

| PublisherInfo | | |
|----------------------|---|----------------|
| PublisherName | : | BioMed Central |
| PublisherLocation | : | London |
| PublisherImprintName | : | BioMed Central |

Unraveling chromosome 22

| ArticleInfo | | |
|-----------------------|---|--|
| ArticleID | : | 3553 |
| ArticleDOI | : | 10.1186/gb-2000-1-1-reports019 |
| ArticleCitationID | : | reports019 |
| ArticleSequenceNumber | : | 44 |
| ArticleCategory | : | Paper report |
| ArticleFirstPage | : | 1 |
| ArticleLastPage | : | 4 |
| ArticleHistory | : | RegistrationDate : 2000-1-18 Received : 2000-1-18 OnlineDate : 2000-3-17 |
| ArticleCopyright | : | BioMed Central Ltd2000 |
| ArticleGrants | : | |
| ArticleContext | : | 130591111 |

Abstract

The first publication of the genomic sequence of an entire human chromosome.

Significance and context

There are two technically distinct strategies being championed by the different human genome projects. The publicly funded international consortium has chosen a 'clone-by-clone' approach that began with the careful construction of genomic maps and the subsequent systematic sequencing of overlapping clones. In contrast, an alternative strategy is being attempted by a private American company, Celera Genomics, who are directly sequencing random DNA fragments that will subsequently be assembled in a 'random shotgun' approach. It will be only a matter of months until we know who gets there first. The publication of the sequence for human chromosome 22 represents a 'proof of principle' of the clone-by-clone approach, and is a sequencing landmark attesting to the strength of international public collaboration. Several considerations contributed to the choice of chromosome 22. It is the second smallest autosome and has been linked to a number of human diseases including DiGeorge syndrome, spinocerebellar ataxia and schizophrenia. The short arm (22p) encodes tandemly repeated ribosomal RNA genes and no protein-coding genes, whereas the long arm (22q) is thought to be gene-rich. This paper reports the sequencing of 22q.

Key results

Extensive clone maps were constructed using a variety of vectors - cosmids, fosmids, BACs and PACs (bacterial and P1-derived artificial chromosomes). DNA sequencing generated 12 contigs spanning the long arm of chromosome 22 from the centromeric tandem repeats to the 22q telomere. There are 11 small gaps indicating the presence of 'unclonable' sequences. The entire sequence is 33,464 kilobases, representing approximately 97% coverage. The error rate is reported as less than one error per 50 kilobases. The authors report the results of extensive computational analysis using a suite of similarity searches and prediction tools. They identified 545 genes, which they divided into 'known genes' (247), homologues or 'related genes' (150), and 'predicted genes' with homology to expressed sequence tags (ESTs). In addition, there are 134 pseudogenes. The total length of identified genes accounts for 39% of genomic sequence. They predict that perhaps 200-300 more genes may be present, but these genes are not easily located with current methodologies. They also report detailed statistics on

average exon lengths, gene lengths, base composition (GC content), and so on. The sequence also offers insights into chromosomal landscape with respect to repeats, tandem duplications and gene families.

Links

[Human genome project information](#) is available from the [Oak Ridge National Laboratory](#). The [University of Oklahoma's Advanced Center for Gene Technology](#) provides information on [Human and mouse genomic sequencing](#) and the [Sanger Centre: human chromosome 22](#) site is devoted to chromosome 22.

Conclusions

DNA sequence generated in a number of international laboratories (including the Wellcome Trust Sanger Centre, UK, the University of Oklahoma, USA, Washington University, USA, and Keio University, Japan) has been successfully assembled into the largest contiguous segment of DNA sequence produced so far, showing that the clone-by-clone strategy can generate long-range genomic continuity.

Reporter's comments

The successful sequencing of chromosome 22 shows the benefits of the open policy of speedy data release championed by the public genome project, and the value of the clone-by-clone approach. The recent completion of the sequencing of the *Drosophila* genome has, however, shown that the shotgun approach can also be powerful and efficient. The chromosome 22 study authors admit that although qualifying as 'operationally complete' there are still several gaps in the sequence. These are concentrated at the centromere and telomere and may be due to specific sequence features that defy cloning or sequencing.

Table of links

Nature

[Human genome project information](#)

Oak Ridge National Laboratory

University of Oklahoma's Advanced Center for Gene Technology

Human and mouse genomic sequencing

Sanger Centre: human chromosome 22

References

1. Dunham I, Shimizu N, Roe BA, Chissoe S, Hunt AR, Collins JE, Bruskiewich R, Beare DM, Clamp M, Smink LJ, et al: The DNA sequence of human chromosome 22. *Nature*. 1999, 402: 489-494. 0028-0836