

Using Graph Theory to Describe and Model Chromosome Aberrations

Rainer K. Sachs,^{a,1} Javier Arsuaga,^a Mariel Vázquez,^a Lynn Hlatky^b and Philip Hahnfeldt^b

^a Department of Mathematics, University of California, Berkeley, California 94720; and

^b Dana Farber Cancer Institute, Harvard Medical School, Boston, Massachusetts 02115

Sachs, R. K., Arsuaga, J., Vázquez, M., Hlatky, L. and Hahnfeldt, P. Using Graph Theory to Describe and Model Chromosome Aberrations. *Radiat. Res.* **158**, 556–567 (2002).

A comprehensive description of chromosome aberrations is introduced that is suitable for all cytogenetic protocols (e.g. solid staining, banding, FISH, mFISH, SKY, bar coding) and for mathematical analyses. “Aberration multigraphs” systematically characterize and interrelate three basic aberration elements: (1) the initial configuration of chromosome breaks; (2) the exchange process, whose cycle structure helps to describe aberration complexity; and (3) the final configuration of rearranged chromosomes, which determines the observed pattern but may contain cryptic misrejoinings in addition. New aberration classification methods and a far-reaching generalization of mPAINT descriptors, applicable to any protocol, emerge. The difficult problem of trying to infer actual exchange processes from cytogenetically observed final patterns is analyzed using computer algorithms, adaptations of known theorems on cubic graphs, and some new graph-theoretical constructs. Results include the following: (1) For a painting protocol, unambiguously inferring the occurrence of a high-order cycle requires a corresponding number of different colors; (2) cycle structure can be computed by a simple trick directly from mPAINT descriptors if the initial configuration has no more than one break per homologue pair; and (3) higher-order cycles are more frequent than the obligate cycle structure specifies. Aberration multigraphs are a powerful new way to describe, classify and quantitatively analyze radiation-induced chromosome aberrations. They pinpoint (but do not eliminate) the problem that, with present cytogenetic techniques, one observed pattern corresponds to many possible initial configurations and exchange processes. © 2002

by Radiation Research Society

INTRODUCTION

Chromosome Aberrations

Cells exposed to ionizing radiation undergo DNA breakage and misrejoinings that lead to large-scale rearrangements of the genome; such rearrangements have a number

of important health-related implications. For acute irradiation during the G₀/G₁ phase of the cell cycle, the rearrangements typically found are chromosome aberrations [i.e. chromosome-type chromosomal aberrations, as defined in ref. (1)]. Chromosome aberration spectra help characterize DNA repair/misrepair pathways and chromosome geometry during cell cycle interphase [for a review, see ref. (2)].

A colorful diversity of radiation-induced chromosome aberrations is now observed with chromosome painting techniques, revealing a higher degree of complexity than was previously assumed (3, 4). Particularly informative are mFISH² and SKY (5–9), where all pairs of autosome homologues, as well as the sex chromosomes X and Y, are simultaneously assigned their own (pseudo-)color, so that many aberrations become readily detectable. Still further technical developments (e.g. 10–14) include “bar coding” and other techniques in which different parts of a single chromosome are assigned different colors, making intra-chromosomal rearrangements more readily detectable.

Cycles

Biophysical implications of observed aberration spectra for DNA repair/misrepair pathways and interphase chromosome geometry can be analyzed using “cycles” to describe the exchange process. A complete exchange involving just two DSBs (DNA double-strand breaks), i.e. a four-ends pairwise misrejoining as occurs for example in standard scenarios for the formation of a simple dicentric or translocation, is a “2-DSB cycle”. More generally, an “*n*-DSB cycle” characterizes an exchange process where *n* different DSBs, presumably close in time and space, have participated in a single, irreducible reaction (7, 15). In large-scale comparative genomics [for a review, see ref. (16)], a very similar concept is used (e.g. 17), referring primarily to rearrangement configurations rather than rearrangement processes [for a review, see ref. (18) Chapter 10].

Describing Aberrations

How to describe all the different aberrations appropriately is problematic. For example, there are theoretical and

¹ Address for correspondence: University of California, Mathematics Dept., Evans Hall, Berkeley, CA 94720-3840; e-mail: sachs@math.berkeley.edu.

² Abbreviations: ISCN, International System for Cytogenetic Nomenclature; mFISH, multiplex- (multifluor-, multicolor-) fluorescence *in situ* hybridization; SEC, sequential exchange complex (e.g. Fig. 6); SKY, spectral karyotyping.

experimental grounds to suppose that at the first metaphase after acute irradiation, dicentric frequency and translocation frequency are approximately equal (19), but unless one is careful with the way dicentrics and translocations are characterized, this underlying biological relationship can be masked (20). With the complicated aberration spectra that modern techniques are now making available, many similar questions arise. Descriptions based on specific biophysical models and those that are more model-neutral both have drawbacks. Importantly, aberration nomenclature and classifications currently tend to change very substantially as the experimental protocol changes, e.g. from solid staining to banding or to mFISH, even though the underlying biology is the same. The “detailed” ISCN designations for banding (18–21) and mPAINT descriptors (4, 7) have similarities, but they have not previously been unified. Recently developed computer and other quantitative modeling methods for analyzing aberration spectra³ (e.g. 22–28) need improved and more standardized ways to describe aberrations. We will argue that “aberration multigraphs” provide the solution for this.

Graph Theory and Aberration Multigraphs

Graphs have been used in various applied fields and studied mathematically for more than two centuries (29). They have been applied recently in computational biology (30), though not for studying radiogenic aberrations or using the particular type of graph theory discussed below. Figure 1 illustrates some basic definitions used throughout graph theory.

Graph theory can be applied to aberration diagrams that have been considered by Savage⁴ and similar diagrams that have been presented by various authors [e.g. ref. (7), Figs. 3 and 5; refs. (25, 31)]. By abstracting from such diagrams, one is led to aberration multigraphs, which unify three basic aberration elements: an initial configuration, an exchange process, and a final configuration. Each of these three basic elements is important in its own right (7, 32).

An aberration multigraph refers in principle to the actual biophysical process of aberration formation, as it occurred in a cell, not just to a final pattern observed cytogenetically. In any currently feasible experiment, one does not directly observe entire aberration multigraphs, and many different aberration multigraphs can correspond to one observed pattern. Modeling the underlying process explicitly by using aberration multigraphs helps organize the observations. We now show how the unification of three basic aberration elements (initial configuration, exchange process, and final configuration) into an aberration multigraph works, starting with some examples and then generalizing.

³J. Arsuaga, R. Sachs, M. Vazquez and M. N. Cornforth, Comparing DNA damage-processing pathways by computer analysis of chromosome painting data. Manuscript submitted for publication.

⁴J. R. K. Savage, private communication, 1999, and “Some thoughts on ‘Sequential Exchange Complexes’ (SEC) and the principle of exchange configuration ‘reducibility’”. Unpublished monograph, 2001.

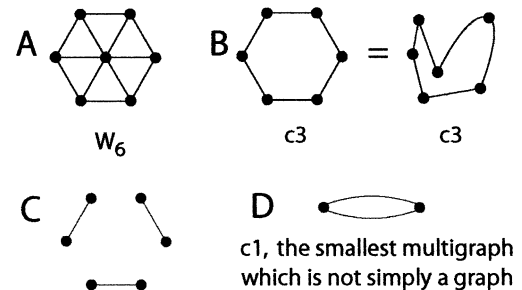


FIG. 1. Graphs and multigraphs. A *graph* consists of vertices and edges, with each edge connecting two different vertices and with at most one edge connecting any given vertex pair. For example, panel A shows a graph known as “the wheel with six spokes” and usually designated W_6 . It has seven vertices (circles) and 12 edges (lines). Panel B shows a graph called the cyclic graph with six vertices. Graph theorists write this graph as C_6 , but we designate it here as c_3 because our later treatment will emphasize DSBs rather than just vertices, with two vertices for one DSB and thus three DSBs for the six-vertex graph shown. c_3 is a *subgraph* of W_6 ; i.e., it is a graph obtained by deleting some of W_6 , namely one vertex and six edges. When discussing graphs, only the pattern of vertices and edges matters, so that in panel B the hexagon and the lopsided figure are considered as the *same* graph, i.e. as two ways to show c_3 . Panel C shows a subgraph of c_3 which is not connected; rather it has three different connected components, each consisting here of two vertices and an edge. In general a *connected component* of a graph is obtained by starting at any vertex and tracing out as much of the graph as one can without lifting pencil from paper (repetitions are allowed). *Multigraphs* are like graphs except that the restriction that at most one edge connect any given pair of vertices is dropped. Every graph is a multigraph, but not vice versa. Panel D shows an example, which we designate c_1 , of a multigraph that is not a graph. Multigraphs that are not graphs are sometimes needed, e.g. when analyzing simple ring aberrations. To include all relevant cases, we will henceforth talk primarily about multigraphs, which automatically leaves the possibility open that any particular multigraph is also simply a graph. *Submultigraphs*, and connected components for multigraphs, are defined in essentially the same way as for graphs.

METHODS

Examples of Aberration Multigraphs

Multigraphs consist of vertices and edges (Fig. 1). A vertex of an aberration multigraph (e.g. Fig. 2F) represents either a DSB free end (“free-end vertex”) or a telomere (“telomeric vertex”). An edge represents a chromosome segment (“chromosome edge”), or an initial partnership between the two free ends of a single DSB (“initial edge”), or a final partnership between DSB free ends that have been misrejoined or restituted (“final edge”). Each of the three basic elements of a chromosome aberration (initial configuration, exchange process, and final configuration) can be considered as a submultigraph, obtained from the unified aberration multigraph by deleting some parts, and having some edges in common with each of the other two basic elements. Figure 2 shows how an aberration multigraph for a simple dicentric is constructed; Fig. 3 gives the aberration multigraphs for a simple centric ring with accompanying acentric fragment, a simple translocation, and a simple inversion; Fig. 4 shows an aberration multigraph for a complex aberration.

mPAINT

Figures 2–4 use mFISH or SKY for illustrative purposes. Specifically, we take mFISH and SKY to mean 24-color whole-chromosome painting, with only color junctions, rings and centromere mismatches scored. For example, inversions and homologue-homologue translocations will both be counted here as cryptic in mFISH. The mPAINT descriptors (7) will

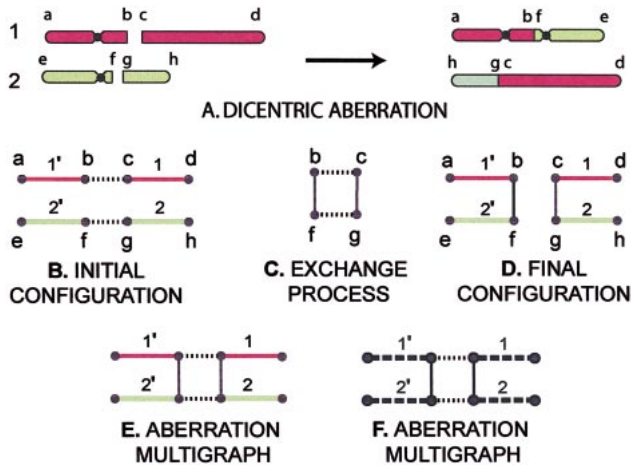


FIG. 2. A dicentric multigraph. Panel A shows the formation of a dicentric from two chromosomes, one painted red (color 1) and the other painted green (color 2). Panels B–F show relevant multigraphs (all of which are also graphs in this particular case). Panel B gives the *initial configuration*, made up of *telomeric vertices* (a, d, e and h), *DSB free-end vertices* (b and c, f and g), *initial edges* (shown dotted) specifying which free ends were originally partners, and *chromosome edges* (solid, colored lines). A prime indicates a centromere. The labels a–h are merely guides for the reader, and such vertex labels will usually not be needed. Panel C shows the *exchange process*; the vertices are the DSB free-end vertices; the edges are the initial edges (dotted lines) from panel B plus two *final edges* (solid lines), specifying that b misjoined with f, and c with g. A connected four-vertex exchange process as shown in panel C characterizes a “two-DSB cycle”, designated c2 and sometimes referred to as a four-ends pairwise exchange or a complete reciprocal exchange. Here the “2” refers to the number of DSBs involved, half the number of free-end vertices in the exchange process multigraph. Panel D shows the *final configuration*. Here the initial (dotted) edges of panel B have been removed and the final (solid) edges of panel C have been inserted instead. The two different connected components of the final configuration define the two rearranged chromosomes, the dicentric on the left and the acentric fragment on the right. The fact that there are corners here is irrelevant (compare the caption to Fig. 1B). Panel E gives the aberration multigraph, unifying panels B–D. Panel F repeats panel E in an equivalent form that is easier to draw and print because dashed lines are used for the chromosome edges instead of colors. The three basic aberration elements B–D can be read off from the multigraph F by the appropriate deletions; the mPAINT descriptors (7) can also be read off directly, as (1':: 2') (1 :: 2), by ignoring the dotted lines. Thus panel F concisely encapsulates all the information given in panels A–D about the aberration. For any aberration, the number of DSBs involved is half the number of vertices in the exchange process submultigraph, which in turn equals the number of vertices in the aberration multigraph minus the number of telomeric vertices. These relationships, and similar ones for the total number of edges, give for any aberration multigraph

$$\begin{aligned} \text{no. of vertices} &= 2(\text{no. of DSBs}) + (\text{no. of telomeres}) \\ &= (\text{no. of free-end vertices}) + (\text{no. of telomeres}). \end{aligned}$$

$$\text{no. of edges} = 3(\text{no. of DSBs}) + (1/2)(\text{no. of telomeres}).$$

An aberration multigraph will be called *simple* if two unrestituted DSBs are involved, *complex* if there are more than two.

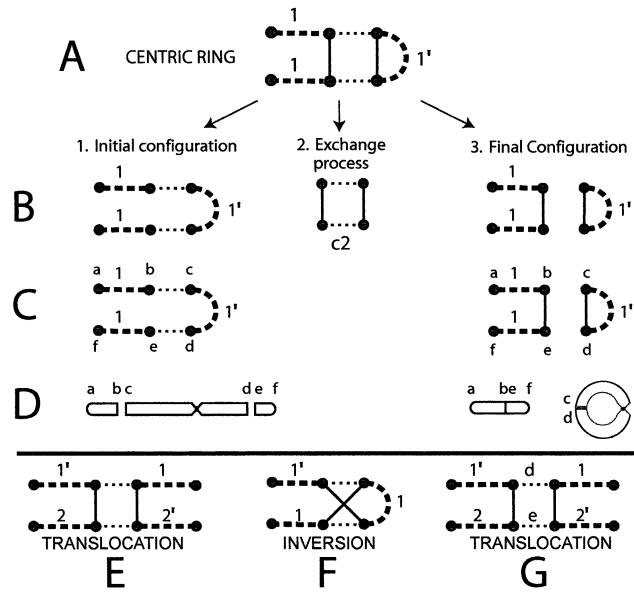


FIG. 3. Interpreting aberration multigraphs. Multigraphs give a way to see all relevant features simultaneously. For example, the multigraph in panel A can be recognized directly as involving a centric ring. Panels B–D show explicitly the steps required, but after some practice with aberration multigraphs such extra diagrams are no longer needed. The detailed interpretation of panel A is obtained as follows. (1) Ignoring the final edges (solid lines, representing misjoinings) characterizes the initial configuration as a single chromosome with one DSB on each side of the centromere (panel B1), since, as discussed in the legend to Fig. 2, the dashed lines represent chromosome segments and the dotted lines represent initial partnerships between two free ends of a DSB. The fact that the chromosome curves back in this figure is irrelevant; compare the legend to Fig. 1B. The identification as one chromosome with two DSBs is shown even more explicitly in panels C1 and D1. C1 repeats B1, adding labels for the vertices; D1 shows the chromosome edges and the DSBs of C1 more intuitively, with chromosome segments for the dashed edges, gaps for the DSBs, and a constriction for the centromere. (2) In panel A, ignoring the chromosome edges (dashed, labeled lines) and the telomeric vertices (vertices with only one incident edge) shows that the exchange process multigraph consists of a two-DSB cycle c2, as in Fig. 2. (3) In panel A, ignoring the initial edges (dotted lines) shows that the final configuration consists of two rearranged chromosomes, an acentric fragment and a centric ring (Panels B3–D3, where C3 and D3 show the misjoinings in detail). The mPAINT descriptors can be read off directly as (1) r(1'), where “r” refers to ring. Thus the aberration multigraph in panel A already contains all relevant information given in panels B–D. One can start with the intuitive picture of the aberration (panel D) and work upward to construct the multigraph in panel A (see Figs. 2 and 4), but a preferred technique is to work directly with aberration multigraphs, which can also be constructed differently. For example, aberration multigraphs are often constructed by starting with an exchange process submultigraph and then adding chromatin edges. Panels E and F give two further examples, a reciprocal translocation and a paracentric inversion, on which the reader can practice. The crossing of two edges in panel F does not represent an extra vertex, just a convenient way to draw the multigraph. Panel G gives an example of a multigraph applicable to the recombinational misrepair model. The only difference from panel E is that one initial edge is labeled “d” to indicate a DSB produced directly by radiation, and the other initial edge is labeled “e” to indicate a DSB produced by enzymatic action during a cell’s attempt to repair the first DSB. The multigraph in panel A is not simply a graph because there is a pair of vertices (namely the free-end vertices on the right) connected by two different edges (compare Fig. 1D), but the multigraphs in panels E, F and G are graphs.

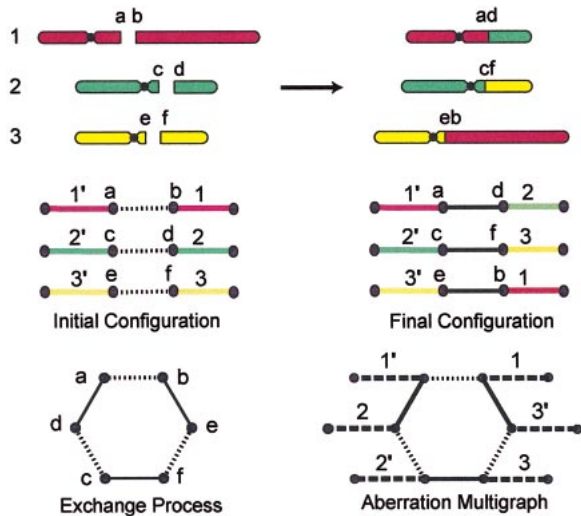


FIG. 4. The aberration multigraph for a complex, three-way aberration. The method used to construct the multigraph is essentially the same as in Fig. 2, but only a few key vertex labels are shown here for the intermediate steps (even the vertex labels shown are in principle not considered part of the multigraph). The exchange process submultigraph shows the initial edges (dotted lines) and the final (i.e. misrejoining) edges connecting appropriate free-end vertices; for example, the initial partner of a is b and the final partner of a is d. The free end vertices are arranged here in a hexagon for visual clarity, but in any case the exchange process submultigraph is the cyclic graph with six vertices (see Fig. 1B). This exchange process submultigraph characterizes the exchange process as a three-DSB cycle, c_3 . Such “three-way” or “musical chair” aberrations are among the most familiar complex aberrations. The reader can check that all relevant aberration features can be read off directly from the aberration multigraph by ignoring appropriate elements. For example, ignoring the initial edges (dotted lines) in the aberration multigraph gives the final configuration, with mPAINT descriptors $(1'::2)$ $(1::3')$ $(2'::3)$; the use of angles in drawing the rearranged chromosomes $(1'::2)$ and $(1::3')$ is again irrelevant (compare Fig. 1B). When the three basic elements of an aberration (initial configuration, exchange process and final configuration) are unified in the aberration multigraph, additional mathematical aspects, discussed in the Results section, emerge—the whole is more than just the sum of its parts.

be used and will define the *observed mPAINT pattern*. For example, the mPAINT descriptors for a simple dicentric involving chromosomes 1 and 5 together with its acentric fragment are $(1'::5')$ $(1::5)$, where the primes indicate centromeres. Other examples are given in the figure legends. We here use double colons ($::$) instead of dashes to indicate observed misrejoinings for two reasons: facilitating computer searches, and indicating similarities of mPAINT descriptors (7) to ISCN designations for aberrations scored with banding (18).

Completeness

For the time being, we assume that all exchanges are complete and that all chromosome edges are long enough to be observable. If instead there are chromosome segments so short they are not seen in the assay, a different mathematical situation arises. In that case, if there are enough short segments, every observed aberration pattern is consistent with an exchange process consisting entirely of 2-DSB cycles, as can be shown by generalizing an argument given earlier (15); a quite elegant mathematical theory is available for determining how many 2-DSB cycles are needed to produce a given observed pattern (21). These results are not directly relevant when all chromosome edges involved are long enough to be observable, and they will not play any further role in the analysis below.

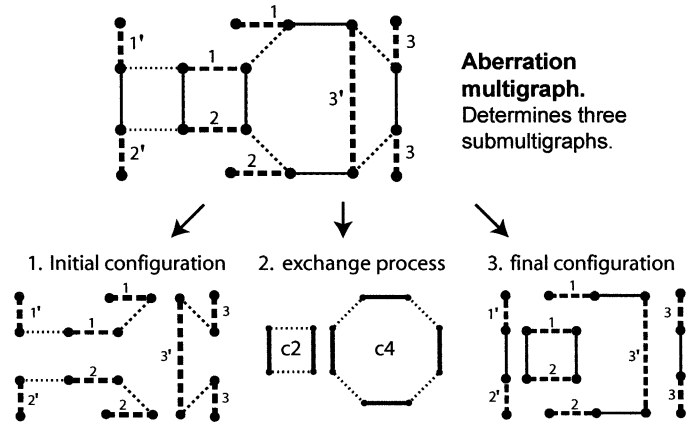


FIG. 5. A complex aberration involving six DSBs. The aberration multigraph gives the following properties for the three basic elements of the aberration: (1) The initial configuration, obtained by ignoring the final edges (solid lines), shows that originally there were three different chromosomes, with three different colors, each having two DSBs, for a total of six DSBs in all. The fact that there are various strange corners is not relevant (compare Fig. 1B). (2) The exchange process is obtained by ignoring the chromosome edges (labeled dashed lines) and ignoring the telomeric vertices (vertices with only one incident edge). The exchange process cycle structure here is c_2+c_4 , confirming that there were six DSBs. The four-DSB cycle suggests a situation where four DSBs were nearby in space and time and their free ends misrejoined in an irreducible interaction. When the exchange process has more than one connected component, as here where there are two connected components, one has, by definition, a sequential exchange complex (SEC). For any SEC there are several different reactions, here c_2 and c_4 , counted as part of the same aberration because they involve shared chromosomes (dashed, labeled edges in the aberration multigraph). (3) Final configurations are obtained by ignoring the initial edges (dotted lines) in an aberration multigraph. Here the final configuration has mPAINT descriptors (7) $(1'::2')$ $r(1::2)$ $(1::3'::2)$ (3), where “r” again refers to ring, describing the four rearranged chromosomes shown. An aberration involving as many as six DSBs is complicated, no matter what description method is used. In this case, as in all complicated cases, the aberration multigraph (top) unifies the relevant information as simply as possible, at the cost of being somewhat abstract.

Defining SECs

The aberration multigraph for the entire altered karyotype in one cell may have more than one connected component (connected components are defined in the legend to Fig. 1). In that case, we are dealing with several different aberrations that are presumably independent of each other. Formally, we can define an individual aberration as a connected aberration multigraph. Even for a connected aberration multigraph, the exchange process submultigraph may have more than one connected component (Fig. 5 shows an example). Having more than one component for the exchange process then identifies the aberration as a sequential exchange complex (SEC). This mathematical definition captures the current biology usage [for a review, see ref. (33)] even in complicated situations.

Biophysical Models

Possible biophysical pathways for producing chromosome aberrations are breakage and reunion, based on nonhomologous end joining, recombinational misrepair, and the exchange theory pathway (34). The examples in Figs. 2–4 are easily interpreted using a breakage-and-reunion scenario. However, aberration multigraphs can also be used to analyze recombinational misrepair scenarios as follows: (1) restricting attention to those multigraphs where each connected component of the exchange pro-

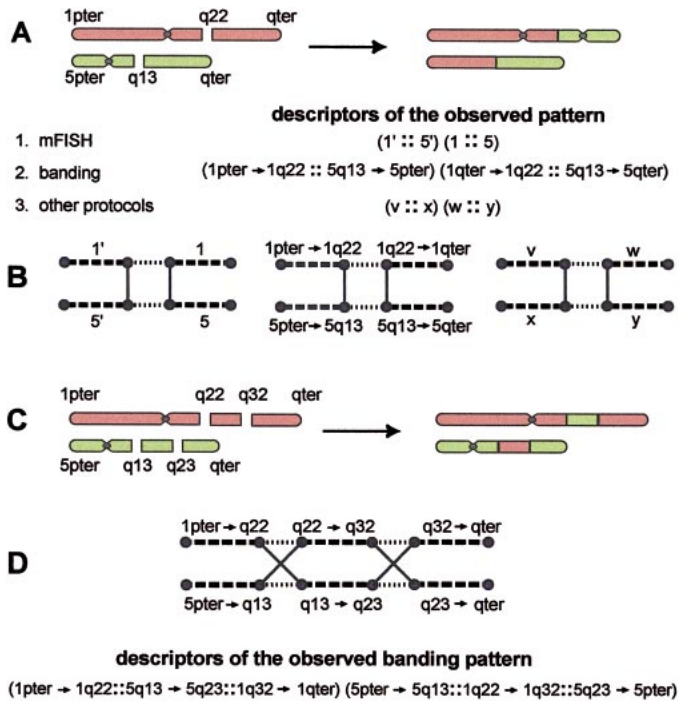


FIG. 7. Examples of nomenclature. Panel A shows an intuitive picture of the formation of a simple dicentric with breakpoints at 1q22 (band 22 in the long arm of chromosome 1) and 5q13; pter and qter designate the ends of the short and long arms, respectively. The mPAINT descriptors are given in A1. A2 gives the suggested descriptors if the protocol is banding instead. These suggested descriptors are similar to, but not identical to, “detailed” ISCN designations (21). Some of the differences between the suggested descriptors and the ISCN designations are designed to improve computer searchability in databases; others are suggested by the need for a general nomenclature that is applicable not just to banding but also to solid staining, mFISH/mPAINT, bar coding, etc. In general, for any protocol in which neither misjoining is cryptic, the descriptors would have the form given in A3. For example, $v = 1'$ in the mPAINT case, $v = 1\text{pter} \rightarrow 1\text{q22}$ in the banding case, and for other protocols v could have other forms, in principle even an entire DNA sequence, as indicated by some further examples in the text. For some protocols the misjoinings could be cryptic. For example, suppose chromosomes 1 and 5 are both counterstained in an experiment with some other chromosomes painted by FISH, and centromeres on counterstained chromosomes are not scored. Then the entire aberration is cryptic for this protocol and no descriptors apply. Panel B shows corresponding aberration multigraphs. The actual multigraph, given by vertices and edges, is the same for all protocols, even including protocols which make the dicentric cryptic; only the labels for the chromosome edges are different for different protocols. The descriptors in panel A are obtained from the final configuration by following the chromosome and misjoining edges (dashed and solid edges) of each rearranged chromosome. In some protocols, such as banding, the chromosome edge labels on the multigraph have a direction, and this direction is then taken into account when going from the multigraph to the descriptors. For example, “1q22 → qter” labeling a chromosome edge of the multigraph here becomes “1qter → 1q22” in the descriptor, where the “1” in 1q22, implicit in the edge label and in detailed ISCN, is made explicit in the descriptor for the sake of computer searchability. Panel C shows an intuitive picture for another example. Panel D gives the corresponding multigraph and gives the observed pattern in a banding protocol, which can be read off the multigraph by following the dashed and solid lines. For an mFISH protocol, this same aberration would have observed pattern (1' :: 5 :: 1) (5' :: 1 :: 5). For any protocol in which none of the four misjoinings is cryptic, the observed pattern has the form (u :: v :: w) (x :: y :: z) illustrated by the

string for each rearranged chromosome (i.e. for each connected component of the final configuration submultigraph of the aberration multigraph; compare Fig. 6). If the labels have a direction, as in the examples given above, the direction is taken into account appropriately in the descriptors of the observed pattern. Figure 7 gives two examples for the case of a banding protocol.

This generalization of mPAINT should suffice for the foreseeable future, no matter what advances are made in aberration detection techniques. It allows considerably more systematic comparisons among different kinds of experiments or among different types of scoring within one experiment than was possible previously. It is compatible with the notation used for global comparisons of gene order in modern comparative genomics (16, 18).

Cycles

The *cycle structure* (7, 15, 17) of an aberration can be defined mathematically in terms of aberration multigraphs as follows. Consider the exchange process, a submultigraph of the aberration multigraph, obtained by deleting the chromosome edges and the telomeric vertices (Figs. 2–6). In general this exchange process can have several different connected components, corresponding to several different reactions that could be separated in time and/or space. A theorem in graph theory³ (18, 29) shows that for an aberration multigraph each connected component of the exchange process must have an even number of vertices in a cyclic arrangement. The decomposition of the exchange submultigraph into connected components thus defines the cycle structure in the way exemplified by Figs. 2–5. Specifically, the cycle structure for each aberration multigraph in Figs. 2 and 3 is c_2 , that in Fig. 4 is c_3 , and that in Fig. 5 is c_2+c_4 . A cycle structure c_5 would correspond to an exchange process that is a cyclic graph with 10 vertices (for five DSBs); cycle structure c_1 would correspond to a single restitution (one DSB); cycle structure $c_2+c_3+c_4$ describes an aberration which starts from nine DSBs and involves three separate reactions that have some chromosomes in common; etc.

Classifying Aberrations

There are so many different kinds of aberrations that it has long been common to classify them into groups. Unlike primarily descriptive nomenclature such as mPAINT (or its generalization suggested above), and in parallel with such nomenclature, classifications have sometimes been based on “mechanistic” considerations. We next discuss classification methods that are mechanistic in the sense that they

←

banding and mPAINT examples, where the six edge labels u, . . . , z are appropriate to the particular protocol; the aberration multigraph has the same vertices and edges as in panel D no matter what protocol is used.

TABLE 1
Examples of Diameter and Girth

Aberration	Figure	No.	DSBs	Diameter	Girth
Ring	3A	1	2	3	2
Inversion	3F	1	2	4	3
Dicentric	2F	2	2	4	4
Translocation	3E	2	2	4	4
Three-way	4	3	3	5	6
Any c2+c2	—	1–2	4	4–7	2–4
Any c4	—	1–4	4	4–9	2–5 or 8
Any c2+c3	—	1–3	5	4–7	2–4
Star-fighter	8A	3	5	6	4
c5	8B	3	5	6	5
c2+c4	5	3	6	8	4

Notes. This is how a typical graph-theory classification, using diameter and girth, looks when it is applied to aberration multigraphs. No. is the number of chromosomes involved and diameter is the multigraph diameter. The reader is asked to check the assignments of the multigraph diameter and multigraph girth in those cases where a figure is referenced. In all cases it is assumed that the initial configuration consists of linear, rather than circular, chromosomes. Often, girth 2 corresponds to a ring being one of the rearranged chromosomes.

can use an entire aberration multigraph, whereas the descriptive nomenclature suggested above uses only the non-cryptic aspects of the final configuration submultigraph (Fig. 6).

Aberrations have often been classified according to the number of chromosomes involved (i.e. half the number of telomeric vertices in the aberration multigraph) and the number of DSBs involved (i.e. half the number of free-end vertices in the multigraph) (37). Recently (7, 15) they have also been classified according to the cycle structure (given by the exchange process submultigraph of the aberration multigraph). Graph theory makes available additional ways to classify aberrations.

Simple examples of quantities that could be used to help classify aberrations, often used by graph theorists, are “diameter” and “girth” (29). Table 1 shows values of these particular quantities for some aberrations. The *diameter* of an aberration multigraph is the longest distance between any pair of vertices, where the *distance* between two given vertices is defined as the number of edges in a shortest connected path between the vertices. High diameter for a given number of DSBs corresponds to many different cycles in the exchange process, linked by the fact that they have some chromosomes in common. Thus, serendipitously, an aberration whose formation was spread out in space and/or time tends to have a higher diameter, in the graph-theory sense, for its multigraph than does an aberration involving the same number of DSBs in a tighter reaction. The *girth* of an aberration multigraph is the number of edges in a shortest cyclic submultigraph. Girth 2 usually corresponds to ring formation and girth 3 to inversions. Large girth can occur only if every reaction of the exchange process is complicated.

Diameter and girth are used throughout graph theory, but special properties of aberration multigraphs also make a

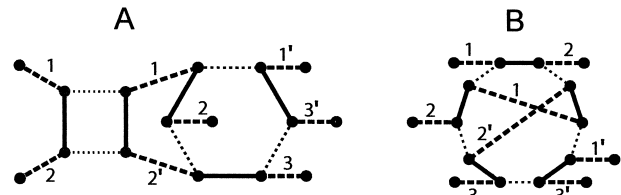


FIG. 8. An ambiguity. Each of the two panels shows an aberration multigraph. In each panel the initial configuration can be read off by following the dashed and dotted lines. The initial configuration is the same in both panels, as follows: Chromosome 1 has two breaks, the centromere being on an end segment; chromosome 2 has two breaks, the centromere being on the middle segment; and chromosome 3 has one break. The final configuration multigraphs can be read off by following the solid and dashed lines and are the same in both panels; the mPAINT descriptors are $(1 :: 2) (2 :: 1 :: 2' :: 3) (1' :: 3')$, and there is no cryptic damage. Nevertheless, the aberrations are different, as can be seen from the dotted and solid lines: The exchange process for the “starfighter” multigraph in A has cycle structure $c2+c3$ whereas the exchange process in B has cycle structure $c5$.

somewhat deeper classification possible. All DSB free-end vertices of an aberration multigraph have three incident edges (compare Figs. 2–5), so aberration multigraphs are closely related to what are called *cubic* multigraphs, defined as multigraphs where every vertex has three incident edges. An enormous amount is known about cubic multigraphs (29), mainly because they happen to be related to the famous four-color problem for maps. Results on cubic multigraphs suggest a mathematical classification of aberration multigraphs, as will be described elsewhere.

Cryptic Damage and Ambiguities

The connected components of the final configuration submultigraph represent the rearranged chromosomes (Figs. 2–6). One of the main problems in quantitative aberration analyses is that often the rearranged chromosomes have cryptic damage, where the experimental protocol does not uncover a misjoining. An example is given by the paracentric inversion in Fig. 3F. The misjoinings will be cryptic in an mFISH/mPAINT experiment, though a banding or bar coding experiment might uncover them. Typically, a more sophisticated experimental protocol decreases such ambiguities but does not eliminate them. Thus the observed pattern does not determine the final configuration multigraph of an aberration multigraph uniquely (given a specific protocol, the converse does hold, as represented in Fig. 6 by an arrow *from* the final configuration multigraph *to* the observed pattern).

Even when the final configuration is completely known, the initial configuration and/or the exchange process (i.e. the other two basic submultigraphs of the aberration multigraph) are not in general uniquely determined. Figure 8 gives an example. In both panels the final configuration is the same. In this particular case the initial configurations are also the same in both panels. Thus the two aberrations involve the same number of chromosomes and DSBs, as well as having

the same observed final mFISH/mPAINT pattern. Nonetheless, the two aberrations shown are different, as shown by the fact that the cycle structures of the exchange processes differ, being $c2+c3$ in Fig. 8A and $c5$ in Fig. 8B. The difference could, at least in principle, be uncovered by other kinds of experiments, e.g. premature chromosome condensation experiments.

Since different aberrations can correspond to the same final configuration multigraph and different final configurations can correspond to the same observed final pattern, the number of different aberrations giving a particular observed pattern can be large. We next consider some methods of coping with this ambiguity, trying to determine, or at least give probabilistic estimates for, the actual damage processing events from the observable patterns.

Obligate and Probabilistic Cycle Structures

Because aberration exchange processes are informative about enzymatic DNA repair/misrepair³ (27, 34–36), attention has recently been focused on determining cycle structures. In cases where an observed final pattern defined by mPAINT descriptors is consistent with different possible cycle structures for the exchange process, Cornforth introduced the concept of *obligate* cycle structure, essentially as that consistent cycle structure which maximizes the number of cycles involved³ (7); the concept is closely related to the concept of maximal cycle decomposition (18) in comparative genomic analyses. For example, the obligate cycle structure in Fig. 8 is the cycle structure $c2+c3$ of the aberration shown in panel A, since the alternative method (panel B) of making the given mPAINT pattern has cycle structure $c5$. The obligate cycle structure is considered the most conservative possibility for generating a given observed mPAINT pattern, being closest to the older view that all, or at least most, complete exchange processes are simply composed of one or more 2-DSB cycles, with higher-order cycles rare or altogether absent.

However, the obligate cycle structure is not always the one that occurs in aberration radiogenesis, so probabilistic analyses are needed (15). Specifically, suppose we are given an observed mPAINT pattern. We can first carry out the following four steps³ (7) to assign configurations:

1. Count telomeres in the observed pattern, to find out how many chromosomes were involved.
2. By analyzing color junctions and cases of two contiguous centromeres involving a single color, infer the minimum number of DSBs that could have given rise to the observed pattern.
3. Assuming the minimum break number, deduce the final configuration.
4. Work out, for purposes of estimating the cycle structure, initial configurations which are consistent with that final configuration and have the minimum break number.

For example, in Figs. 2, 3A, 3E, 4, 5 and 7, the mPAINT

descriptors and this procedure lead to the initial and final configurations given in the figures. Figure 8 shows that even given the initial and final configurations, the cycle structure is not necessarily determined uniquely. However, at this point one can determine all the different ways the given final configuration can be generated from the given initial configuration by different exchange processes involving various possible cycle structures, and probabilities can thereby be assigned to the different cycle structures.³ Here are some examples that we have tried:

1. The initial and final configurations of Fig. 8, involving five DSBs, gave $c2+c3$ (obligate cycle structure, two cases out of eight for a probability of 1/4) compared to $c5$ (six cases out of eight, probability 3/4).
2. An example involving four DSBs gave $c2+c2$ (obligate cycle structure, two cases out of four for a probability of 1/2), compared to $c4$ (two cases out of four for a probability of 1/2).
3. A six-DSB example gave $c2+c2+c2$ (obligate cycle structure, probability 1/4), compared to $c2+c4$ (probability 1/2), compared to $c6$ (probability 1/4).

Thus, in each case, there is a substantial probability of a cycle structure with cycles of orders higher than that given by the obligate cycle structure. This result was pointed out by Savage,⁴ who analyzed a number of examples in detail.

In an actual experiment, there will be not only a probabilistic mixture of cycle structures given the initial and final configurations, but also a probabilistic mixture of configurations given the observed mPAINT pattern. At low doses, there will be a bias toward configurations having the minimum number of DSBs consistent with the observed pattern, as assumed in the simple probabilistic approach just described. For higher doses, aberrations with more DSBs than the minimum consistent number may play an important role, and more extensive probabilistic calculations are required, using Monte Carlo simulations (e.g. 15, 38). Then attention shifts from individual aberrations to probabilistic estimates for the entire observed aberration spectrum.

An n-Color Theorem

In contrast to the analysis just discussed, where the objective is to estimate frequencies of particular kinds of exchange processes, one can also ask if the existence of higher-order cycles can be deduced unambiguously, at least in principle, from observed patterns. It used to be implicitly assumed in many papers that all exchange processes have cycle structure $c2+c2+\dots+c2$ (i.e. consist exclusively of four-end pairwise exchanges). However, by using three-color painting (two FISH colors and a counterstain), it was possible to show that some $c3$ exchange processes occur (39).

Aberration multigraph analysis has shown that this situation is more general. For example, for aberrations involving three FISH colors (and no counterstain), we can un-

ambiguously identify three-DSB cycles but not four-DSB cycles (this statement is contingent on the condition that no chromosome segments too short to be detected are involved; whether centromeres are scored or not is irrelevant in the present argument). Still more generally, a given observed pattern whose mPAINT descriptors involve n colors can never unambiguously identify the occurrence of an m -DSB cycle with $m > n$. If we assume enough DSBs, we can always produce any n -color pattern observed using at worst n -DSB cycles. Stated more formally, we have the following.

Theorem 1. Consider mPAINT observed patterns involving n colors, with $n \geq 2$.

a. There are observed patterns such that any exchange process producing the observed pattern involves at least one n -DSB cycle.

b. Given any observed pattern, there is an exchange process producing that pattern which involves only m -DSB cycle(s) with $m \leq n$.

The proof of the theorem, generalizing a proof given earlier of a special case (15), involves modifying multigraphs by adding extra DSBs. The proof is rather long and will be given elsewhere. In lieu of the proof, we give examples here for both part a and part b.

For part a, consider the aberration of Fig. 4. The mPAINT descriptors for the observed pattern are $(1':: 2)$ $(1 :: 3')$ $(2':: 3)$. Thus we are considering $n = 3$ in part a of the theorem. There is an aberration multigraph involving a $c3$ that can produce this observed pattern, namely just the aberration multigraph in Fig. 4. Are there any aberration multigraphs which give this observed pattern but have an exchange process submultigraph involving only two-DSB cycles (possibly many $c2$ s in some clever combination)? The answer is “no”. In the observed pattern the number of color junctions between color 1 and color 3 is an odd number, namely 1. But every $c2$ makes either zero or two color junctions between color 1 and color 3. Consequently, making the observed pattern with mPAINT descriptors $(1':: 2)$ $(1 :: 3')$ $(2':: 3)$ requires at least one cycle involving at least three DSBs. Thus part a of the theorem has been shown to be true at least in case $n = 3$.

To get an example for part b, suppose $n = 2$ and suppose the observed pattern has mPAINT descriptors $(1':: 2')$ $(1 :: 2')$ (2) . The observed pattern is of a familiar kind (32, 37) and is commonly considered to imply a $c3$ (Fig. 9A and B). However, this same observed pattern is also generated by a different aberration (Fig. 9C) whose exchange structure is $c2+c2$ (Fig. 9D). The trick here is to use four DSBs instead of just three DSBs, with one misrejoining cryptic; by using an extra DSB, the maximum number of DSBs in the cycles is reduced from three to two. Figure 9 shows that part b of the theorem is true at least in the special case where $n = 2$ and the observed pattern is $(1':: 2')$ $(1 :: 2')$ (2) . Notice here that the aberration of Fig. 9C need not involve chromosome segments too short to observe. As dis-

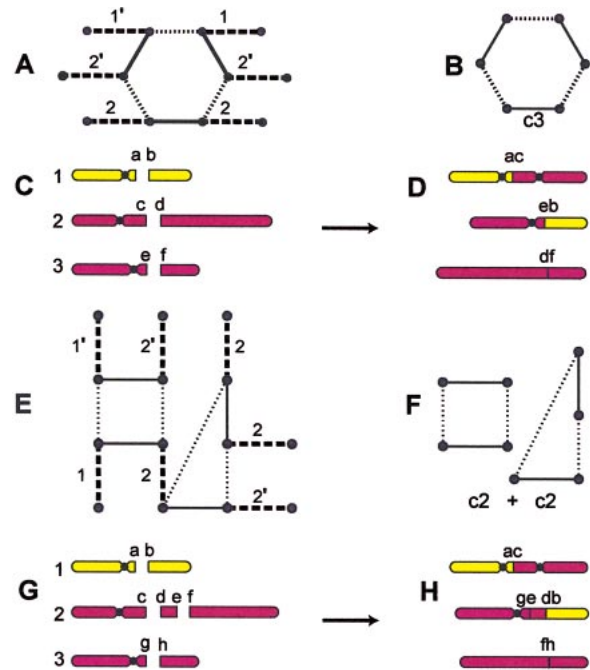


FIG. 9. Multigraphs illustrating part b of theorem 1. The figure shows two different aberrations; they have the same observed mFISH pattern, with mPAINT descriptors $(1':: 2')$ $(1 :: 2')$ (2) . The multigraph in panel A has only three DSBs but there is a three-DSB cycle (panel B). In panel E, four DSBs are involved, but only two-DSB cycles occur (panel F). The information of panels A and E is repeated in pictorial form by panels C, D, G and H. Panels A and E by themselves would be enough to give all the relevant information.

cussed earlier, cryptically short chromosome segments correspond to a mathematically different situation.

The Telomere Removal Trick

As we have discussed, it is generally hard to infer from an observed pattern the aberration cycle structure (which is the feature of an aberration most directly informative about enzymatic misrejoining pathways or interphase chromosome geometry), and the results can be ambiguous. If we somehow have some extra information on the initial configuration, the situation is sometimes better. Here are some examples.

Suppose, in an mFISH/mPAINT experiment, we know that the initial configuration has no color with more than one break. Then the cycle structure of the exchange process can be obtained directly from the mPAINT descriptors by what we will call the *telomere removal trick*,⁵ which involves hypothetical fusions of pter with qter segments. For example, in the final configuration of Fig. 4, take the two red chromosome segments, both of which terminate in telomeres, and join them into a single segment, mentally removing the two telomeres. Proceed similarly for the green

⁵ The telomere removal trick was suggested by the methods Radivoyevitch introduced for analyzing telomeres and DSBs simultaneously in experiments designed to measure length distributions for DNA fragments.

and yellow segments. The result is a big ring with three color junctions. This hypothetical ring, although constructed directly from the observed pattern without using the rest of the aberration multigraph, correctly identifies the exchange process as a c3. In terms of mPAINT descriptors, the procedure is the following. We start with the descriptors $(2':: 1) (1':: 3') (3 :: 2)$ of Fig. 4. Fusing the telomeres of color 1 now gives $(2':: 1':: 3') (3 :: 2)$; fusing the telomeres of color 3 gives $(2':: 1':: 3':: 2)$; and, finally, fusing the telomeres of color 2 gives $r(2':: 1':: 3')$, where “r” refers to ring. This mentally constructed ring with three color junctions identifies the exchange process as a c3. Had we started by removing telomeres of color 2 or 3 instead of color 1 the same final conclusion would have been reached. The reader can check that essentially the same procedure works in Figs. 2 and 3B, giving cycle structure c2 in both cases. In fact, the following fairly general result holds.

Theorem 2. In an mFISH/mPAINT experiment suppose no color initially has more than one DSB; then the cycle structure of the exchange process is given directly by the telomere removal trick applied to the mPAINT descriptors.

The proof of the theorem uses the fact that when telomeres are removed, one has a cubic multigraph (i.e. every vertex has three incident edges, as discussed in a different context above). This cubic multigraph structure leads to a remarkable symmetry: There is no longer, in an aberration multigraph with telomeres removed, any essential mathematical distinction between initial edges, misrejoining edges, and final edges. This unusual form of symmetry allows one to prove a number of unexpected results, including theorem 2. However, the argument is again lengthy and will be given elsewhere. In its place, we add two comments.

1. Similar but more complicated results can be proven for cases where a few colors have two DSBs rather than just one.
2. For fairly low doses, the theorem’s presupposition that all colors initially have one DSB at most tends to hold, because if there are not many DSBs, the chance that there is more than one on any one color is small. However, in general there could be a more complicated initial situation, so the telomere removal trick is only a partial answer to the problem of inferring exchange processes from observed final patterns, even assuming an mFISH/mPAINT protocol.

Incompleteness

Two kinds of incompleteness can be incorporated into an aberration multigraph. When true incompleteness occurs, some DSB free ends have no final partners; i.e., some of the final edges in the aberration multigraph are missing. Apparent incompleteness, caused by chromosome segments too short to observe, can be incorporated simply by giving the corresponding chromosome edges of the aberration multigraph an additional label (e.g. “s” for short). Incompleteness in general can then be defined as a situation where

a free-end vertex has no detectable final edge incident on it. These possibilities give a substantially more natural way to describe incompleteness than has hitherto been available.

DISCUSSION

We introduced aberration multigraphs, which give the most systematic method to date for analyzing radiation-induced chromosome aberrations. Familiar aspects of aberrations, such as initial or final configurations, exchange processes, SECs, cycle structure, rearranged chromosomes, incompleteness, etc. re-emerge automatically as mathematical aspects of aberration multigraphs (Fig. 6). Using the multigraphs allows drawing some comparatively subtle but essential distinctions; for example, there is a difference between the final configuration of rearranged chromosomes, as it actually occurs in a cell, and the observed pattern, which usually characterizes some aspects of the final configuration but not cryptic aspects (Figs. 3F and 6). We here emphasized graph-theoretical results relevant to the hard problem of trying to work out actual aberration exchange processes, which are the focus of interest biologically, from observed aberration patterns. Two theorems bearing on the problem, and also probabilistic estimates, were discussed.

Aberration multigraphs suggested a generalization of mPAINT descriptors flexible enough to include all current cytogenetic scoring methods for exchange-type G_0/G_1 -phase rearrangements and, as far as can be foreseen, future methods, including in principle even full DNA sequencing. The generalization, given in the Results section above, can be restated very simply, avoiding any reference to multigraphs. For a rearranged chromosome that is not a ring, the descriptor is obtained by the following steps: (1) Start at either end and write open parentheses, “(”, to represent the end; (2) describe the chromosome segment to the first recognizable breakpoint as explicitly as is possible given whatever cytogenetic protocol is being used (e.g. banding or mFISH); (3) write “::” to represent a detected misrejoining; (4) describe the next chromosome segment as explicitly as possible; (5) continue until the other end, represented by “)”, is reached. For a ring, an analogous recipe holds.

Descriptors such as mPAINT descriptors or their generalizations, being determined by observed patterns, are not by themselves enough to classify aberrations fully. The present approach allows a sharp distinction between nomenclature, which depends only on an observed pattern, and “mechanistic” classifications, based on a multigraph as a whole. Mechanistic classifications involve various well-known mathematical concepts, for example the diameter of a multigraph.

Examples showed that using aberration multigraphs is more informative than studying separately the three basic aberration elements (initial configuration, exchange process, and final configuration; see Fig. 6). One main reason is that, serendipitously, aberration multigraphs are very closely related to cubic multigraphs, which happen to have

received an unusually large amount of attention from mathematicians. In fact, by using the telomere removal trick discussed in the Results section, an aberration multigraph can be made into a cubic multigraph. Some normal genomes, and human mitochondrial DNA, consist of one or more circular chromosomes. Such a genome is especially amenable to aberration multigraph analysis because its aberration multigraphs are automatically cubic multigraphs, and the telomere removal trick, whose use can involve some loss of information, is never needed.

Graph theory has recently been applied to various aspects of computational biology (30), but the present paper's use of graph theory to analyze chromosome aberrations and its use of cubic multigraphs are both new. Compared to the analyses used in modern large-scale comparative genomics (16), a key difference is that aberration multigraphs specify the overall process of aberration formation, not just the final configuration. A unique cycle structure is associated with the process, whereas in the comparative genomics case, cycle decompositions can be ambiguous (18). The radiation cytogenetics approach suggests that, in comparative genomics, searches in DNA sequences for details on breakpoints should be undertaken. Zooming in on breakpoints manifest in coarse-grained assays such as Zoo-FISH can in principle determine cycle structures, breakpoint multiplets due to clustering or consecutive nearby breaks, and, for rearrangements made by ionizing radiation, characteristic features of particular kinds of radiation.

To summarize, we argue that in the long run a graph-theoretical approach will give the best way to describe aberrations, to analyze them mathematically, and to compare aberration data obtained using different protocols.

ACKNOWLEDGMENTS

We are grateful to Drs. M. Cornforth and J. R. K. Savage and to M. Chan for discussions and corrections. This research was supported by NSF grant DMS 9971169 (MV and JA), NSF grant DMS 9896163 (PH), NSF grant DBI 9904842 and NIH grant CA86823 (LH), NIH grant GM 57245 and grant DE-FG03-00-ER62909 in the Low Dose Radiation Research Program, Biological and Environmental Research (BER), U.S. Department of Energy (RKS).

Received: May 13, 2002; accepted: June 19, 2002

REFERENCES

1. J. R. K. Savage, Classification and relationships of induced chromosomal structural changes. *J. Med. Genet.* **13**, 103–122 (1976).
2. L. Hlatky, R. Sachs, M. Vazquez and M. Cornforth, Radiation-induced chromosome aberrations: insights gained from biophysical modeling. *Bioessays*, in press.
3. J. R. K. Savage, Enhanced perspective: Proximity matters. *Science* **290**, 62–63 (2000).
4. B. D. Loucas and M. N. Cornforth, Complex chromosome exchanges induced by gamma rays in human lymphocytes: An mFISH study. *Radiat. Res.* **155**, 660–671 (2001).
5. K. M. Greulich, L. Kreja, B. Heinze, A. P. Rhein, H-U. G. Weier, M. Briickner, P. Fuchs and M. Molls, Rapid detection of radiation-induced chromosomal aberrations in lymphocytes and hematopoietic progenitor cells by mFISH. *Mutat. Res.* **452**, 73–81 (2000).
6. D. O. Ferguson and F. W. Alt, DNA double strand break repair and chromosomal translocation: Lessons from animal models. *Oncogene* **20**, 5572–5579 (2001).
7. M. N. Cornforth, Analyzing radiation-induced complex chromosome rearrangements by combinatorial painting. *Radiat. Res.* **155**, 643–659 (2001).
8. E. Schrock and H. Padilla-Nash, Spectral karyotyping and multicolor fluorescence *in situ* hybridization reveal new tumor-specific chromosomal aberrations. *Semin. Hematol.* **37**, 334–347 (2000).
9. M. R. Speicher, S. G. Ballard and D. C. Ward, Karyotyping human chromosomes by combinatorial multi-fluor FISH. *Nat. Genet.* **12**, 368–375 (1996).
10. C. Johannes, I. Chudoba and G. Obe, Analysis of X-ray-induced aberrations in human chromosome 5 using high-resolution multicolor banding FISH (mBAND). *Chromosome Res.* **7**, 625–633 (1999).
11. J. Wiegant, V. Bezrookove, C. Rosenberg, H. J. Tanke, A. K. Raap, H. Zhang, M. Bittner, J. M. Trent and P. Meltzer, Differentially painting human chromosome arms with combined binary ratio-labeling fluorescence *in situ* hybridization. *Genome Res.* **10**, 861–865 (2000).
12. K. Saracoglu, J. Brown, L. Kearney, S. Uhrig, J. Azofeifa, C. Fauth, M. R. Speicher and R. Eils, New concepts to improve resolution and sensitivity of molecular cytogenetic diagnostics by multicolor fluorescence *in situ* hybridization. *Cytometry* **44**, 7–15 (2001).
13. R. Karhu, M. Ahlstedt-Soini, M. Bittner, P. Meltzer, J. M. Trent and J. J. Isola, Chromosome arm-specific multicolor FISH. *Genes Chromosomes Cancer* **30**, 105–109 (2001).
14. C. Fauth and M. R. Speicher, Classifying by colors: FISH-based genome analysis. *Cytogenet. Cell Genet.* **93**, 1–10 (2001).
15. R. K. Sachs, A. M. Chen, P. J. Simpson, L. R. Hlatky, P. Hahnfeldt and J. R. Savage, Clustering of radiation-produced breaks along chromosomes: Modelling the effects on chromosome aberrations. *Int. J. Radiat. Biol.* **75**, 657–672 (1999).
16. D. Sankoff and J. H. Nadeau, *Comparative Genomics*. Kluwer Academic, Boston, 2000.
17. V. Bafna and P. A. Pevzner, Genome rearrangements and sorting by reversals. *Siam. J. Comput.* **25**, 272–289 (1996).
18. P. Pevzner, *Computational Molecular Biology: An Algorithmic Approach*. MIT Press, Cambridge, MA, 2000.
19. J. R. K. Savage and D. G. Papworth, Frequency and distribution studies of asymmetrical versus symmetrical chromosome aberrations. *Mutat. Res.* **95**, 7–18 (1982).
20. J. N. Lucas, A. M. Chen and R. K. Sachs, Theoretical predictions on the equality of radiation produced dicentrics and translocations detected by chromosome painting. *Int. J. Radiat. Biol.* **69**, 145–153 (1996).
21. F. Mitelman, *ISCN 1995: An International System for Human Cytogenetic Nomenclature*. S. Karger, New York, 1995.
22. D. Brenner, Track structure, lesion development, and cell survival. *Radiat. Res.* **124** (Suppl.), S29–S37 (1990).
23. A. M. Chen, J. N. Lucas, F. S. Hill, D. J. Brenner and R. K. Sachs, Chromosome aberrations produced by ionizing radiation: Monte Carlo simulations and chromosome painting data. *Comput. Appl. Biosci.* **11**, 389–397 (1995).
24. A. A. Edwards, Modeling radiation-induced chromosome aberrations. *Int. J. Radiat. Biol.* **78**, 000–000 (2002).
25. W. R. Holley, L. S. Mian, S. J. Park, B. Rydberg and A. Chatterjee, A model for interphase chromosomes and evaluation of radiation induced aberrations. *Radiat. Res.* **158**, 568–580 (2002).
26. J. Y. Ostashevsky, Higher-order structure of interphase chromosomes and radiation-induced chromosomal exchange aberrations. *Int. J. Radiat. Biol.* **76**, 1179–1187 (2000).
27. M. Vázquez, K. Greulich-Bode, J. Arsuaga, M. Cornforth, M. Brckner, R. Sachs, P. Hahnfeldt, M. Molls and L. Hlatky, Computer analysis of mFISH chromosome aberration data uncovers an excess of very complex metaphases. *Int. J. Radiat. Biol.*, in press.

28. A. Ottolenghi, F. Ballarini and M. Biaggi, Modelling chromosomal aberration induction by ionising radiation: The influence of inter-phase chromosome architecture. *Adv. Space Res.* **27**, 369–382 (2001).
29. N. Hartsfield and G. Ringel, *Pearls in Graph Theory: A Comprehensive Introduction*. Academic Press, San Diego, 1994.
30. M. Kanehisa, *Post-Genome Informatics*. Oxford University Press, Oxford and New York, 2000.
31. R. M. Anderson, D. L. Stevens and D. T. Goodhead, M-FISH analysis shows alpha-particle induced complex chromosome rearrangements are cumulative products of localized rearrangements. *Proc. Natl. Acad. Sci. USA*, in press.
32. J. R. K. Savage and P. J. Simpson, FISH “painting” patterns resulting from complex exchanges. *Mutat. Res.* **312**, 51–60 (1994).
33. A. A. Edwards and J. R. K. Savage, Is there a simple answer to the origin of complex chromosome exchanges? *Int. J. Radiat. Biol.* **75**, 19–22 (1999).
34. J. R. K. Savage, A brief survey of aberration origin theories. *Mutat. Res.* **404**, 139–147 (1998).
35. M. N. Cornforth, Radiation-induced damage and the formation of chromosomal aberrations. In *DNA Damage and Repair* (J. A. Nickoloff, Ed.), pp. 559–585. Humana Press, Totowa, NJ, 1998.
36. R. K. Sachs, A. Rogoff, A. M. Chen, P. J. Simpson, J. R. Savage, P. Hahnfeldt and L. R. Hlatky, Underprediction of visibly complex chromosome aberrations by a recombinational-repair (‘one-hit’) model. *Int. J. Radiat. Biol.* **76**, 129–148 (2000).
37. J. R. K. Savage and J. D. Tucker, Nomenclature systems for FISH-painted chromosome aberrations. *Mutat. Res.* **366**, 153–161 (1996).
38. R. K. Sachs, D. Levy, A. M. Chen, P. J. Simpson, M. N. Cornforth, E. A. Ingerman, P. Hahnfeldt and L. R. Hlatky, Random breakage and reunion chromosome aberration formation model; an interaction-distance version based on chromatin geometry. *Int. J. Radiat. Biol.* **76**, 1579–1588 (2000).
39. J. N. Lucas and R. K. Sachs, Using three-color chromosome painting to test chromosome aberration models. *Proc. Natl. Acad. Sci. USA* **90**, 1484–1487 (1993).