6 Results: Investigation of viral integration using microarrays and custom-made libraries

6.1 Introduction

The human herpes virus is ubiquitous in the human population and normally exists as an extra chromosomal element in the host. HHV-6 is a lymphotropic virus which has been identified as a causative agent in the illness exanthum subitum (also known as Roseola infantum) whose patients present with rash, fever and febrile convulsions. Restriction analysis and FISH studies have shown that in approximately 3% of the British population the HHV-6 virus has integrated into the human genome (Luppi et al. 1993; Torelli et al. 1995; Morris et al. 1999).

This chapter demonstrates how the techniques developed to map chromosome rearrangement breakpoints can be applied to genomic rearrangements such as the integration of viral genomes into the human genome. FISH onto metaphase chromosomes was used to identify the chromosomes with viral integration, microarrays were used to assess the viral copy number within the human genome, and a custom-made library was used to investigate the composition of the integrated virus. These techniques, along with FISH onto extended chromatin fibres have generated data that may help to elucidate the mechanism by which the HHV-6 virus integrates into the human genome.

6.1.1 The HHV-6 virus structure

The HHV-6 virus was extensively mapped by restriction enzyme digestion using a combination of 5 different enzymes (Figure 6.1) (Lindquester et al. 1996). The virus consists of a central unique segment of approximately 144.5Kb in length, flanked by 2 direct repeats either side (DR_L and DR_R). Whole genomic DNA isolated from the virus and cloned fragments covering approximately 103.9Kb (72%) of the unique segment were obtained for use in this study from Duncan Clark at University College London and Scott Schmid at CDC, Atlanta.

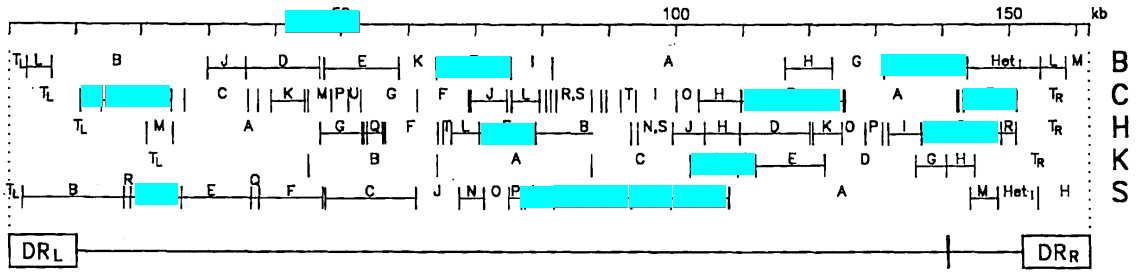


Figure 6.1 Restriction digest map of the HHV-6B genome using enzymes B; BamHI, C; ClaI, H; HindIII, K; KpnI and S; SalI (taken from Lindquester et al. 1996). Cloned fragments (blue) were obtained for use in this study.

6.1.2 Patient details

Lymphoblastoid cell-lines were obtained from 3 patients (AMD, AS and KK) suspected of containing integrated HHV-6 after HHV-6 PCR testing.

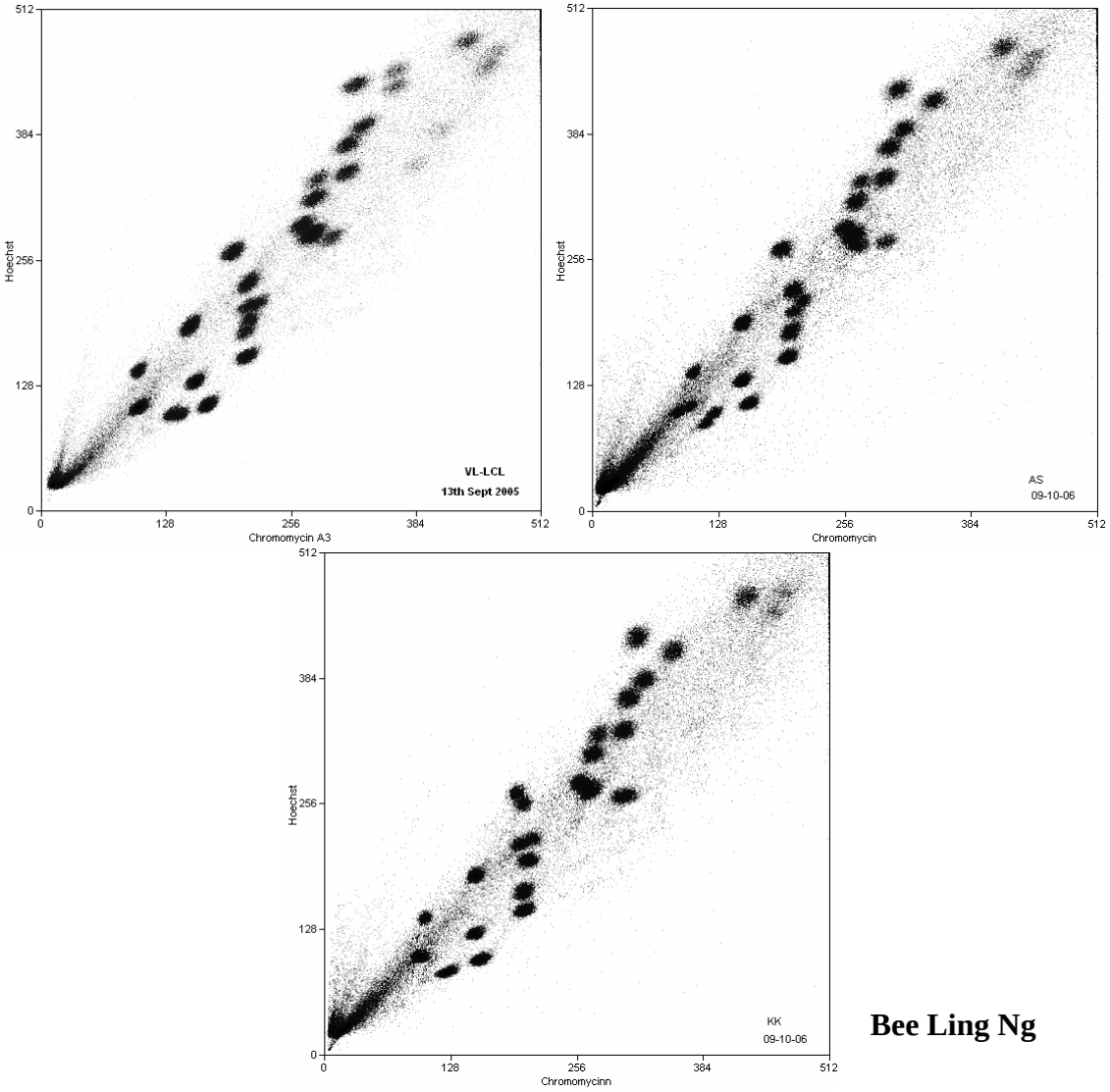
Patient AMD was a healthy 35 year old male with no known past medical history. HHV-6 specific quantitative competitive PCR had established that a high viral load equating to approximately 5 copies per cell was retained by the patient over a 10 month period (Clark et al. 1996). PCR primers used to amplify each of the 87 predicted open reading frames of the virus showed that all were present.

Patient AS was a 24 year old male who had previously developed chronic renal failure and received a renal transplant at the age of 18 years. Following the transplant, the patient presented with lethargy and severe weight loss. PCR tests for HHV-6 showed a raised level of $12x10^6$ copies/ml, with a subsequent sample 8 months later showing a persistently high level of $8x10^6$ copies/ml.

Patient KK was a 25 year old male who developed chronic renal failure and received a cadaveric renal transplant aged 19 years. Following a bout of lethargy, HHV-6 PCR testing revealed a viral load of $8.6x10^6$ copies/ml and a follow-up sample 2.5 years later showed a level of $12x10^6$ copies/ml.

6.1.2.1 Flow karyograms

The flow karyograms were assessed for all 3 patients (Figure 6.2), however in all cases there was no notable change in the profile when compared to a karyogram from a normal individual, therefore no indication was given of which chromosome contained the integrated viral DNA.

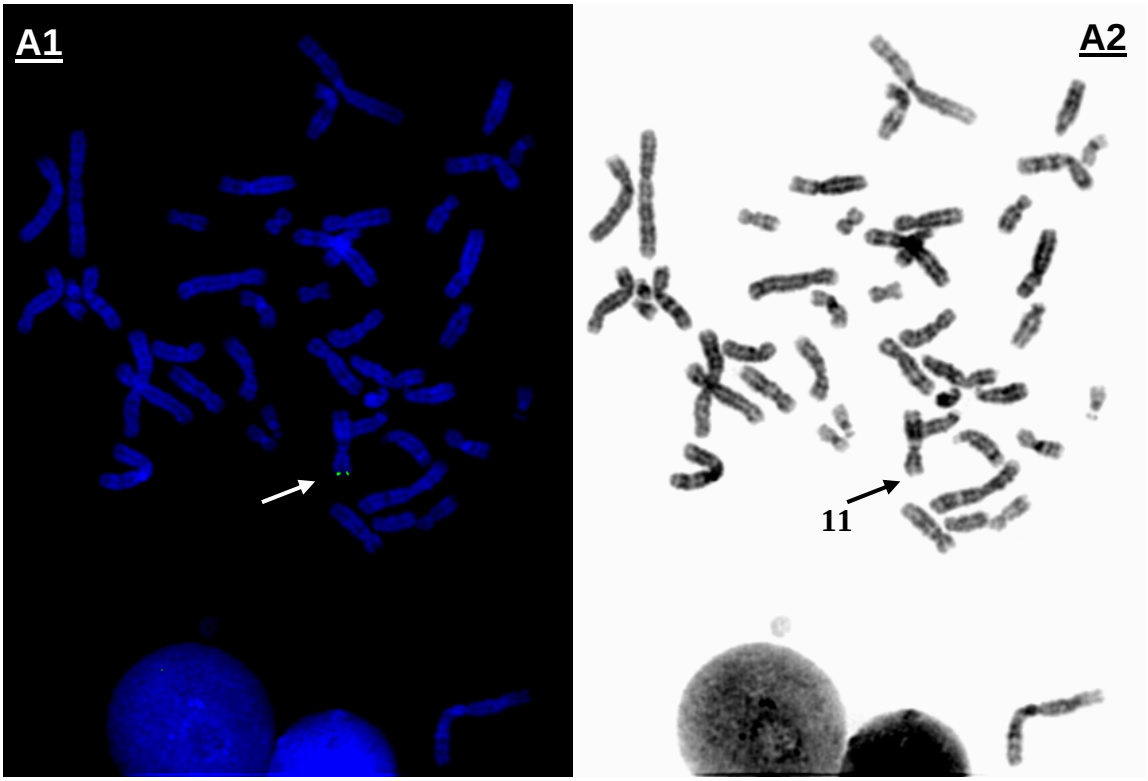


Bee Ling Ng

Figure 6.2 Flow karyograms for patients AMD, AS and KK. In all 3 patients, the integration of the virus had a negligible effect on the karyogram and therefore the chromosome carrying the integration was unable to be determined by this method.

6.2 FISH analysis of viral integration sites in 3 patients

A FISH probe created from viral genomic DNA was hybridised onto metaphase chromosome spreads isolated from the patients' cell-lines. Analysis of the signal revealed that in each patient the virus had integrated into one homologue of a particular chromosome only. In patient AMD the virus had integrated into 11p15.2-11p15.5, in patient AS it had integrated into 9q34.12-9q34.3 and in patient KK it had integrated into 17p13.1-17p13.3 (Figure 6.3). In all cases, the integration sites were seen to be within or close to the telomeres of human chromosomes consistent with published data where integration had been observed at 1q44, 22q13 and 17p13.3 (Daibata et al. 1998; Daibata et al. 1999; Morris et al. 1999).



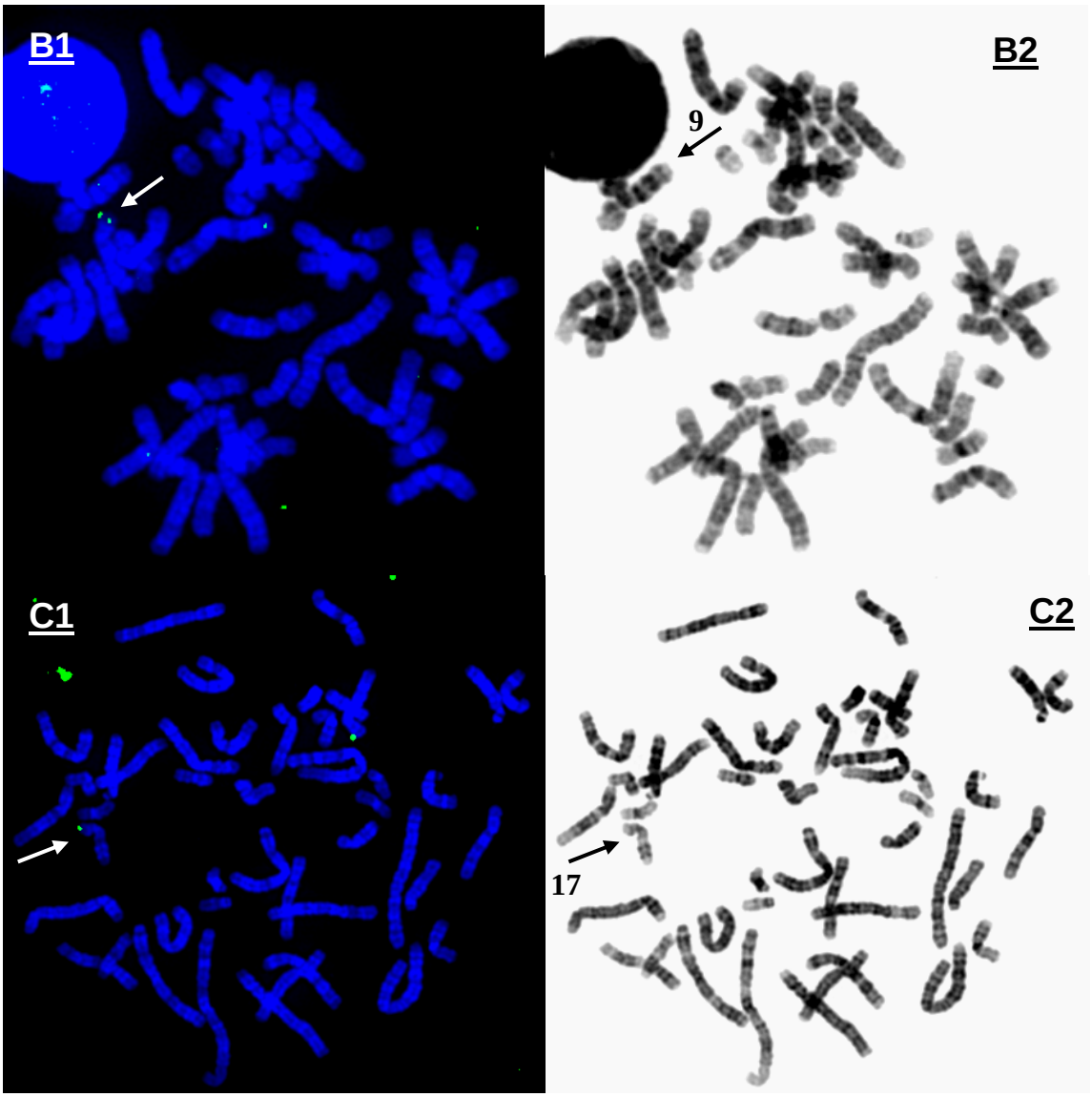


Figure 6.3 Metaphase chromosome analysis of viral integration in patients AMD, AS and KK. 1 metaphase FISH image, 2 inverted DAPI banding of metaphase chromosomes. Results for A patient AMD at 11p15.2-11p15.5, B patient AS at 9q34.12-9q34.3 and C patient KK at 17p13.1-17p13.3.

6.3 Investigation of integration using genomic microarrays

A custom-made genomic microarray was constructed for the quantification of HHV-6B integration into the human genome. The microarray contained a selection of plasmid clones covering 103,922bp of the genome as shown in Figure 6.1. In addition, to increase the resolution of the analysis, the microarray contained 158 PCR products designed to amplify approximately 800-1200bp regions of DNA to tile across the 144.5Kb of unique sequence of the HHV-6 genome. The minimum overlap between PCR products was 9bp and the maximum overlap was 1145bp.

6.3.1 Verification of the microarray

Prior to using the microarray for analysis of viral integration in the 3 patients, the clones on the array assessed for Cy3 versus Cy5 labelling bias using a self versus self hybridisation. The experiment was performed using DNA from patient AMD.

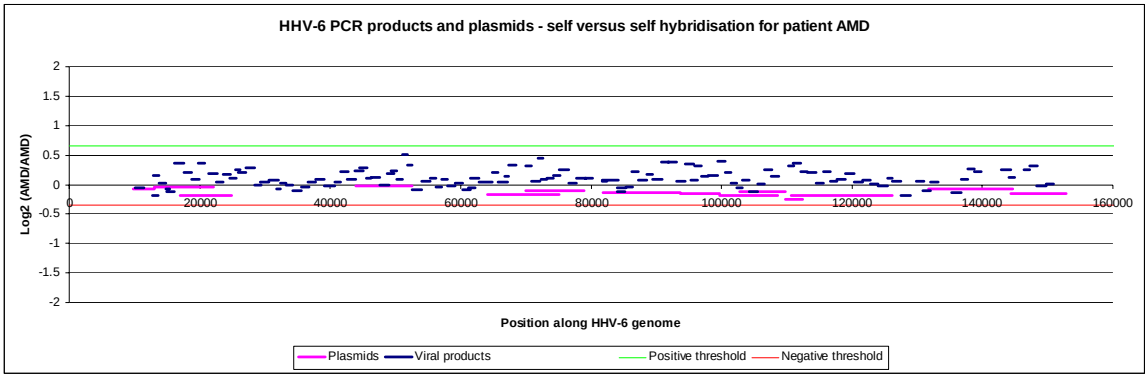


Figure 6.4 Self versus self hybridisation onto custom-made PCR product and plasmid microarray using patient AMD DNA.

All plasmid clones and PCR products were seen to lie within 2 standard deviations of the mean $\log_2$ ratio and therefore no elements on the microarray were excluded on the basis of this experiment.

6.3.2 Comparison between the 3 patients

In order to investigate any differences between the viral copy number integrated into each patient, all three patients were hybridised to the array against each other. Patient AMD was hybridised against patient AS, patient AS hybridised against patient KK and patient AMD against patient KK. Figure 6.5 shows the results of this comparison.

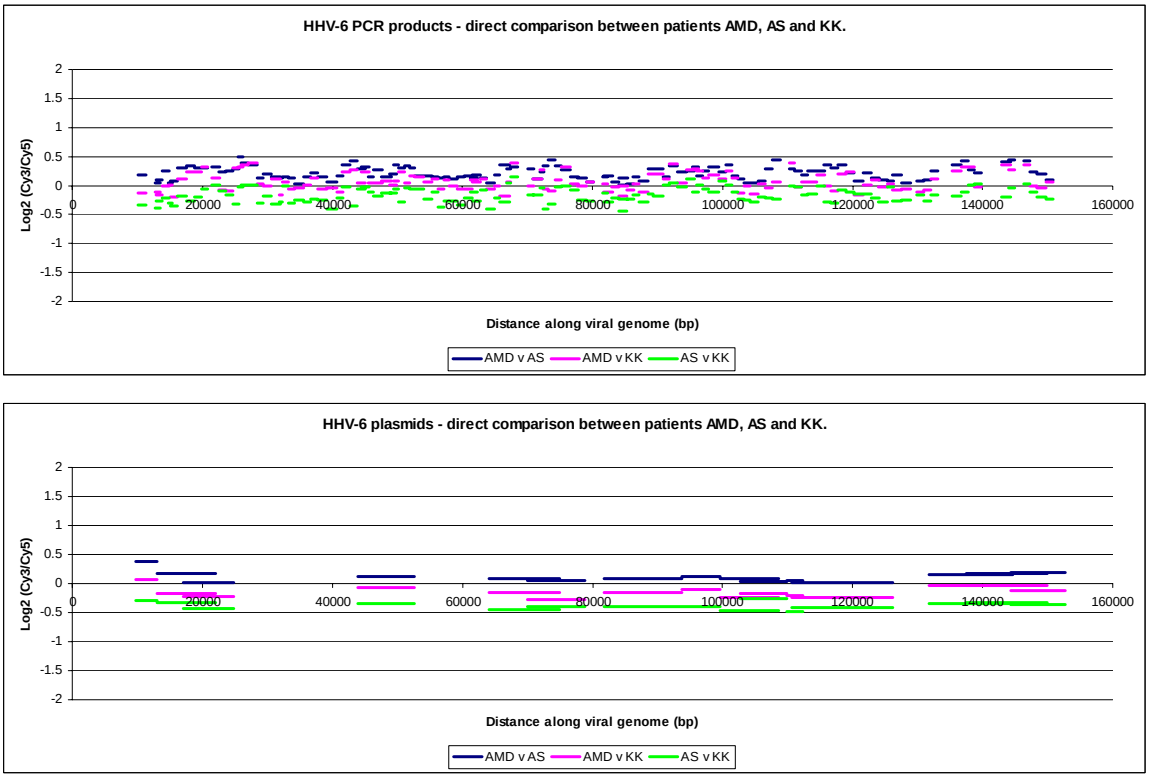


Figure 6.5 Comparative analysis of viral copy number in patients AMD, AS and KK on a custom-made array containing PCR products and plasmids.

Analysis of the average ratio across the PCR products for each hybridisation suggested that patient AMD had a greater number of viral genomes integrated than patient KK which had more than patient AS. The patient samples were analysed comparatively as it was not possible to determine the exact number of integrated viral genomes per patient because there was no sample with a known copy number.

6.4 Investigation of integration using a custom-made library

Patient KK was selected for further analysis because previous FISH analysis had shown that the virus had integrated into chromosome 17p13.1-17p13.3 (Figure 6.3). Patients AMD and AS had shown integration sites on chromosomes 11 and 9 respectively. Neither of these chromosomes can be flow sorted cleanly due to their similar size and constitution to chromosomes 10 and 12. A custom-made fosmid library was created using flow sorted DNA from patient KK to investigate the structure of the integrated virus.

6.4.1 Generation of custom-made library

Chromosomes containing the integrated viral genome along with their normal homologues were isolated from the remainder of the genome by flow sorting (Figure 6.6).

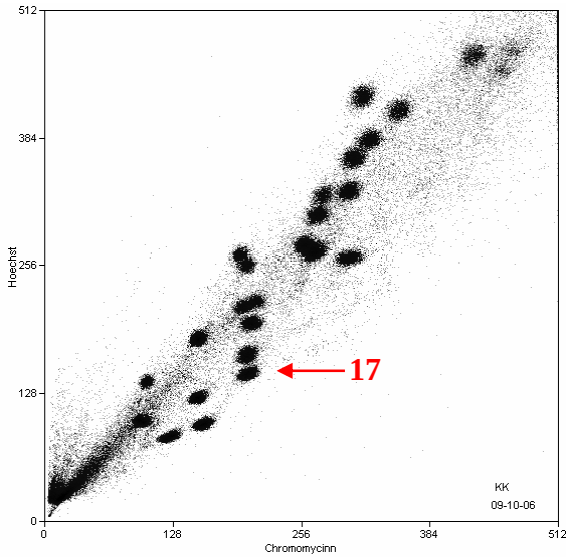


Figure 6.6 Flow karyogram for patient KK with integrated virus on 17p13.1-17p13.3. Flow sorting isolated chromosome 17 from the remainder of the genome, but only half of the flow sorted chromosomes contained integrated virus.

The flow sorted chromosomes were sheared by passing twice through a syringe needle to create an average fragment size of approximately 36Kb (Figure 6.7) and used to generate a custom-made fosmid library.

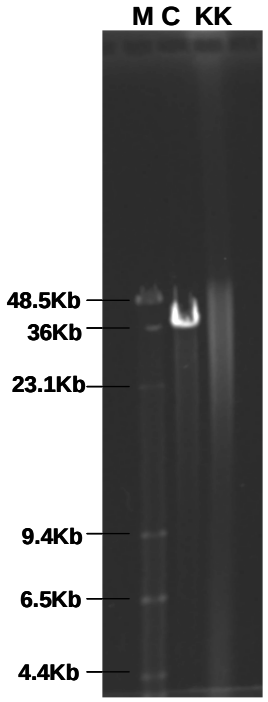


Figure 6.7 Results showing flow sorted chromosome 17 DNA after shearing. KK; sheared patient KK DNA run against M; λ DNA digested with HindIII and C; control DNA of 36Kb supplied with the library production kit.

6.4.2 Screening of custom-made library

After plating of the library, 18,432 colonies were picked into gridded plates by robots, resulting in a theoretical eightfold coverage along the length of chromosome 17. These gridded clones were used to generate filters for library screening (as described in the Methods).

PCR primer pairs were designed to amplify products every 10Kb along the central unique segment of the viral genome. 15 products (Appendix A6) were labelled and hybridised to the gridded library. Eight clones were identified as containing viral DNA; KK_1H19, KK_7B16, KK_8J11, KK_17G11, KK_23M7,

KK_23M8, KK_26O3 and KK_28P16. These clones were end sequenced to verify whether they contained viral DNA only or whether they were chimeric for human and viral DNA and therefore contained the sequence across the integration sites. The end sequence results are summarised in Table 6.1.

Clone name	Left end read	Right end read
KK_1H19	Human; 17;51,582,334-51,582,799	Human; 17;51,538,661-51,539,359
KK_7B16	Virus; 58,660-58,813	Virus; 23,275-23,983
KK_8J11	Virus; 85,739-86,404	Virus; 47,116-47,693
KK_17G11	Virus; 35,848-35,914	Virus; 73,313-73,685
KK_23M7	Virus; 60,127-60,705	Virus; 94,673-95,421
KK_23M8	None	None
KK_26O3	None	None
KK_28P16	None	None

Table 6.1 Alignment of end sequence data to human and viral genomes.

The end sequence reads for clone KK_1H19 aligned to chromosome 17q23.2 approximately 20Mb proximal to the integration site detected by FISH analysis. This data suggested that either the well contained a mixed clone or that whilst the PCR primers used to generate products for library screening were specific for the viral genome, the product created had sufficient homology to hybridise to chromosome 17q23.2.

Four clones; KK7_B16 (35,538bp), KK_8J11 (39,288bp), KK_17G11 (37,465bp) and KK_23M7 (35,294bp) aligned to viral DNA only (Figure 6.8). Analysis of the alignments revealed that the data was consistent with the viral genome integrating intact with no rearrangements.

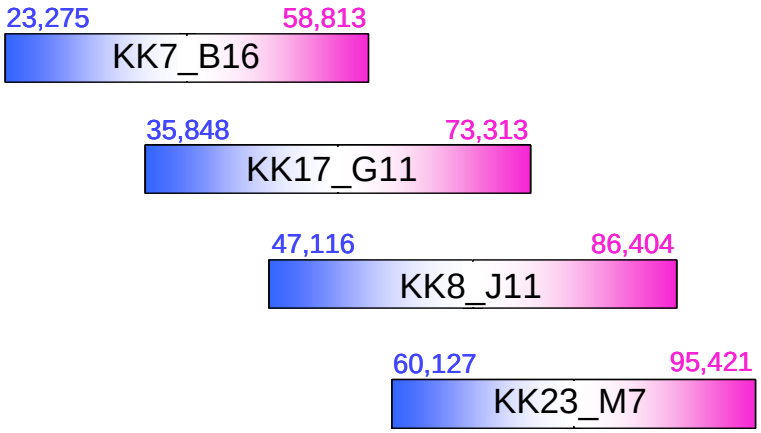


Figure 6.8 Relative positions of the 4 fosmid clones identified through screening of the patient KK library using PCR products spaced every 10Kb along the viral genome.

Three further clones failed to produce end sequence reads; KK_23M8, KK_26O3 and KK_28P16. These clones showed only weak hybridisation signals during the screen suggesting that these are not genuine positive clones.

No clones chimeric for viral and human chromosome 17 sequence were identified.

A second screen of the library was performed using primers designed approximately every 1Kb in the 10Kb of unique sequence close to the direct repeats of the viral genome (Appendix A7). A pool of 9 products was generated for the left end and a pool of 7 products for the right end of the viral genome (pictured in Figure 6.9 with the whole genome primers used in the original screen). However, these PCR product pools did not hybridise to any clones on the gridded filters. Due to time constraints, this experiment was not repeated to determine if this result was due to technical issues, or if the lack of hybridisation signal was indicative of the absence of fosmid clones containing sequence from the first and last 10Kb of the unique region of the viral genome.

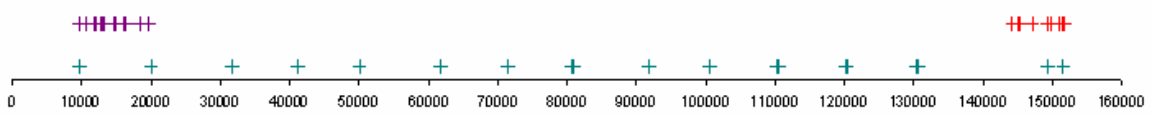


Figure 6.9 Summary of PCR products used for screening the patient KK custom-made fosmid library. The pool of 15 products used in the original screen are shown in green, the pool of 9 products at the left end of the genome are shown in purple and the pool of 7 products at the right end of the genome are shown in red along the length of the viral genome.

6.5 Investigation of integration using FISH onto extended DNA fibres

To verify whether the viral genome was integrated intact and with no rearrangement as hypothesised in section 6.4.2, probes were created from 2 plasmid clones and hybridised to extended chromatin fibres generated from patient KK (Figure 6.10). The plasmid clone from the left end of the viral genome was pHZV14, with an insert size of 8.7Kb and the plasmid clone from the right end of the genome was pH6Z-204 with an insert size of 12.3Kb.

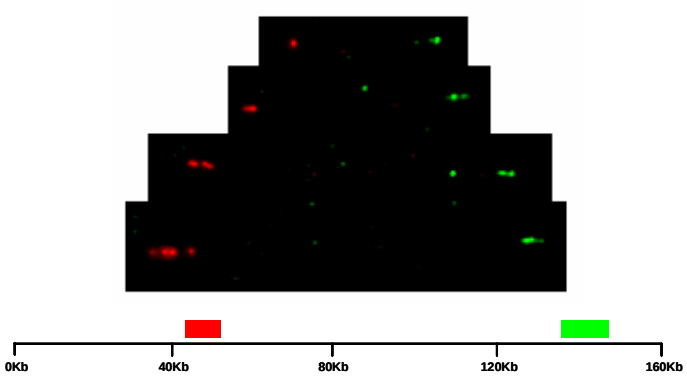


Figure 6.10 Results of plasmid clone FISH hybridisation onto extended chromatin fibres created from patient KK. pHZV14 was detected with Texas Red (red) and pH6Z-204 detected with FITC (green).

The FISH results showed that the 2 plasmids used as probes show a relationship with a consistent gap of approximately 15 times the size of the probes. This suggested that the integrated virus was complete with no rearrangement.

Occasional relationships where the 2 probes were seen to lie extremely close to each other were also observed, however, it was unclear whether this relationship was real suggesting that multiple copies of the virus integrated tandemly, or whether the signals were chance events resulting from the distribution of fibres along the slide.

6.6 Conclusions

The techniques used in this chapter were initially utilised to map rearrangement breakpoints in the human genome and have also been used to investigate viral integration. FISH onto metaphase chromosomes identified the sites of integration in 3 patients into or close to the telomeres of 9q, 11p and 17p and array CGH distinguished that there was variation in the number of integrated viral genomes between the 3 patients. The generation and screening of a custom-made fosmid library showed that the structure of the integrated virus appeared to be intact without rearrangement within the unique region of the genome. Further work to isolate the fosmid clones chimeric for human and viral DNA containing the integration sites would elucidate the sequence across the integration junction and confirm or refute the hypothesis that the HHV-6 virus integrates by way of its human telomeric like repeats.