Title: METHODS FOR HIGH RESOLUTION GENE MAPPING

Abstract

Two convenient methods for ordering discrete DNA sequences at a high resolution are described. In one embodiment, the invention features a) contacting a eukaryotic cell containing a chromosome of interest with an agent that effects decondensation; b) preparing the cell containing decondensed chromatin for hybridization; and c) hybridizing the cell produced in step b) with at least two DNA probes, each probe being complementary to at least one discrete DNA sequence present on the chromosome of interest; and d) detecting the presence and relative order of the probes as an indication of the relative and absolute locations of the DNA sequences on the chromosome. In another embodiment, the invention features ordering at least two discrete DNA sequences on a fragment of DNA that has been extended, for example by physical means, chemical means or a combination of both. A preferred physical means for extending DNA is by gentle smearing.
### FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>AT</td>
<td>Austria</td>
<td>FR</td>
<td>France</td>
<td>MR</td>
<td>Mauritania</td>
</tr>
<tr>
<td>AU</td>
<td>Australia</td>
<td>GA</td>
<td>Gabon</td>
<td>MW</td>
<td>Malawi</td>
</tr>
<tr>
<td>BB</td>
<td>Barbados</td>
<td>GB</td>
<td>United Kingdom</td>
<td>NL</td>
<td>Netherlands</td>
</tr>
<tr>
<td>BE</td>
<td>Belgium</td>
<td>GN</td>
<td>Guinea</td>
<td>NO</td>
<td>Norway</td>
</tr>
<tr>
<td>BF</td>
<td>Burkina Faso</td>
<td>GR</td>
<td>Greece</td>
<td>NZ</td>
<td>New Zealand</td>
</tr>
<tr>
<td>BG</td>
<td>Bulgaria</td>
<td>HU</td>
<td>Hungary</td>
<td>PL</td>
<td>Poland</td>
</tr>
<tr>
<td>BJ</td>
<td>Benin</td>
<td>IE</td>
<td>Ireland</td>
<td>PT</td>
<td>Portugal</td>
</tr>
<tr>
<td>BR</td>
<td>Brazil</td>
<td>IT</td>
<td>Italy</td>
<td>RO</td>
<td>Romania</td>
</tr>
<tr>
<td>CA</td>
<td>Canada</td>
<td>JP</td>
<td>Japan</td>
<td>RU</td>
<td>Russian Federation</td>
</tr>
<tr>
<td>CF</td>
<td>Central African Republic</td>
<td>KP</td>
<td>Democratic People's Republic of Korea</td>
<td>SD</td>
<td>Sudan</td>
</tr>
<tr>
<td>CG</td>
<td>Congo</td>
<td>KR</td>
<td>Republic of Korea</td>
<td>SE</td>
<td>Sweden</td>
</tr>
<tr>
<td>CH</td>
<td>Switzerland</td>
<td>KZ</td>
<td>Kazakhstan</td>
<td>SK</td>
<td>Slovak Republic</td>
</tr>
<tr>
<td>CI</td>
<td>Côte d'Ivoire</td>
<td>LK</td>
<td>Sri Lanka</td>
<td>SN</td>
<td>Senegal</td>
</tr>
<tr>
<td>CM</td>
<td>Cameroon</td>
<td>LU</td>
<td>Luxembourg</td>
<td>SO</td>
<td>Soviet Union</td>
</tr>
<tr>
<td>CS</td>
<td>Czechoslovakia</td>
<td>MC</td>
<td>Madagascar</td>
<td>TD</td>
<td>Chad</td>
</tr>
<tr>
<td>CZ</td>
<td>Czech Republic</td>
<td>MG</td>
<td>Madagascar</td>
<td>TG</td>
<td>Togo</td>
</tr>
<tr>
<td>DE</td>
<td>Germany</td>
<td>ML</td>
<td>Mali</td>
<td>UA</td>
<td>Ukraine</td>
</tr>
<tr>
<td>DK</td>
<td>Denmark</td>
<td>MN</td>
<td>Mongolia</td>
<td>US</td>
<td>United States of America</td>
</tr>
<tr>
<td>ES</td>
<td>Spain</td>
<td></td>
<td></td>
<td>VN</td>
<td>Viet Nam</td>
</tr>
<tr>
<td>FI</td>
<td>Finland</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
METHODS FOR HIGH RESOLUTION GENE MAPPING

Background of the Invention

The human genome consists of about 3 billion base pairs of DNA carried on 46 chromosomes. This genetic blueprint encodes the information required for the growth, differentiation, maintenance and proper functioning of human cells. To aid in identifying genes associated with disease, there has been great interest in mapping the human genome. Indeed, in creating the Human Genome Project, the United States government has announced the elucidation of the human genome as a national objective.

The two main types of human genome maps are genetic linkage and physical. Genetic linkage maps are generated mainly by studying families and measuring the frequency with which two different traits are inherited together, or linked. Physical maps, on the other hand, are derived from measurements made on the DNA molecules that form the human genome.

Physical maps can be created based on information provided by restriction fragment polymorphisms or a collection of ordered clones of genomic fragments (e.g. cosmids). Maps based on ordered clones are especially useful, since each fragment is available as a clone that can be propagated and distributed. These clones can serve as a starting point for gene isolation, functional analyses and for the determination of nucleotide sequences. Preparing an ordered-clone collection involves cloning DNA fragments, determining their order in the genome and propagating the fragments in pure form to make them widely available for subsequent analysis. Because of the limited resolution currently available, the most difficult aspect involves determining the order of DNA fragments.

In situ hybridization analysis has been useful for genome mapping, because it allows direct detection of the location at which a DNA fragment maps on a particular chromosome. However, the resolution for localizing probes to chromosomal bands in
metaphase chromosomes is only several megabases (Lichter, P., et. al., et. al., Science 247:64-69 (1990); Lawrence, J.B., et. al., Science 249:64-69 (1990)). To achieve higher resolution, FISH has been applied to cell targets collected in interphase (Lawrence, J.B. et. al., Cell 52:51-61 (1988)). At interphase, as opposed to at metaphase, chromatin within the cell nucleus exists in an uncondensed form. Probes separated by distances from <100 K up to at least 1Mb have been ordered on the X chromosome in somatic interphase nuclei (Lawrence, J.B., et. al., Science 249:64-69 (1990); Trask, B.J., et. al., Am. J. Hum. Genet. 48:1-15 (1991)). Using male interphase pronuclei generated by gamete fusion, probes separated by less than 50kb have been resolved. (Brandriff, B., et. al., Genomics 10:75-82 (1991)). However, this technique is time consuming, requires a high level of skill and requires materials that are not readily available to most researchers.

Numerous probes are generated in efforts to identify gene loci. A need exists for a simple and effective high resolution method for determining the relative order of these probes and to determine their position relative to known markers.

**Summary of the Invention**

In general, the invention relates to methods for ordering at least two discrete DNA sequences on a chromosome. In one embodiment, the invention features a) contacting a eukaryotic cell containing a chromosome of interest with an agent that effects decondensation; b) preparing the cell containing decondensed chromatin for hybridization; and c) hybridizing the cell produced in step b) with at least two DNA probes, each probe being complementary to at least one discrete DNA sequence present on the chromosome of interest; and d) detecting the presence and relative order of the probes as an indication of the relative and absolute locations of the DNA sequences on the chromosome. By decondensing interphase chromatin according to the method of the invention, high resolution interphase mapping can be carried out. In effect, fragments containing DNA sequences that abut one another or even overlap can be resolved. In addition, because the methods disclosed herein render
the DNA more accessible to hybridization, results are more readily obtained.

In another embodiment, the invention features ordering at least two discrete DNA sequences on a fragment of DNA that has been extended, for example by physical means, chemical means or a combination of both. A preferred physical means for extending DNA is by gentle smearing. A preferred chemical means for obtaining extended DNA from a eukaryotic genome is by contacting the chromatin with a solution that removes histones. Once extended, the DNA can be prepared for hybridization and hybridized with at least two DNA probes, each probe being complementary to at least one discrete DNA sequence present on the chromosome of interest. Once hybridized, the presence and relative order of the probes can be determined as an indication of the relative and absolute locations of the DNA sequences on the fragment of DNA. By extending DNA according to the method of the invention, mapping can be carried out at an even higher resolution.

Detailed Description of the Invention

The subject invention is based on the discovery that chromosomes which are "relaxed" or DNA which is "extended" can be analyzed using a hybridization assay to determine the order in which a set of DNA sequences (i.e. at least two) map on a particular chromosome or fragment of DNA. The invention, therefore, features two convenient methods for ordering a set of discrete DNA sequences at a high resolution. The ordering of discrete DNA sequences is useful for example for physical mapping. As used herein, the term "ordering" means establishing the linear relationship of discrete DNA sequences relative to one another and/or relative to a known marker on a chromosome or portion of a chromosome.

In one aspect of the subject invention, a cell is contacted with an agent that effects decondensation, prepared for hybridization and analyzed using at least two DNA probes, each probe being complementary to at least one discrete DNA sequence present on a chromosome of interest. Preferably the cells are in interphase, so that the chromatin is already in a somewhat extended form. The growth of cells in
culture can be synchronized using known methods. By using synchronized cells, a higher proportion of cells in the interphase stage of the cell cycle can be obtained.

According to the method of the subject invention, a cell is first contacted with an agent that effects decondensation. For example, an agent that effects decondensation can be a chemical which inhibits proteins that deacetylate histones (i.e. the set of proteins that interact with DNA to form chromatin). When histones are hyperacetylated, the chromatin becomes less condensed or relaxed. A salt of a short chain fatty acid (such as sodium butyrate, sodium propionate and sodium valerate) when added to cells appears to relax chromatin, apparently by inhibiting the proteins that deacetylate histones.

Cells containing relaxed chromatin are then prepared for in situ hybridization analysis using methods, which are well-known in the art. In general such methods involve depositing cells on a solid substrate, fixing, drying and denaturing them to yield single stranded DNA that is therefore available for hybridization. Cells can be deposited on a substrate such as glass, plastic or nitrocellulose. A glass microscope slide is preferred, because it can be readily manipulated and viewed under a microscope. The substrate can be pretreated with a "cell adherent" which improves the likelihood that a cell settling onto the surface remains attached during subsequent manipulations. A preferred cell adherent is 3-Aminopropytriethoxysilane. Treatment with this adherent results in "silanized" substrates. Other adherents include poly-L lysine and mussel adhesin. Pretreatment of solid substrates can be accomplished using any method that ensures that a cell adherent is deposited (e.g. submersion, transferring using a dropper, etc.). Pretreated solid substrates can be stored in a dust free environment at room temperature.

Cells on a substrate can then be processed through a fixation protocol to preserve the nuclei/ chromosome in a morphologically stable state so that nucleic acids are retained through the rigorous conditions present during in situ hybridization. Appropriate fixatives are well-known in the art and include, for example, 4%
paraformaldehyde or glutaraldehyde in phosphate buffered solution (PBS) containing
5 mM MgCl₂, a fixative containing 3 parts ethanol and 1 part acetic acid, Carnoy’s
fixative, 1% osmium tetroxide, Bouin’s fixative, Zenker’s fixative.

Prepared cells containing relaxed chromatin can then be hybridized with
appropriate probes and detected using methods which are well-known in the art, e.g.
hybridization and detection can be performed basically as described by Lichter et. al.
(Hum. Genet. 80:224-234 (1988)).

For use in the subject invention, deoxyribonucleic acid (DNA) probes, labelled
with a detectable marker can be prepared from single-stranded DNA molecules or
fragments thereof according to procedures which are well known in the art. Such
techniques include incorporation of radioactive labels, direct attachment of
fluorochromes or enzymes, and various chemical modifications of the nucleic acid
fragments that render them detectable immunochemically or by other affinity
reactions. A preferred method of labelling is by nick translation using a haptened
nucleotide triphosphate (e.g. biotin labelled dUTP) or by random primer extension
(Feinberg & Vogelstein, Anal. Biochem. 137:266-267 (1984) (e.g. multiprime DNA
labeling system (Amersham) substituting dTTP with Bio-11-dUTP. (Langer, P.R., et. al.,
(1983)).

For DNA sequence ordering, each of the probes should be distinctly labelled.
Three sets of distinguishable fluorophores, emitting in the green (e.g fluorescein), in the
red (e.g. rhodamine or Texas Red), and in the blue (e.g. AMCA or Cascade Blue) are
typically used for fluorescent in situ hybridization (FISH). Therefore using standard
procedures, at least three probes can be used in combination for DNA sequence
ordering. More than three probes can be detected simultaneously using the methods
only can be accomplished by determining the position of probes relative to one
another, but ordering can also make use of known chromosome markers (e.g.
centromeres and telomeres).
In another embodiment, the invention features ordering at least two discrete DNA sequences on a fragment of DNA that has been extended. Such DNA can be obtained, for example from a eukaryotic or a prokaryotic chromosome, although histones must first be removed from DNA obtained from a eukaryotic chromosome. Histones can be removed chemically using a histone removing solution (e.g. see Paulson and Laemmli Cell 12:817-828 (1977)).

DNA free in solution (e.g. DNA extracted from a prokaryote or DNA extracted from a eukaryotic chromosome and treated to remove proteins) can be extended by physical means. A preferred physical means for extending DNA is by gentle smearing or pulling of the DNA, e.g. using a pipette. Once extended, DNA can be prepared for hybridization and hybridized with appropriate probes as described above to determine the relative and absolute locations of the DNA sequences on the fragment of DNA.

The present example will now be further illustrated by the following examples, which are not intended to be limiting in any way.

**Example 1: Preparation of Nuclei Containing Decondensed Chromatin**

1. Short term lymphocyte cultures were established using standard protocols and incubated for 72 hrs at 37°C.

2. Sodium butyrate was added to cultures to a concentration of 7mM (from a 500mM stock) and incubated an additional 5-6 hrs at 37°C.

3. Cells were resuspended in BT buffer (20mM sodium butyrate, 1mM Tris, 25mM KCl, 0.9mM MgCl₂ and 0.9mM CaCl₂ at pH 7.6).

4. Cells were resuspended in BT/PMSF (BT buffer plus 0.5mM phenylmethylsulfonyl fluoride (PMSF)).

5. Cells were resuspended in BT buffer, then several drops of Carnoy’s fix (3 parts methanol: 1 part acetic acid) was added to the cell suspension and gently mixed.
6. Cells were suspended in Carnoy's, followed by several changes of Carnoy's.

5 7. The cell suspension was dropped from a height of about 2 ft onto humid slides and allowed to dry slowly in a humid atmosphere.

8. Slides were aged overnight on a 60°C warmer before use.

Example 2: Preparation of Extended DNA Molecules

1. Cells were to grown to confluence on glass coverslips in small petri dishes.

2. The culture media was removed by aspiration and 0.075M KCl with 0.01% Triton was added, and culture incubated for 20 m at 37°C.

3. KCl/Triton was removed, then cold chromosome isolation buffer (1.0M hexylene glycol, 0.5mM CaCl₂, and 0.1mM PIPES; Wray and Stubblefield Exp. Cell Res. 59:469-478 (1970)) was gently added and incubated for 20m at 4°C.

4. The buffer was removed, then a cold histone removing solution (0.2 mg/ml dextran sulphate, 0.02 mg/ml heparin, 10mM EDTA, 10mM Tris-HCl pH 9.0, 0.1% Nonidet P-40, and 1.0mM PMSF; Paulson and Laemmli, Cell 12:817-828 (1977)) was added, and incubated 30m at 4°C.

5. Most of the solution (leaving the coverslip submerged) was removed then 50% Carnoy's fix (in dH₂O) was gently added, and allowed to stand for 5 m.

6. This solution was removed and the coverslip was allowed to dry.
7. The coverslip was flooded with Carnoy’s, allowed to stand for 10 m and followed with several changes of Carnoy’s.

8. The coverslip was removed and allowed to air dry.

Example 3: Ordering Cosmids from Contigs which Map to Human Chromosomes 18 and 21 to Demonstrate the Resolution Achieved by the Methods Described in Examples 1 and 2.

Material was prepared as described in Examples 1 and 2.

Probes consisted of cosmids from a chromosome 18 or chromosome 21 contig.

Chromosome 18 contig: This contig maps to chromosome 18q23 and comprises five cosmids spanning 109 kb. This DNA molecule can be obtained from the American Type Culture Collection (Rockville, MD) under Accession Number 68934. The two flanking cosmids (pWE7 (labelled with a green fluorophore)) and pWE50wl (labelled with a green fluorophore) and the central cosmid (pWE50 (labelled with a red fluorophore)) were used in these experiments. There is known to be a large gap between pWE50 and pWE50wl; restriction digests indicate that pWE7 and pWE50 and may overlap.

Chromosome 21 contig: This contig maps to chromosome 21q22.3 and consists of five cosmids extending 118 kb. The two flanking cosmids (chC1-8A and pWE18.3w12) and the central cosmid (pWE 19.2) were used in these experiments. This contig has been restriction mapped, and there is a 2.3 kb overlap of chC1-8A with pWE19.2 and a 1.5 kb gap between pWE19.2 and pWE18.3w12.

All probes were labeled by nick translation. Hybridization cocktails contained 5-8 ng/µL of each flanking cosmid labeled with digoxigenin-dUTP of biotin and 5-8 ng/µL of the central cosmid labeled with the alternate hapten, 200 ng/µL human Cot-1 DNA, and 800 ng/µL of salmon DNA in 6XSSC, 10% Dextran Sulfate. Suppression hybridization and washing were essentially as described in Klinger et al.
Am J Hum Genet. 51(1): in press (1992), with the following modifications. The pre-
detection washes for coverslips containing extended DNA molecules were: 1X5m
each in 2XSSC, 0.2XSSC, 0.1XSSC; 3X5m in 60% formamide in 0.1XSSC; 1X5 in 2XSSC;
blocked with 3% BSA/4XSSC for 5m. All coverslips were washed at room temperature.

Hybridization was detected with 0.5 μg/ml FITC-anti digoxigenin and 2.0 μg/ml Cy3-
streptavidin. Photographs were taken directly from the microscope with Kodak Gold
400 film using FITC-Texas Reid (Omega Optical) or FITC-TRITC (Chroma Technology)
dual band pass filters. Using this labelling and detection scheme, the expected
pattern of fluorescence would be red-green-red or green-red-green.

Results

Resolution of about 70kb was consistently achieved with butyrate treated
material and adjacent cosmids (30-40kb) were occasionally resolved. By "relaxing"
interphase chromatin, sodium butyrate rendered it more accessible to hybridization.
These preparations were very clean; cytoplasm was rarely associated with the nuclei
or metaphase spreads. Metaphase chromosomes typically appeared "fuzzy" in these
preparations, suggesting the presence of loops of decondensed chromatin which
made them less appealing for chromosome mapping, but indicated that butyrate
had the expected effect of decondensing the chromatin.

Hybridization was consistently detected on extended DNA molecules. In
the more extended filaments, the signal appeared as a string of fluorescent spots,
although a considerable variation in the degree of condensation among the
hybridized molecules as well as within the molecule was noticed. In one trial, the
entire length of the green-red-green pattern representing the 18 contig was about
40um; the expected length for 109kb of B-DNA is about 37um. The three cosmids
could be uniquely identified in these preparations. Consistent patterns of
fluorescence (representing the hybridized cosmids) were identified, especially where
the cosmids met. Overlapping or abutting cosmids tended to run together, whereas
non-fluorescence gaps were conspicuous.
A large gap was evident between one red-green pair, consistent with the relationship between pWE50 and pWE50wl. The other red-green pair appeared to be continuous. This general pattern of large gap between one red-green pair and a small or no gap between the other was consistently observed. This resolution enabled the determination of absolute order of the three cosmids.

The three cosmids from the chromosome 21 contig were accurately resolved by hybridization to extended DNA molecules. The 2.3kb overlap and the 1.5kb gap could be identified in many of the filaments.

10 **Equivalents**

Those skilled in the art will recognize or be able to ascertain using no more than routine experimentation many equivalents to the specific embodiments of the invention described specifically herein. Such equivalents are intended to be encompassed in the scope of the following claims.
CLAIMS

1. A method for ordering at least two discrete DNA sequences on an interphase chromosome of interest comprising the steps of:
   a) contacting a eukaryotic cell containing an interphase chromosome of interest with an agent that effects decondensation, thereby producing decondensed chromatin;
   b) preparing the cell containing relaxed chromatin for hybridization;
   c) hybridizing the cell produced in step b) with at least two distinct labelled DNA probes, each probe being complementary to at least one discrete DNA sequence present on the chromosome of interest; and
   d) detecting the presence and relative order of the probes as an indication of the relative and absolute locations of the DNA sequences on the chromosome.

2. A method according to Claim 1 wherein the agent that effects decondensation is a salt of a short chain fatty acid.

3. A method according to Claim 1 wherein the agent that effects decondensation is a salt of a short chain fatty acid.

4. A method for ordering at least two discrete DNA sequences on a fragment of DNA, comprising the steps of:
   a) smearing a fragment of DNA; thereby producing smeared DNA;
   b) preparing the smeared DNA for hybridization;
   c) hybridizing the DNA produced in step b) with at least two distinct labelled DNA probes, each probe being complementary to at least one discrete DNA sequence present on the chromosome of interest; and
   d) detecting the presence and relative order of the probes as an indication of the relative and absolute locations of the DNA sequences on the fragment of DNA.
5. A method of Claim 4, wherein step a), the smearing of the DNA is accomplished by physical means.

6. A method of Claim 4, wherein step a), the smearing of the DNA is accomplished by chemical means.

7. A method for ordering at least two discrete DNA sequences on an interphase chromosome of interest, comprising the steps of:

   a) contacting a eukaryotic cell containing an interphase chromosome of interest with an agent that causes histone hyperacetylation, thereby producing decondensed chromatin;

   b) preparing the cell containing the decondensed chromatin for hybridization;

   c) hybridizing the cell produced in step b) with at least two distinct labelled DNA probes, each probe being complementary to at least one discrete DNA sequence present on the chromosome of interest; and

   d) detecting the presence and relative order of the probes as an indication of the relative and absolute locations of the DNA sequences on the chromosome.

8. A method of Claim 7, wherein the agent that causes histone hyperacetylation is a salt of a short chain fatty acid.

9. A method of Claim 8, wherein the salt of a short chain fatty acid is selected from the group consisting of sodium butyrate, sodium propionate and sodium valerate.

10. A method of Claim 8, wherein the salt of the short chain fatty acid is sodium butyrate.
11. A method for locating at least two discrete DNA sequences on an interphase human chromosome of interest comprising the steps of:

a) contacting a human cell containing an interphase chromosome of interest with an agent, that causes histone hyperacetylation, thereby producing decondensed chromatin;

b) preparing the cell containing the decondensed chromatin for hybridization;

c) hybridizing the cell produced in step b) with at least two distinct labelled DNA probes, each probe being complementary to at least one discrete DNA sequence present on the chromosome of interest; and

d) detecting the presence and relative order of the probes as an indication of the relative and absolute locations of the DNA sequences on the chromosome.

12. A method of Claim 11, wherein the agent that causes histone hyperacetylation is a salt of a short chain fatty acid.

13. A method of Claim 12, wherein the salt of a short chain fatty acid is selected from the group consisting of sodium butyrate, sodium propionate and sodium valerate.

14. A method for ordering at least two discrete DNA sequences on an interphase chromosome comprising contacting a eukaryotic cell containing an interphase chromosome with a histone removing agent and analyzing the cell by fluorescence in situ hybridization using at least two distinct labelled probes which are complementary to at least two discrete DNA sequences on the chromosomes.

15. A method for ordering at least two discrete DNA sequences on a fragment of
DNA, thereby producing smeared DNA and analyzing the smeared DNA by fluorescence in situ hybridization using at least two distinct labelled probes which are complementary to at least two discrete DNA sequences on the smeared DNA.

16. A method of Claim 2, wherein the salt of a short chain fatty acid is sodium butyrate.

17. A method of Claim 8, wherein the salt of a short chain fatty acid is sodium butyrate.