3: Results 1
Pilot Replication Timing Studies Utilising a Genomic Array Representing 4.5Mb of Chromosome 22 Sequence.

3.1: Introduction.

Initially a small array was constructed to test the use of microarrays for the evaluation of replication timing. This array was designed to cover 4.5Mb of chromosome 22 spanning a boundary between a G dark and a G light band. A total of 83 overlapping clones, including cosmids, fosmids, PACS and BACS, were chosen between BACs CTA-415G2 and CTA-390B3 (inclusive) spanning the region from approximately 17Mb to 21.5Mb along the q arm of chromosome 22q.

3.2: Initial verification experiments on the 4.5Mb array.

To assess the systematic variation in measurements on the array a self:self hybridisation was performed. In this assay, DNA from the same source is used both as the test and the reference DNA and by definition all clones should report a 1:1 ratio. In this case, a self:self hybridisation was conducted on the 4.5Mb test array using differentially labelled DNA from the same sort of G1 nuclei. Analysis of the array rejected 11 out of the 83 points because the triplicate values were not within the rejection criteria described in section 2.5.2 . Briefly, to be included in the analysis, all loci were required to report intensities at least twice the values reported by the *Drosophila* clone DNA loci represented on the array. Triplicate spots were also required to report ratios within 5% of each other. The remaining 72 clones were available for further analysis. The mean G1:G1 ratio was 1.035 with a standard deviation of 0.0714. The G1:G1 ratios plotted against the mid point position of each clone on the array is shown in Figure 3.1. The distance between data points reflects the size of the clone used in the golden path of sequencing clones – a high proportion of cosmids were used between 17 and 18.5Mb while larger insert clone BAC and PACs predominate between 18.5 and 22.6Mb.

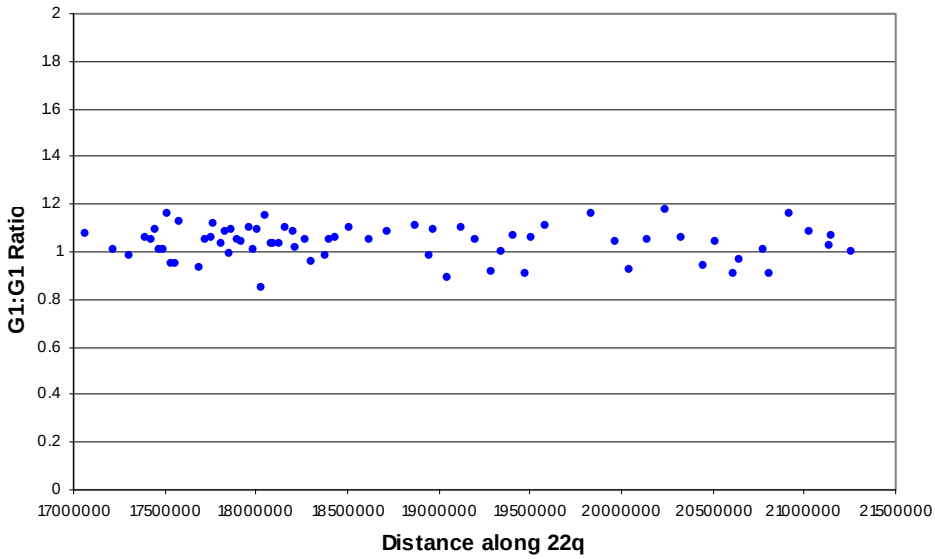

Figure 3.1: G1 self:self hybridisation performed on a 4.5Mb array. Data points on the X axis correspond to the position of the midpoint of each clone, and the Y axis shows the G1:G1 ratio.

3.3: S phase DNA: G1 phase DNA Hybridisation on the 4.5Mb Test Array.

Two S:G1 hybridisations were conducted on the 4.5Mb test array using differentially labelled S and G1 DNA sorted from the same preparation of nuclei. All 84 data points were included in the analysis for both replicates; no data points were rejected at the analysis stage. The mean ratio for each clone is shown in Figure 3.2a, vertical error bars representing one standard deviation on each clone, whilst horizontal error bars represent the extent size of the clone. Replicate experiments are shown in Figure 3.2b. Ratios close to 2:1 indicate early replicating regions whilst loci reporting ratios close to 1:1 replicate late.

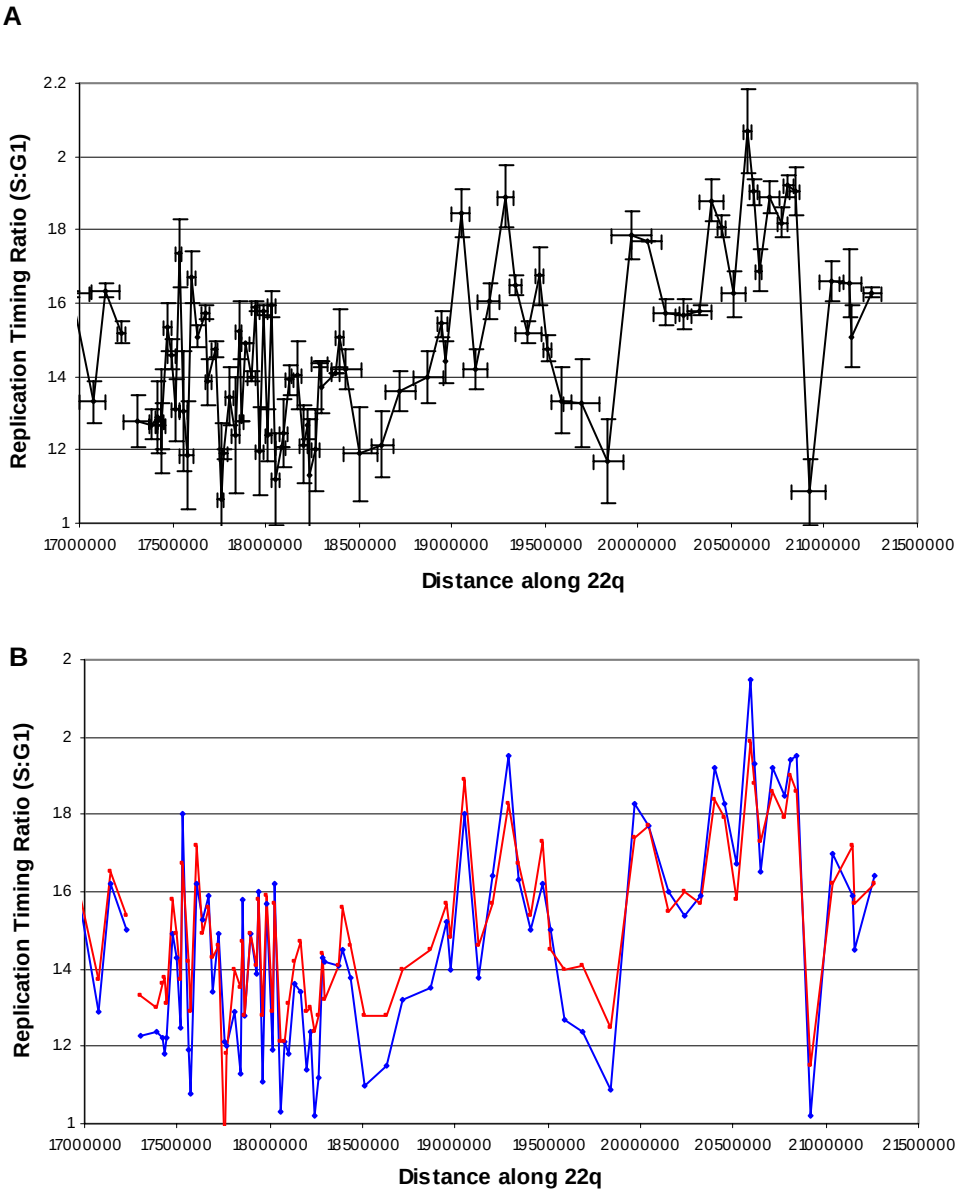

Figure 3.2: replication timing profiles for a 4.5 Mb region of chromosome 22q. A: Average S:G1 ratio of two arrays. A single standard deviation of the two arrays is indicated by the Y error bars. The X error bars represent the length of the clone and indicates the size of the overlap between clones. B: Replication profiles of two individual replicates.

3.4: Correlation between replication timing and sequence features.

The replication timing across the 4.5Mb region of chromosome 22 was also correlated with the guanine and cytosine (GC) content of each clone and the density of introns of genes within each clone.

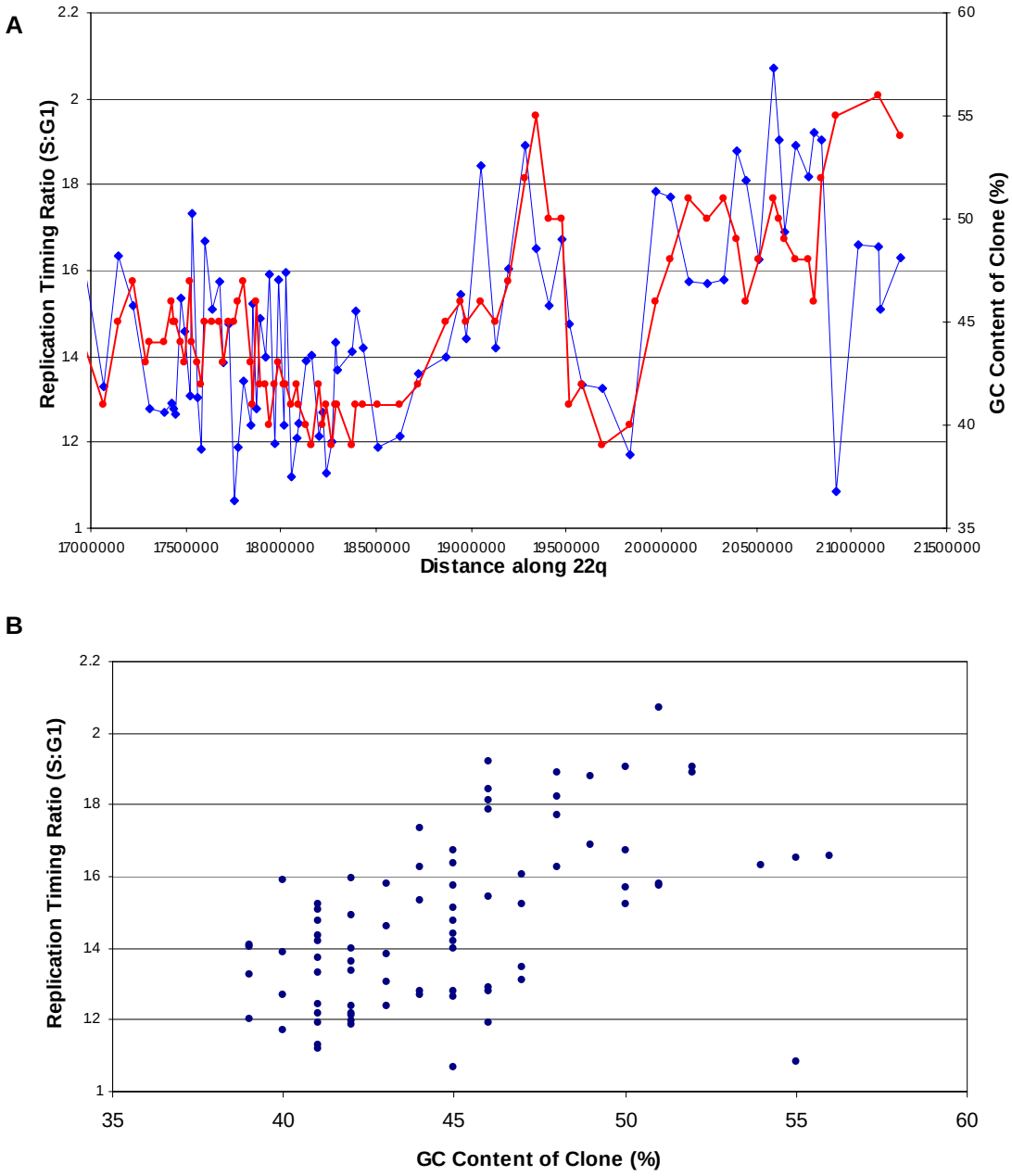

Figure 3.3: Correlation between replication timing and GC content over a 4.5 Mb region. A: Replication timing (blue) and GC content profile (red). B: Replication Timing versus GC Content. The equation of the best fit line through the data points is y = 0.03x + 0.127 with a correlation coefficient of 0.53.

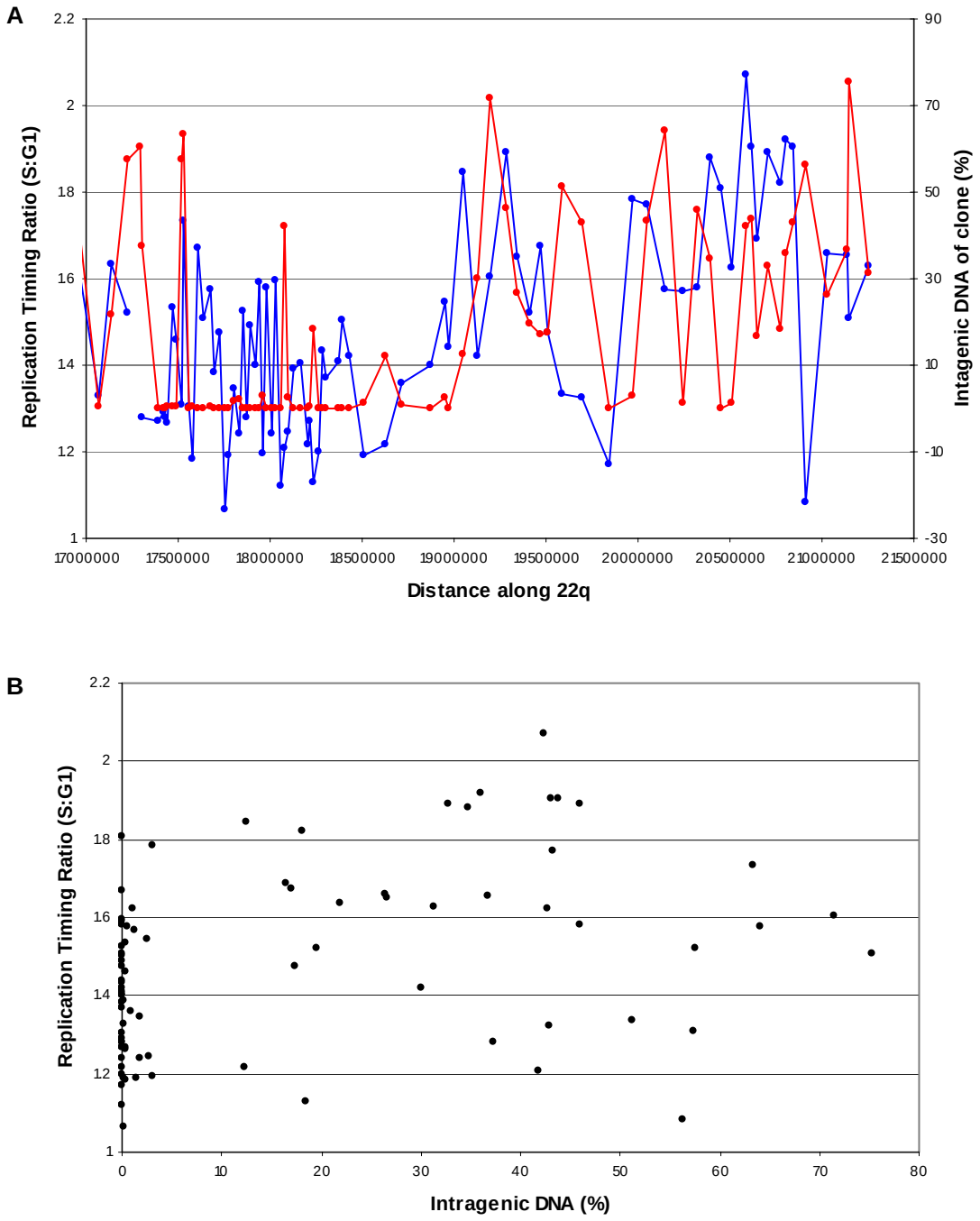

Figure 3.4: Correlation between replication timing and intragenic DNA over a 4.5 Mb region. A: Replication timing (blue) and intragenic DNA (red) profile. B: Replication timing versus intragenic DNA. The equation of the best fit line through the data points is y = 0.003x + 1.41. The correlation coefficient is 0.36

These preliminary experiments showed clear reproducible differences between the replication timing ratios reported for the different loci represented on the array.

3.5: Discussion.

To verify the assessment of replication timing on arrays a small region of chromosome 22 was chosen for study. A microarray was constructed from sequencing clone DNAs across a 4.5Mb region located 17.5 - 21Mb along chromosome 22. The 4.5Mb clone array was initially verified with a G1 self:self hybridisation. A 1:1 ratio should be reported on all clones. An average ratio of 1.035 was reported on the array. A low standard deviation of 0.0714 was observed indicating that all the clones hybridise in a similar fashion.

An early replicating region of the genome will contain twice as much DNA throughout S phase as it will during G1 phase and so the ratio reported will be 2:1. Conversely, a late replicating region of the genome will not duplicate its DNA until the end of S phase so that in this assay a ratio of close to 1:1 will be generated. In this way, the replication timing can be reported as a ratio of S:G1 DNA which should vary between 1 and 2. On this test array, all the ratios (except one) were between 1 and 2. The clones with mid points between 17-18.5 Mb along the q arm of chromosome 22 replicated later than clones with midpoints between 18.5-21Mb. This is consistent with the fact the proximal region corresponds to a G dark band and the distal region corresponds to a G light band (Strehl, LaSalle et al. 1997).

Comparison of the S:G1 ratio with GC content showed that GC rich clones generally reported earlier replication timing ratios than GC poor clones. This is in agreement with of previous studies (Tenzen, Yamagata et al. 1997; Watanabe, Fujiyama et al. 2002). The correlation with density of intragenic DNA was less clear. However a positive correlation was still observed as has been reported previously (Strehl, LaSalle et al. 1997; Cook 1999; Gilbert 2002).

These initial experiments on the 4.5Mb region validated this assay as having the potential to accurately assess replication timing and it was decided to expand this approach by developing a microarray covering the whole of human chromosome 22q.