to isolate large-insert bacterial clones for contig construction. Initially, each PCR assay was tested by amplification of genomic DNA and DNA from a chromosome 1p hybrid in order to establish optimal conditions for specificity (figure 4.4a). PCR products from the assays were then radiolabelled by PCR and hybridised in pools of 20 – 30 to genomic arrays of PAC clones (fig 4.4b). Positive PAC clones were picked into microtitre plates and re-arrayed on a bacterial clone polygrid filter. Probes were then hybridised individually to the polygrid filter. A total of 110 publicly available and flow-sorted markers were screened in pools across RPCI 1, 3 – 5 PAC libraries identifying 878 PAC clones (the remaining 86 available markers were PAC screened by others as part of the chromosome 1 mapping project). Hybridisation data was entered into a chromosome specific ACeDB database (Durbin and Thierry-Meig 1994), 1ace. In parallel, positive bacterial clones identified by the pool hybridisation were subjected to restriction digest fingerprinting (Gregory *et al*., 1996) (figure 4.4c) and marker hybridisation data from 1ace was assimilated within FPC (Soderlund *et al*., 1997) following contig assembly. Minimum tile path clones representing bacterial clone coverage of island contigs generated by marker hybridisation, were selected for sequencing.

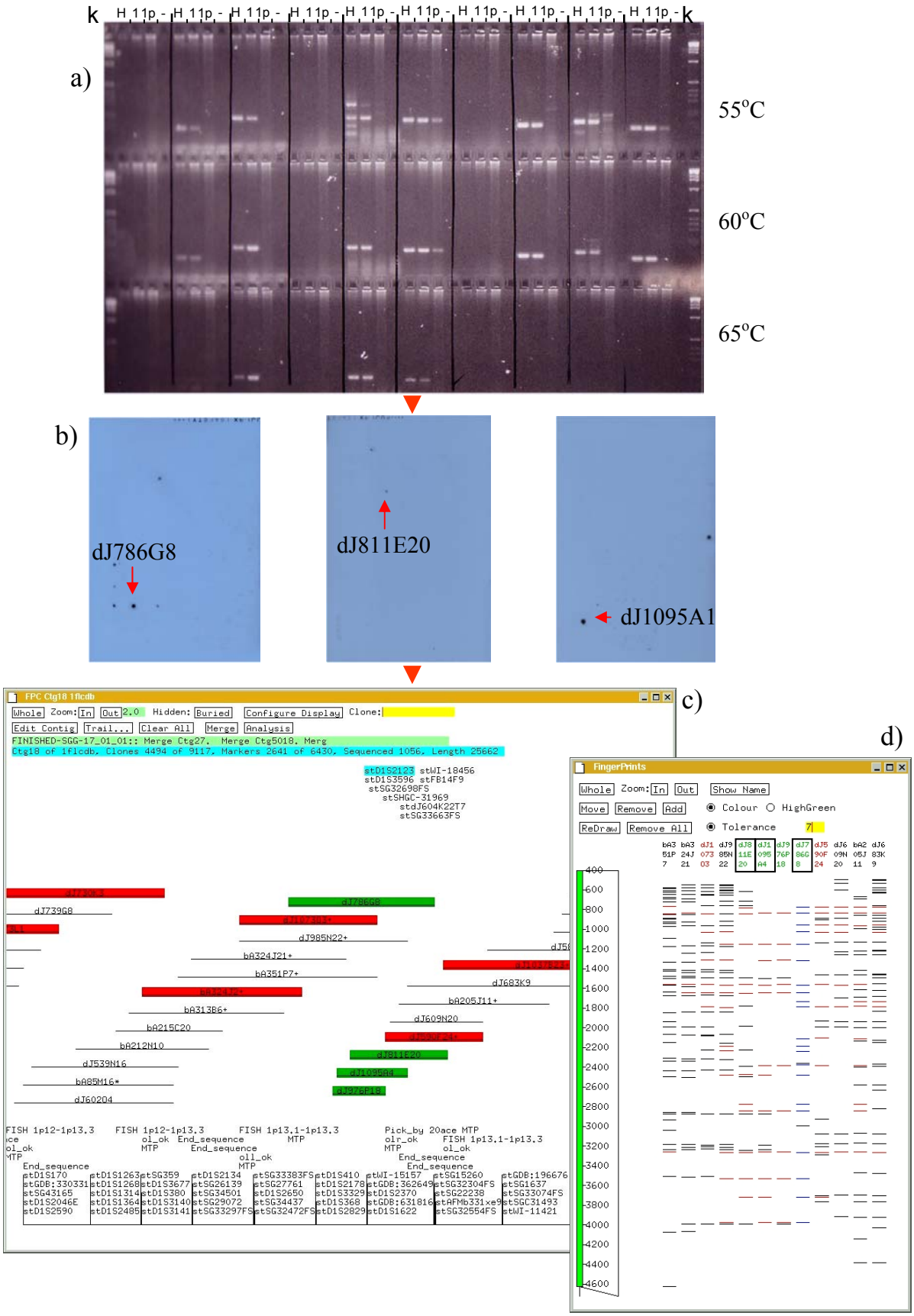
k H 11p - H 11p - H 11p - H 11p - H 11p - H 11p - H 11p - H 11p - H 11p - k
a)
55°C
60°C
65°C
b)
dJ786G8
dJ811E20
dJ1095A1
c)
d)
FPC Ctg18 1flcdb
Whole Zoom: In Out 2.0 Hidden: Buried Configure Display Clone:
Edit Contig Trail... Clear All Merge Analysis
FINISHED-SGG-17_01_01:: Merge Ctg27. Merge Ctg5018. Merg
Ctg18 of 1flcdb. Clones 4494 of 9117. Markers 2641 of 6430. Sequenced 1056. Length 25662
stD1S2123 stWI-18456
stD1S3596 stFB14F9
stSG32698FS
stSHGC-31969
stdJ604K22T7
stSG33663FS
dJ739G8
dJ786G8
dJ985N22+
bA324J21+
bA351P7+
dJ683K9
bA205J11+
bA313B6+
dJ609N20
bA215C20
bA212N10
dJ539N16
bA85M16*
dJ602O4
dJ811E20
dJ1095B4
FISH 1p12-1p13.3
FISH 1p12-1p13.3
FISH 1p13.1-1p13.3
Pick_by 20ace MTP
FISH 1p13.1-1p13.3
End_sequence
stD1S170 stD1S1263 stSG359 stD1S2134 stSG33383FS stD1S410 stWI-15157 stSG15260 stGDB:196676
stGDB:330331 stD1S1268 stD3677 stSG26139 stSG27761 stD1S2178 stGDB:362649 stSG32304FS stSG1637
stSG43165 stD1S1314 stD1S380 stSG34501 stD1S2650 stD1S3329 stD1S2370 stSG22238 stSG33074FS
stD1S2046E stD1S1364 stD1S3140 stSG29072 stSG34437 stD1S368 stGDB:631816 stAFMb331xe9 stSGC31493
stD1S2590 stD1S2485 stD1S3141 stSG33297FS stSG32472FS stD1S2829 stD1S1622 stSG32554FS stWI-11421
FingerPrints
Whole Zoom: In Out Show Name
Move Remove Add Colour HighGreen
ReDraw Remove All Tolerance 7
bA3 51P 7 bA3 24J 21 dJ1 073 03 dJ9 85N 22 dJ8 11E 20 dJ1 095 A4 dJ9 76P 18 dJ7 86G 8 dJ5 90F 24 dJ6 09N 20 bA2 05J 11 dJ6 83K 9
400 600 800 1000 1200 1400 1600 1800 2000 2200 2400 2600 2800 3000 3200 3400 3600 3800 4000 4200 4400 4600

Figure 4.4: Generation of sequence ready bacterial coverage using the hierarchal strategy. a) An agarose gel showing PCR products generated at annealing temperatures of 55^{o}C, 60^{o}C and 65^{o}C. The PCR reaction used total human DNA (H), chromosome 1 (1), chromosome 1p (1p) and a negative control (-) as template. b) Autoradiographs of an STS pool hybridisation to genomic PAC library filters, positives are indicated by red arrows. c) PACs identified by hybridisation were restriction digest fingerprinting, assembled into contigs and assimilated with STS data within FPC. d) The digitised fingerprints (black boxes) of 3 clones identified by pool screening and assembled within FPC.

Bacterial clone coverage of 1pcen-1p13 within the whole genome (WG) fingerprint database (McPherson *et al*., 2001) was identified either by using the chromosome 1 markers associated with assembled contigs (where they were available), or by adding a representative set of experimental or virtual *Hin*d III fingerprints of PAC clones isolated by the hierarchical strategy (Figure 4.5a). Placement of PACs with associated RH mapped STS data within the stringently assembled WG fingerprint database facilitated the localisation and joining of WG BAC contigs. A three-fold redundant tiling path of BAC clones was picked from the WG database (Figure 4.5b) and incorporated into the chromosome 1 specific fluorescent fingerprint database (as part of this thesis and the chromosome 1 project). Once WG BACs were assimilated into the 1pcen-1p13 PAC contigs a more optimal series of BAC and PAC clones were chosen (with more minimal overlaps) as the sequencing tile path (Figure 4.5c). The use of both hierarchical and WG fingerprint approaches resulted in the 1pcen-1p13 interval being covered by 2 bacterial clone contigs. The two bacterial clone contigs were fortuitously linked by 2.3 kb of unfinished sequence from BAC clone bA722J12, produced by

RIKEN Genomic Sciences Center as part of their chromosome 18 sequencing project. The clone appears to have been erroneously chosen by RIKEN as the unfinished sequence is placed uniquely on chromosome 1 by high BLAST score alignment.

Figure 4.5. Can be viewed in a separate PDF

Figure 4.5: The assimilation of PAC contigs into whole genome and chromosome specific fingerprint databases. a) PAC clones representing bacterial clone coverage from the hierarchical mapping technique were added to the WG fingerprint database by either experimental or virtual restriction digest fingerprinting (light blue clones). b) a three fold redundant set of BACs were selected from the WG fingerprint database and incorporated into the chromosome 1 specific database by fluorescent fingerprinting. c) Minimum tile path clones are boxed in red, alignment between the 3 maps is shown by selected clones (circled).

The final 13 Mb bacterial clone contig covering human 1pcen-1p13.2 (figure 4.6a) includes 1130 bacterial clones (480 BACs, 648 PACs and 2 cosmids) and contains 250 markers by hybridisation. Two hundred and four of these markers (157 publicly available and 47 flow sorted) have associated RH mapping data. The final map contained an additional 46 markers without RH information; but these markers were added to the contig map by hybridisation. All markers have a unique chromosomal placement, i.e. either all or the majority of bacterial clones identified by hybridisation have fingerprints that assemble within the 1pcen-1p13.2 contig and to no other location the chromosome. A minimum tile path of 136 minimum clones (figure 4.6b), including 2 previously sequenced cosmids (AC000031 and AC000032), was selected for sequencing (by others) within the Sanger Institute. Figure 4.6c represents an FPC display of a 3 Mb region of the 13 Mb contig which illustrates the extensive coverage of independent map information which anchors clones across the 1pcen-1p13.2 contig. The 3 Mb interval is defined by two framework markers, D1S221 and D1S2746 (markers are denoted by yellow boxes, figure 4.6c). Also shown are markers with RH mapping data (light blue boxes) that were used to identify large insert bacterial clones by hybridisation to genomic library filters.

Figure 4.6 can be viewed in a separate PDF

Figure 4.6: A representation of PAC and BAC contig coverage of 1pcen – 1p13. a) An ideogram of the region of bacterial clone contig coverage generated by this study as represented at 850 band resolution of human chromosome 1. b) An AceDB display of the minimum tile path of 136 clones from the 13 Mb contig. Represented are the sequencing statuses of minimum tile path clones (as of 14th October 2002), pre-shotgun (clear boxes), pre-finished (green boxes), finished sequence (red boxes) and submitted sequence (black boxes). c) a 3 Mb region of the final contig depicted in FPC reflecting the density of framework markers (yellow boxes), RH mapped markers (light blue boxes) and non-RH mapped markers (clear) positioned by hybridisation within the bacterial clone contig. Minimum tile path clones are highlighted in red.

4.3 Evaluation of SIL marker distribution in chromosome 1

The chromosomal distribution of 776 RH mapped STSs, originating from the flow sorted small insert library, was evaluated based on the fingerprint map (figure 4.4). The histogram in figure 4.7b depicts the number of STSs in 5 Mb intervals along the length of chromosome 1. The interval size was calculated directly from the fingerprint map using 4 bands / kb (see section 2.22.2). The average number of markers (14.7 / 5 Mb) varies appreciably next to the centromere and telomeres. This may be due to difficulties in cloning bias encountered when cloning chromosomal centromeric and telomeric repeat sequences (Doggett *et al*., 1995). The overrepresentation of flow-sorted markers in some intervals may relate to sequence content. For example, a region of genomic sequence with a particular base composition (e.g. AT-rich)

may be particularly amenable to digestion with *Hin*d III, and thus yield a high fraction of short (easily cloned) *Hin*d III region compared to a GC-rich region. This bias would be reflected in the representation of inserts in the SIL, as the SIL was prepared from HindIII-digested DNA.

The distribution of the 776 flow sorted markers in the fingerprint-based map was compared to the radiation hybrid map of chromosome 1 (figure 4.7c). There is very good agreement between the two maps, with 12 STSs (1.5%) showing incongruent placement (blue dots). These discrepancies may result from low copy repeats (duplicates not being identified within radiation hybrid vectors) or by experimental error where the wrong STS has been hybridised to the gridded arrays of bacterial clones. The density of markers that have associated RH mapping data and unique chromosome hybridisation data within 1pcen-1p13, was increased from 12 / Mb to 17 / Mb upon addition of flow-sorted markers.

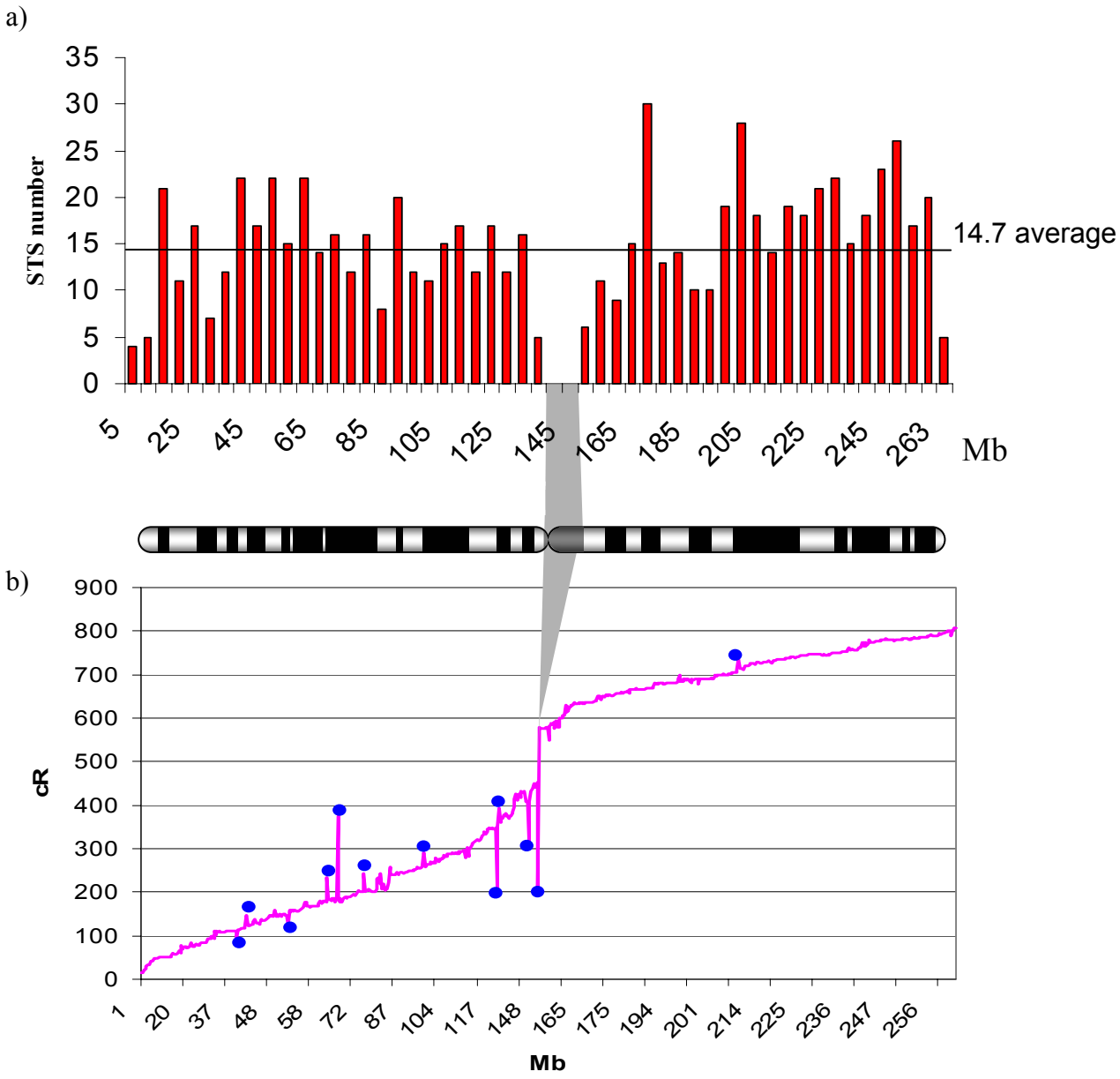

Figure 4.7: Chromosomal distribution of flow sorted markers and comparison to of radiation hybrid and physical maps. a) The chromosomal distribution of 776 RH mapped STSs originating from the flow sorted small insert library along the length of the chromosome in 5 Mb map intervals. b) A comparison of the radiation hybrid map of 778 flow sorted markers and their localisation to bacterial clone contigs by hybridisation, outliers are denoted by blue dots.

4.4 Comparison of Published Maps:

4.4.1 Physical maps

Two physical maps, consisting primarily of YACs, have previously been published within 1pcen-1p13 (Carrier *et al*., 1996, Brintnell *et al*., 1997). Carrier *et al.,* (1996) used genetic markers and STSs designed from genes localised to 1p13 to screen CEPH YAC libraries. Overlaps between the eight non-chimeric YACs used to build the 3 Mb contig around the NGFB locus were determined by long-range restriction mapping and by STS content data. Brintnell *et al.,* (1997) used a similar approach to generate YAC coverage (from CEPH and ICI YAC libraries) across an interval of 1p13.1 but also included CIT library BACs within the contig. Though the two contigs overlap, based on their shared content of marker D1S252, the Brintnell *et al.,* (1997) map extends distally by 1.7 Mb. Both maps, with the exception of one marker, were in broad agreement with the physical map described here (within the limits of localising a marker or gene by hybridisation to a YAC or YAC restriction fragments). One conflicting marker, D1S3347, was positioned within 1p13 by Brintnell *et al.,* (1997) but has subsequently been localised to 1p35.3 by RH mapping. BLAST analysis of D1S3347 places the marker in 1p35.1 and 1p12 via a high BLAST score and 95% sequence alignment (the parent sequence contains ambiguous bases, thus the <100% BLAST score). D1S3347 (synonymous with WI-8708) was derived from an EST. Subsequent alignment of the EST to genomic sequence (by Ensembl analysis) indicates the marker is derived from the 3' UTR of an mRNA isolated from small cell carcinoma lung tissue by the IMAGE consortium. The three exon gene is localised to 1p35.1 whilst a processed mRNA, inserted into the genomic

sequence in the reverse orientation, is contained within 1p12 and thus explains the discrepant placement of the marker in the Brintnell *et al.,* (1997) study.

4.4.2 Genetic map

The generation of a bacterial clone contig within 1pcen-1p13 permitted a comparison to be made with the genetic map of chromosome 1 (Dib *et al*., 1996). The order and distance between 27 genetic markers placed uniquely in the bacterial clone contig by hybridisation was determined. There was very good agreement between the two maps in relation to marker order (figure 4.8), with the higher resolution bacterial clone map facilitating the separation of markers that had previously been placed within the same genetic interval. Two genetic markers showed discrepant placement; the first, D1S2852 (152.2 cM, 8.2 Mb) localises to its current position in the physical map by hybridisation and sequence alignment; the second, D1S418 (152.2 cM, 11.3 Mb), was positioned by weak hybridisation to two PAC clones within the contig. BLAST analysis of the sequence from which D1S418 was derived (which contains ambiguous bases) localises the marker within PAC clone dJ671G15 placing it in the correct physical map location. Incorrect placement of these markers within the Généthon genetic map may be attributed to the level of resolution provided by the number of meioses used to construct the map.

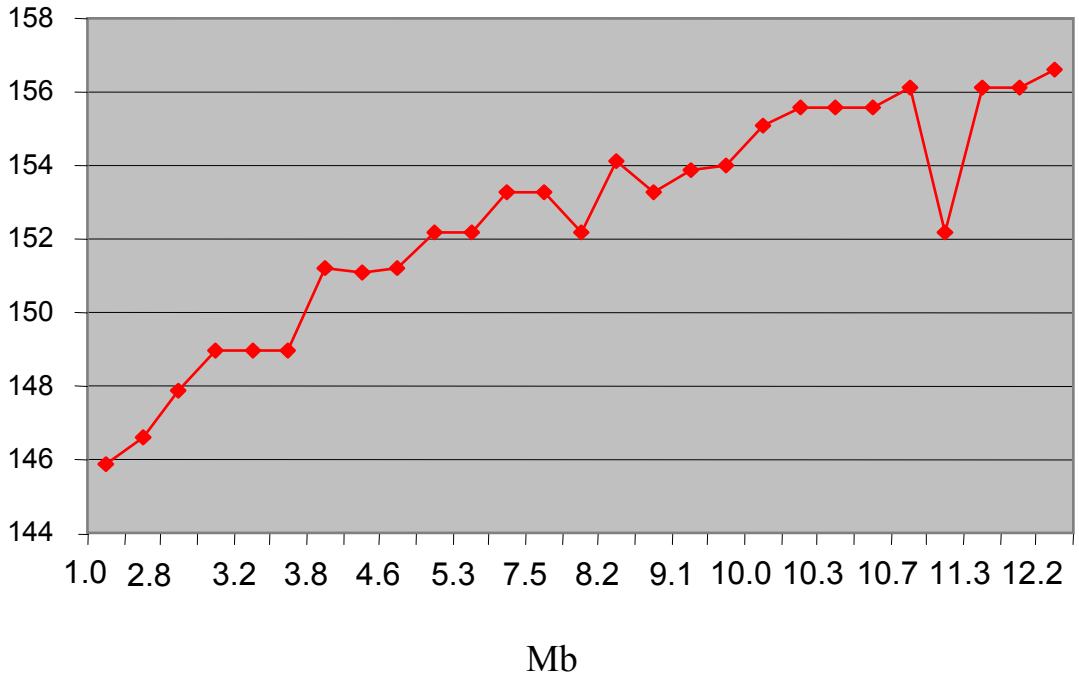

Figure 4.8: The distribution of genetic mapped markers positioned within the 1pc – 1p13 contig by hybridisation. The physical map is plotted on the X-axis in megabase (Mb) and the Y-axis in centimorgans (cM).

4.4.3 Radiation hybrid map

To investigate the order and resolution of radiation hybrid markers on the physical map, markers with unique hybridisation positions on the fingerprint map were plotted against their RH map locations. Figure 4.9 shows a good correlation for the most part between physical and RH maps but indicates a number of markers that appear to be poorly resolved on the RH map. Twenty seven markers show discrepant placement of >15 cR from their physical map position, the resolution of the chromosome 1 RH map (Panos Deloukas personal communication). One marker was subsequently placed correctly by BLAST analysis and a

second marker incorrectly placed because one of the primer pairs was designed within a repeat. The remaining 25 are placed on the physical map by unambiguous hybridisation and or electronic PCR (ePCR) (Schuler *et al*., 1997).

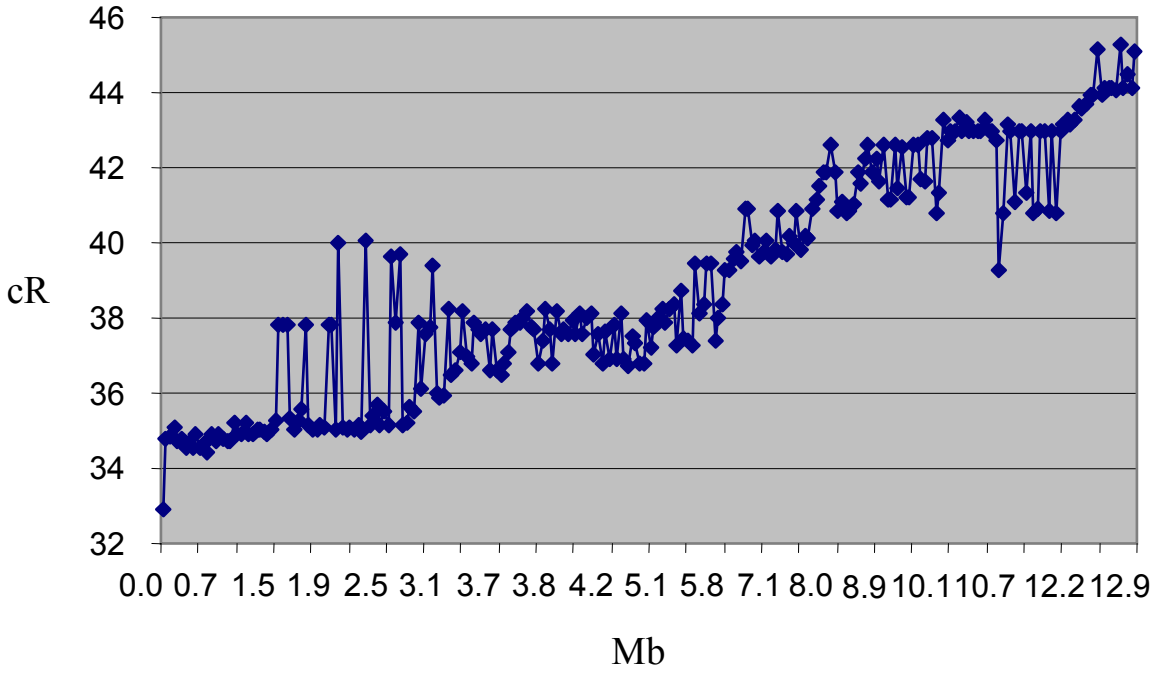

Figure 4.9: The distribution of radiation hybrid mapped markers positioned within the 1pc – 1p13 contig by hybridisation. The physical map is plotted on the X-axis in megabase (Mb) and the Y-axis in centi-ray (cR).

4.4.4 A comparison of three maps

The comparative accuracy of the genetic and radiation hybrid maps can be evaluated by plotting a set of markers contained within all three maps from the analyses above (figure 4.10). The genetic map shows good agreement with the physical map (discrepant markers being accounted for above) with markers placed within the same genetic interval being

resolved. The RH map provides higher resolution and is also mostly in good agreement with the other map and provides both more markers and a higher resolution than the genetic map. However it also shows more discrepancies in marker order (see chapter 7).

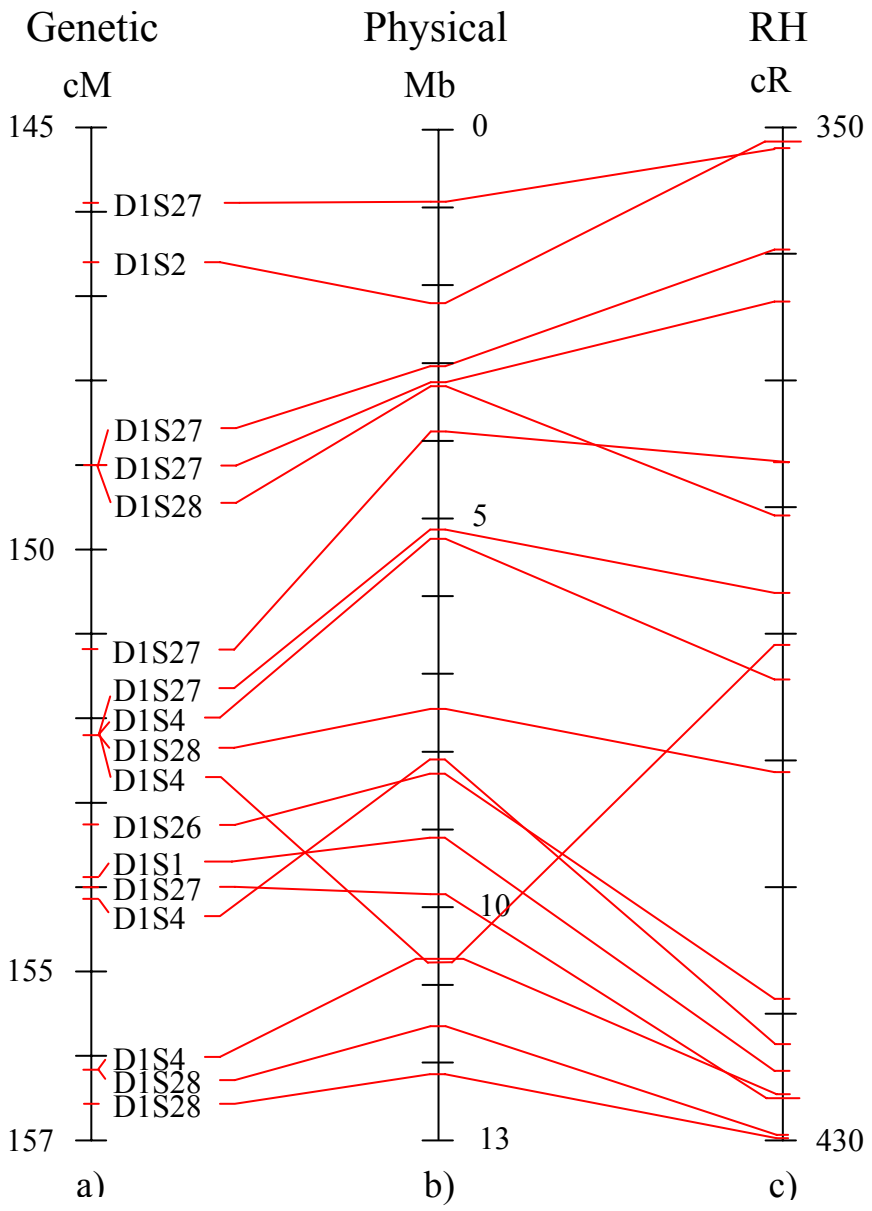

Figure 4.10: A comparison of marker distribution between genetic a), physical b) and radiation hybrid c) maps of 1pcen – 1p13. Only markers that were present on all three maps have been represented.

4.5 Discussion

Two different strategies were used to construct a 13 Mb bacterial clone contig of human chromosome 1pcen-1p13. Bacterial clone coverage was initiated by utilising publicly available markers, localised to 1pcen - 1p13 by RH mapping, to screen large insert PAC libraries. The non-random distribution of these markers (a large proportion of which had been derived for positional cloning projects or from ESTs, thus enriching marker density in disease regions or gene-rich regions, respectively) required the generation of flow sorted STSs. A total of 776 markers derived from a SIL, were successfully placed on the chromosome 1 specific radiation hybrid map and 47 of them mapped within the region 1pcen-1p13 markers. This supplemented the publicly available markers in the region (157) and increased the density of markers from 12 per Mb to 17 per Mb, above the target density of 15 per Mb (Olson *et al*., 1993, Bentley *et al*., 2000). PACs identified by genomic library screening of all markers within the interval were fingerprinted and assembled into contigs. Selected PACs from these island contigs were incorporated within the whole genome BAC fingerprint database by *Hin*d III fingerprinting and overlap analysis. Clones from this database were retrieved and incorporated into the local study. At the end of the project, a three-fold redundant tile path of 480 BAC clones, from the whole genome fingerprint database, in addition to the 648 PACs, identified by library screening, constitute the 13 Mb contig.

The generation of contiguous physical map coverage has permitted a high resolution comparison to be made of genetic and radiation hybrid maps within the interval. The algorithms used to construct the RH map assume random distribution of breaks along the

length of the chromosome. Experimental data (Deloukas *et al*., 1998) has shown that the high retention rate at the centromere, coupled with variation of DNA fragment sizes (in comparison to the rest of the chromosome), results in an overestimation of the cR distances between markers adjacent to the centromere. These anomalies explain why there are a relatively high percentage of markers that show discrepancies between the two maps and why the usual size estimate of 1 cR_{3000} to 250 kb (Deloukas *et al*., 1998) does not hold within 1pcen-1p13. According to the RH map the 105 cR interval should be 26 Mb where as it has been shown to be ~13 Mb.

The relative uniformity of the physical map metric facilitated the resolution of markers that had previously been localised to the same interval on the genetic map. Apparent differences in recombination rate across the interval, as shown by the step-wise comparison of genetic and physical maps in figure 4.8, cannot be determined from the resolution of the bacterial clone contig and would have to be resolved by placement of the markers within contiguous genomic sequence. The physical map is not a uniform metric as it relies upon the existence of *Hin*d III sites for the generation and overlapping of fingerprint bands; a clone with a large number of bands may contain the same amount of genomic sequence as a shorter clone with fewer bands.

It has been proven that there is a large variation in the rate of recombination along the length of a chromosome (IHGSC, 2001). Recombination rates at telomeres are greater than within chromosome arms which are in turn greater than regions adjacent to the centromere.

Therefore, it was not unexpected that a shorter genetic distance, 8.7cM defines a larger physical distance, 13 Mb, within 1pcen-1p13.

The bacterial clone map has provided a means by which a more accurate estimate of the physical size of the interval can be made, overcoming inherent levels of lower resolution within genetic and radiation hybrid mapping. High resolution sequence analysis of the 136 minimally overlapping bacterial clones selected for sequencing from the contig constructed here forms the basis of the next chapter.