

Selectively Advantageous Mutations and Hitchhikers in Neoplasms: p16 Lesions Are Selected in Barrett's Esophagus

Carlo C. Maley,¹ Patricia C. Galipeau,¹ Xiaohong Li,¹ Carissa A. Sanchez,¹ Thomas G. Paulson,¹ and Brian J. Reid^{1,2}

¹Divisions of Human Biology and Public Health Sciences, Fred Hutchinson Cancer Research Center, and ²Departments of Medicine and Genome Sciences, University of Washington, Seattle, Washington

ABSTRACT

Neoplastic progression is an evolutionary process characterized by genomic instability and waves of clonal expansions carrying genetic and epigenetic lesions to fixation (100% of the cell population). However, an evolutionarily neutral lesion may also reach fixation if it spreads as a hitchhiker on a selective sweep. We sought to distinguish advantageous lesions from hitchhikers in the premalignant condition Barrett's esophagus. Patients (211) had biopsies taken at 2-cm intervals in their Barrett's segments. Purified epithelial cells were assayed for loss of heterozygosity and microsatellite shifts on chromosomes 9 and 17, sequence mutations in CDKN2A/MTS1/INK4a (p16) and TP53 (p53), and methylation of the p16 promoter. We measured the expanse of a lesion in a Barrett's segment as the proportion of proliferating cells that carried a lesion in that locus. We then selected the lesion having expanses >90% in the greatest number of patients as our first putative advantageous lesion. We filtered out hitchhikers by removing all expanses of other lesions that did not occur independent of the advantageous lesion. The entire process was repeated on the remaining expanses to identify additional advantageous lesions. p16 loss of heterozygosity, promoter methylation, and sequence mutations have strong, independent, advantageous effects on Barrett's cells early in progression. Second lesions in p16 and p53 are associated with later selective sweeps. Virtually all of the other lesion expansions, including microsatellite shifts, could be explained as hitchhikers on p16 lesion clonal expansions. These techniques can be applied to any neoplasm.

INTRODUCTION

Neoplastic progression has long been recognized as an evolutionary process (1). Cells in neoplasms acquire heritable genetic and epigenetic lesions that affect their survival and reproduction, typically generating clonal heterogeneity (2–6) with evolution by natural selection (7). Indeed, most of the alterations posited as hallmarks of cancer provide a selective advantage to the neoplastic cells (8). However, fundamental principles of evolutionary biology have rarely been used to elucidate factors permissive for clonal expansion of genetic lesions in neoplasms.

Classically, the temporal evolution of somatic genetic lesions during human neoplastic progression has been inferred from frequencies of abnormalities in samples from different patients (9–12). For example, one of the best characterized molecular models, human colon cancer, was developed by comparing the frequencies of molecular lesions in adenomas of different sizes and histological grades from different patients (9). In many conditions, such “intuitive frequency” analyses are the only way to study progression because many premalignant lesions, including colonic adenomas, are removed when they are detected. Although much has been learned about multistage progression from these studies, other model human neoplasms in which

multistage evolution can be studied in the same patient spatially and temporally could improve our understanding of the evolutionary dynamics of neoplastic progression.

Barrett's esophagus (BE) is an ideal model in which to investigate clonal evolution in human epithelial neoplasia *in vivo* (13). Patients with BE present with symptoms of gastroesophageal reflux, such as heartburn, and the premalignant epithelium can be visualized at diagnostic upper endoscopy (14). Multiple biopsies can be safely obtained at endoscopy according to standardized protocols that permit localization of genetic lesions in the premalignant epithelium, and these lesions can be followed over time because periodic endoscopic biopsy surveillance is recommended for early detection of cancer (14, 15). We have previously used clonal ordering, a technique in which the order of mutations can be deduced in individual Barrett's segments, to determine the order in which genetic lesions developed in neoplastic progression (2, 13, 16). We then additionally evaluated candidate lesions, such as 17p (p53) loss of heterozygosity (LOH) and flow cytometric abnormalities, in prospective cohort studies to determine whether they predict progression to esophageal adenocarcinoma (17–19). These results indicate that neoplastic progression in BE is associated with a process of genomic instability and progressive clonal evolution similar to that hypothesized by Nowell (1) and Barrett *et al.* (2). Genomic instability in BE is associated with widespread LOH, point mutations, alterations in microsatellite alleles (shifts), and epigenetic changes, including hypermethylation of promoter regions (2, 20–24). Furthermore, genetic lesions develop very early during progression in BE; lesions in the p16 (CDKN2A/MTS1/INK4a) tumor suppressor gene are found in >85% of Barrett's segments (25).

Clonal expansion of cells with genetic lesions in a tumor suppressor or oncogene increases the probability of progression because it generates a large field of cells in which the subsequent genetic lesions can arise (26–30). Mutations that increase the ratio between rates of cell division of a clone and cell death will tend to cause the mutant clone to expand in the neoplasm (31–33). When a lesion has spread throughout an entire population it is said to have “gone to fixation” because, with no competing alleles left, natural selection cannot change the frequency of that lesion in the population. A “selective sweep” is the phenomenon of natural selection driving an allele to fixation. Therefore, we can recognize selectively advantageous mutations in a neoplasm as those mutations that tend to expand throughout a neoplasm (7).

Not all of the mutations that have gone to fixation are advantageous. Neutral mutations do not affect the fitness of a cell, but they can spread to fixation through two mechanisms. The first is by genetic drift, a purely random process that occurs very slowly and with low probability relative to a selective sweep (31, 34, 35). The second, more likely mechanism, is through linkage to a selectively advantageous lesion (Fig. 1). Such neutral mutations are called hitchhikers (36). In any given neoplasm, it is impossible to distinguish whether a genetic lesion has expanded to fixation through selection or as a hitchhiker; however, it is unlikely that the same neutral mutation would co-occur (hitchhike) with a selectively advantageous mutation across multiple independent neoplasms. Thus, the consistent expansion of a mutation in many neoplasms is evidence for an advantageous mutation.

Received 10/16/03; revised 2/19/04; accepted 3/1/04.

Grant support: NIH P01 CA91955, NIH R01 CA61202, NIH K01 CA89267-02, NIH K07 CA89147-03, NSF ANIR-9986555, ONR N00014-99-1-0417, DARPA AGR F30602-00-2-0584, the Intel Corporation, and the Santa Fe Institute.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Carlo Maley, Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, WA 98109. Phone: (206) 667-4615; Fax: (206) 667-6132; E-mail: cmaley@alum.mit.edu.

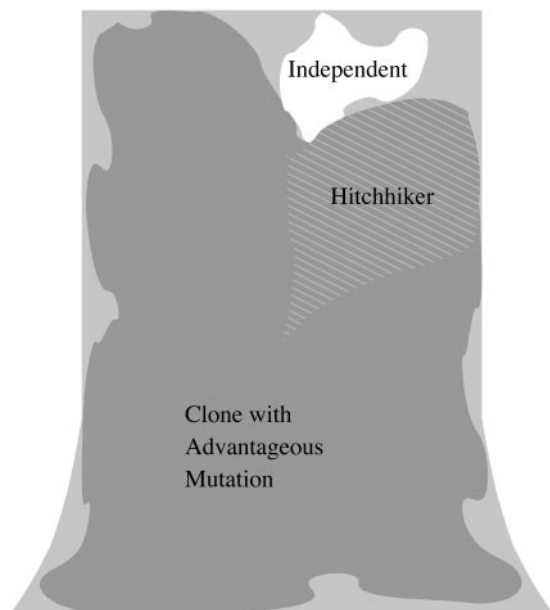


Fig. 1. Depiction of expanses of three lesions in a Barrett's segment. The *dark gray* expanse represents the area covered by cells with an advantageous lesion. The *striped expanse within the dark gray* represents cells with both the advantageous lesion and a lesion at a different locus. The *striped expanse* thus represents a potential hitchhiker on the advantageous lesion. Whether or not it really is a neutral lesion can only be inferred by analysis of multiple neoplasms to determine whether clones with the lesion consistently expand. The *white expanse* represents the area covered by cells with another lesion distinct from the advantageous lesion. This *independent expanse* cannot be explained by the advantageous lesion and would not be considered a potential hitchhiker.

The evolution of a cancer during human neoplastic progression appears to require multiple advantageous genetic lesions (8, 37–39). However, some evolutionarily neutral genetic lesions, such as shifts in the allele sizes of microsatellite markers (*i.e.*, new alleles), which are generally in noncoding regions, have been observed to undergo clonal expansion in BE and a large number of other human neoplasms as have low-frequency LOH events (2, 22, 40–47). Although it has been postulated that such expansions may be hitchhikers on advantageous lesions elsewhere in the genome, methods have not been available previously to evaluate this hypothesis or distinguish advantageous lesions from hitchhikers on sequential waves of selective sweeps (44, 45, 48, 49). Thus, analytic methods that distinguish subsequent advantageous lesions from hitchhikers after the initial selective sweep would be useful advances in studying multistage neoplastic progression. Although a second selective sweep will always occur in the background of the initial selective sweep, thus appearing to be a hitchhiker, it should expand even after the initial selective sweep has gone to fixation. Thus, on average, a second selective sweep should appear at fixation more frequently than a hitchhiker.

Here, we present a method for distinguishing selected lesions that can sweep through a BE segment from neutral lesions that may spread as hitchhikers on expansions of the advantageous mutations. This method is used to identify advantageous lesions that occur early during progression in BE. We also describe a method for distinguishing between early and later selective sweeps that may be tested in future prospective studies. It is not always possible to determine the order of events in any one particular BE segment, but by comparing data across many segments we are able to extract regularities in neoplastic progression of BE.

MATERIALS AND METHODS

Patients (Research Participants). Patients with a diagnosis of BE (211) who had a baseline endoscopic biopsy evaluation as part of the Seattle

Barrett's Esophagus Research Program between January 5, 1995 and December 2, 1999 were eligible as defined by the presence of intestinalized metaplasia in biopsies from at least two levels of the esophagus, an informative (heterozygous) marker in at least one of the microsatellites adjacent to p16 evaluated in at least one sample, clear retention or LOH in one of those microsatellites (see "LOH Analysis" below), and no prior history of esophageal malignancy. The cohort included 37 females and 174 males with a median age 64.7 years (range, 30.5–87.3) at entry into the study. BE was typically a continuous segment and the median segment length, defined by the distance between the ora serrata and the distal end of the tubular esophagus, was 7 cm (range, 2–21 cm). The Seattle Barrett's Esophagus Study was approved by the Human Subjects Division of the University of Washington in 1983 and renewed annually thereafter with reciprocity from the Institutional Review Board of the Fred Hutchinson Cancer Research Center from 1993 to 2001. Since 2001, the study has been approved annually by the Institutional Review Board of the Fred Hutchinson Cancer Research Center with reciprocity from the Human Subjects Division of the University of Washington.

Endoscopy and Biopsy. Endoscopy and biopsy were performed as described previously (17). Patients with a history of high-grade dysplasia ($n = 53$ patients) had biopsies at 1-cm intervals in the Barrett's segment, whereas those without a history of high-grade dysplasia ($n = 158$ patients) had biopsies at 2-cm intervals resulting in a total of 782 biopsies for the 211 patients. Patients diagnosed as negative for dysplasia ($n = 46$ patients), indefinite for dysplasia ($n = 67$ patients), and low-grade dysplasia ($n = 45$ patients) were grouped together because: (a) observer variation studies have shown that low-grade dysplasia cannot be reliably distinguished from negative and indefinite for dysplasia in BE (50–53); (b) prospective evaluation has revealed no significant differences in rates of progression to esophageal adenocarcinoma for negative, indefinite, and low-grade dysplasia (17); and (c) ~70% of low-grade dysplasias regress to no dysplasia compared with only 2–3% that progress to cancer (54–56).

Flow Cytometry. Biopsies were processed by DNA content or Ki67/DNA content flow cytometry as described previously (17, 19, 57–59). Cell cycle analysis (58, 60) was performed on 1467 flow-purified epithelial cell fractions with a median of 6 sorted fractions per participant (range, 2–20 fractions) depending on the Barrett's segment length. Cells from the diploid G_1 (Ki67+), 4N, and aneuploid fractions were collected for genetic analysis and counted to determine cell population size. In the rare cases where nonproliferating G_0 (Ki67–) cells were not separated from G_1 cells, we inferred the fraction of $G_1/(G_0+G_1)$ cells from the median value of the cases where G_0 cells were separated from G_1 cells (16.7%).

DNA Extraction and Whole Genome Amplification. DNA was extracted from the flow-purified cell populations using either standard phenol-chloroform or the Puregene DNA Isolation kit as recommended by the manufacturer (Gentra Systems, Inc., Minneapolis, MN). Whole genome amplification using primer extension preamplification was performed as described previously (57, 61) for each sorted fraction and three constitutive controls per participant.

LOH Analysis. LOH analysis was performed on the flow-purified fractions, as described previously (57, 62). Locus-specific primers were labeled with FAM, TET, or HEX from Research Genetics (Huntsville, AL). LOH data for chromosomes 17 and 9 were obtained from 211 patients (1389 and 1333 informative sorted fractions, respectively). Nineteen microsatellite loci were evaluated for LOH, including the 17p arm loci D17S1298 (3.87 Mb), D17S1537 (6.10 Mb), TP53-ALU (AAAAT)_n in intron 1 (7.77 Mb), TP53 (CA)_n (7.77Mb), D17S786 (9.01 Mb), D17S974 (10.72Mb), D17S1303 (11.06 Mb), chromosome 17q loci D17S1294 (28.53 Mb), D17S1293 (32.71 Mb), D17S1290 (56.81 Mb), D17S1301 (73.28 Mb), chromosome 9p loci D9S2169 (5.19 Mb), D9S935 (5.19 Mb), D9S925 (18.28 Mb), D9S932 (24.43 Mb), D9S1121 (25.39 Mb), D9S1118 (31.92 Mb), and chromosome 9q D9S301 (69.26 Mb), and D9S930 (110.62 Mb). Physical map locations were determined from the University of California Santa Cruz version hg16 July 2003 assembly.³ $Q^{LOH} = (\text{Tumor allele ratio})/(\text{Normal allele ratio})$ was determined for each locus. Q^{LOH} values ≤ 0.4 or ≥ 2.5 were considered to be clearly indicative of LOH, as described previously (18, 57, 62). Q^{LOH} values between 0.80 and 1.25 were considered to be clearly indicative of retention of heterozygosity.

³ Internet address: <http://genome.cse.ucsc.edu>.

DNA Methylation Analysis. Genomic DNA from flow-purified Barrett's epithelium was evaluated for p16 promoter methylation in 276 flow-purified fractions from 109 patients using a modification of the methods for bisulfite treatment and methylation-specific PCR as described previously (23, 25, 63). Human genomic DNA treated *in vitro* with *SssI* methyltransferase (New England Biolabs, Beverly, MA) was used as the methylated control. In a subset of cases, promoter methylation was determined and/or verified by directly sequencing PCR products of bisulfite-treated genomic DNA as described previously (25).

DNA Sequencing. Genomic or primer extension preamplification DNA was sequenced using either BigDye or BigDyeV3 Terminator cycle sequencing (Applied Biosystems, Foster City, CA) on an ABI 377, 3730, or 3700 DNA sequencer. Wild-type sequences for each participant were confirmed using constitutive samples. All of the mutations were confirmed by at least two independent PCR and sequencing reactions, and in cases of ambiguity, by direct sequencing of genomic DNA. Evaluation of mutation of exons 5–9 of the *p53* gene was performed on 766 flow-purified fractions from 176 patients using conditions described previously (64). Mutation analysis of exon 2 of the *p16* gene was performed on 1042 flow-purified fractions from 203 patients using an aliquot of genomic DNA that had undergone whole genome amplification (primer extension preamplification), as described previously (25).

Calculating the Expanse of a Lesion. We define the frequency of a lesion within a BE segment as the expanse of that lesion, to distinguish it from the frequency of a lesion across patients. We define the expanse of a lesion as the frequency of that lesion among the cells sampled from a Barrett's segment (range, 0–1), excluding the G_0 diploid cells. The cells in G_0 include the nonepithelial, stromal cells, and other cells that are no longer proliferating. Formally, the expanse E_i of a lesion in locus i for a given participant was defined as,

$$E_i = \frac{\sum_{j \in B_i} p_{ij}}{|B_i|}$$

where, B_i is the set of biopsies taken from that segment for which there are data for locus i . We divide by the total number of biopsies $|B_i|$ from the participant to normalize for the size of the Barrett's segment and derive a frequency value for a lesion between 0 and 1. In addition, we require that more than one sorted sample be informative for locus i ($|B_i| > 1$) to improve the estimate of the expanse E_i . We define p_{ij} as the proportion of $Ki67^+$ or nondiploid cells in biopsy j with a lesion in locus i .

$$p_{ij} = \frac{\Pr[\text{lesion}_i \mid Ki67^+ \text{ or nondiploid}]}{\Pr[Ki67^+ \text{ or nondiploid}]}$$

Although we have measured previously the number of levels over which lesions have spread (64), we have revised our definition of expanse because the analysis of evolution depends on measurements of allele frequencies in populations.

Finding an Initial Selective Sweep of an Advantageous Lesion. The algorithm (Fig. 2) first measures the expanse of each lesion in each locus in each Barrett's segment. LOH lesions in alternate (maternal *versus* paternal) alleles from different sorted fractions were considered to be different expanses, as were different sequence mutations and ploidy differences $>0.2N$. A histogram was constructed for each locus with a bin for each 0.1 range of expanses from 0 to 1. Fixations were defined to be lesions with expanses of 0.9–1 (90–100%) in the segment. The algorithm then infers that the lesion most frequently expanded to fixation in the cohort is the one most likely to be advantageous for a clone. It removes all of the expanses from the data that can be explained as hitchhikers on that lesion. If two different lesions have gone to fixation the same number of times in the cohort then the algorithm chooses the one that has the most partial expanses in the cohort.

Removing Hitchhikers. We identify a potential hitchhiker as a lesion that only occurs in cells that also potentially carry the putative advantageous lesion (Fig. 1). If a lesion occurs in a group of cells that do not carry the putative advantageous lesion, its presence in those cells must be independent of the effects of the putative advantageous lesion (Fig. 1). After identifying a putative advantageous lesion, the algorithm examines every expanse of every locus in every participant. If it finds a sorted fraction that has clear retention of heterozygosity ($0.80 \leq Q^{\text{LOH}} \leq 1.25$) in the putative advantageous locus but

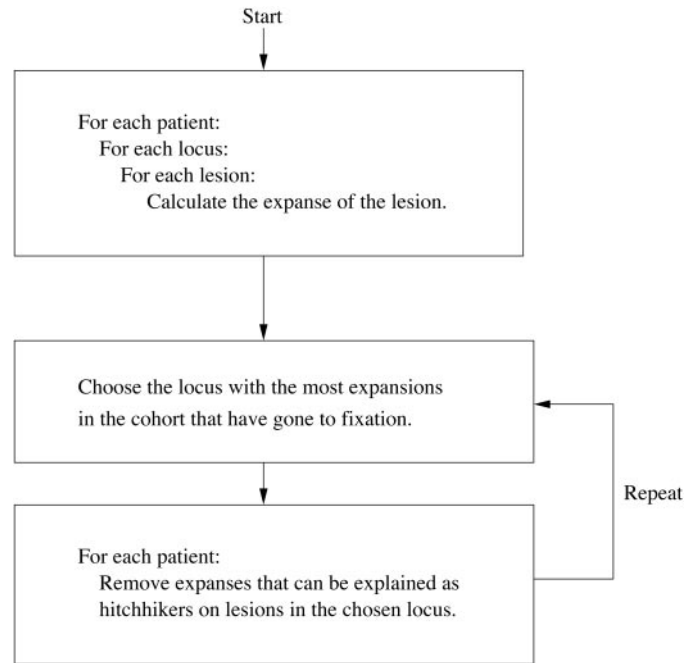


Fig. 2. Algorithm for detecting advantageous lesions by filtering out potential hitchhikers. The expanse of a lesion is the percentage of Barrett's cells, excluding diploid G_0 cells, with the lesion in a Barrett's segment. The algorithm first measures the expanse of lesions at each locus for each patient. The number of expansions in each decile (0–10%, 10–20%, and so forth) are then summed across all patients, and the lesion that has gone to fixation (90–100%) in the most patients in the cohort is selected as the putative advantageous lesion. The algorithm then removes any expanses of lesions that occur within the expanse of the putative advantageous lesion (e.g., the hitchhiker in Fig. 1). Remaining expanses that are independent of the advantageous lesion cannot be explained as hitchhikers and may be additional advantageous lesions. The algorithm then repeats the process with the remaining expanses to identify additional advantageous lesions that cannot be explained as hitchhikers on previously identified advantageous lesions.

clear LOH ($Q^{\text{LOH}} \leq 0.4$ or $Q^{\text{LOH}} \geq 2.50$) in another locus, that lesion is defined as a nonhitchhiker. The same rule holds for sequence mutations and promoter methylation. Only nonhitchhiker lesions are retained, and all of the other expanses are removed from the data. This rule for retaining nonhitchhikers is robust to missing data.

Once the hitchhikers of the putative advantageous lesion have been removed, we repeat the analysis, looking for the next most frequently fixed lesion and filtering out its potential hitchhikers. The repetition of the analysis on the remaining data allows us to detect all of the lesions that may have independent advantageous effects early in progression. A flow chart for the algorithm is shown in Fig. 2.

Finding a Subsequent Selective Sweep. If there has been a second selective sweep in a BE segment, it would be filtered out of the data set as a potential hitchhiker by the algorithm in Fig. 2. The process of finding a second selective sweep is identical to that of finding the first selective sweep when the data are restricted to the expanses that occur within the first selective sweep. Thus, we can use the same algorithm but restrict it to the expanses that occur within the genetic background of the first sweep to search for lesions that have gone to fixation more frequently in this background than would be expected if they were neutral hitchhikers.

The Perl source code for the analysis of the data are freely available upon request from the corresponding author.

Statistical Analyses. The frequency of fixation for a lesion at a locus was defined to be the number of fixations ($E_i > 0.9$) divided by the number of patients that were informative at the locus. We followed the method of Barrett *et al.* (21) iterating a binomial regression to estimate the background fixation rate and excluding outliers. We used the general linear model in the statistics package R,⁴ with each locus weighted by the number of informative patients. The fixation rate for all of the loci combined was first estimated. Then we excluded loci with lesions that were fixed more frequently than would be

⁴ Internet address: <http://www.r-project.org>.

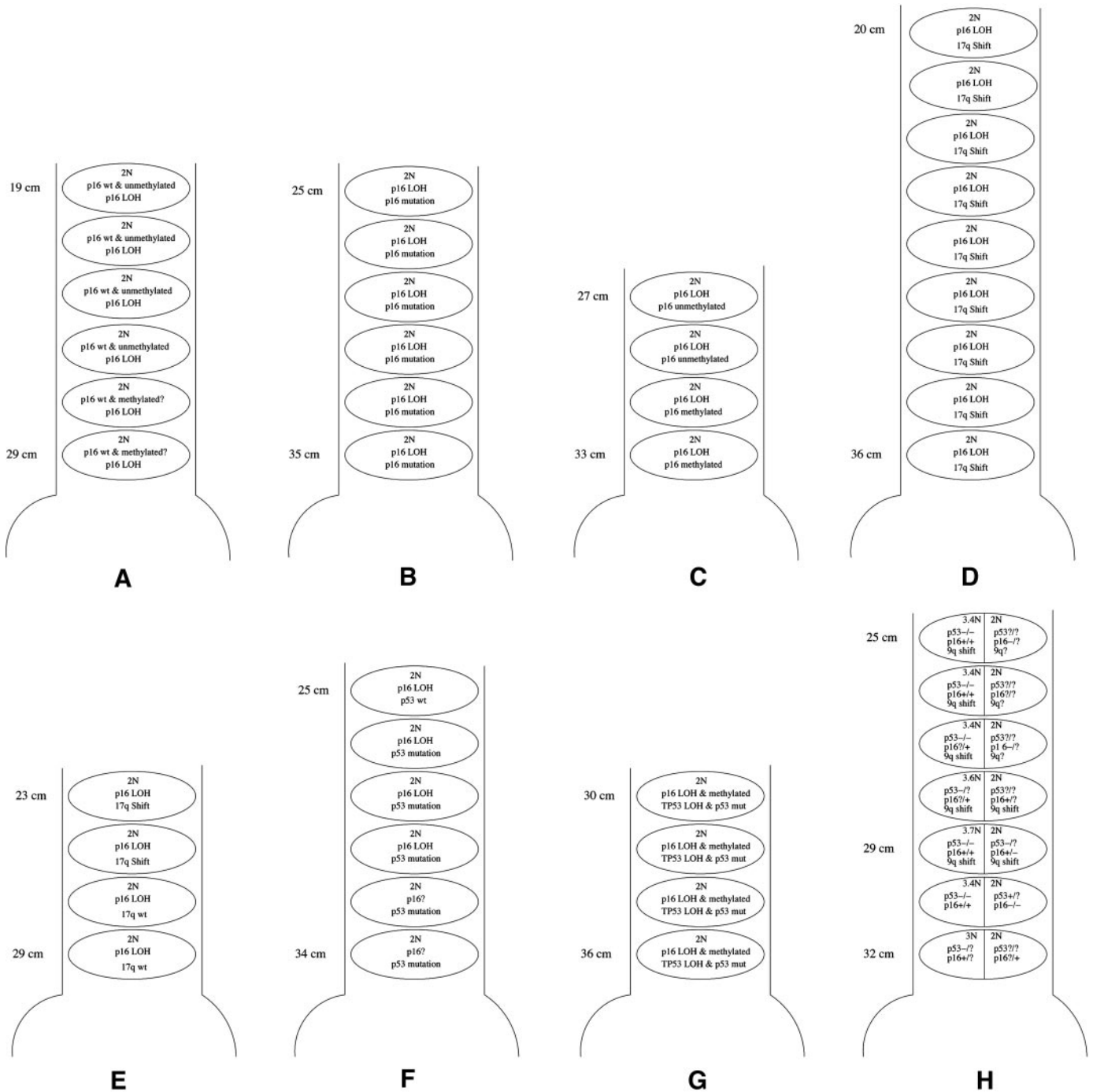


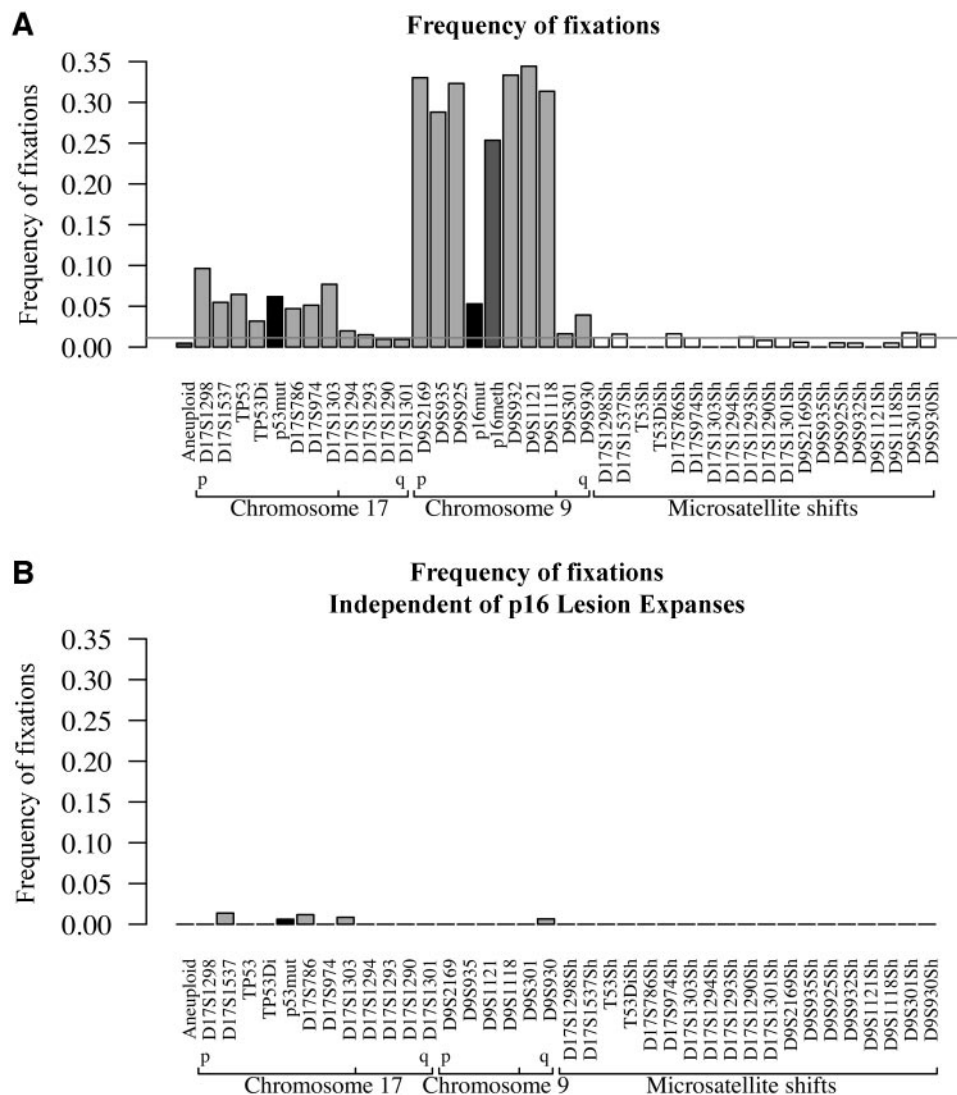
Fig. 3. Representative examples of Barrett's segments. Ploidy is represented by N, and areas marked 2N include both the diploid G_1 and G_2 sorted fractions. Missing data are indicated with a "?". A, p16 hemizygous [p16+/- with loss of heterozygosity (LOH) at D9S932] clone at fixation. B, p16 null (p16-/- with mutation and LOH at D9S925) clone at fixation. C, a p16 methylated clone appears as a partial expanse within the fixation of a p16 LOH (D9S925) lesion resulting in a mosaic of p16+/- and p16-/- clones. D, a microsatellite shift in 17q (D17S1301) and a p16 LOH (D9S925) lesion have both gone to fixation. E, a microsatellite shift in 17q (D17S1290) appears as a partial expanse within the fixation of a p16 LOH (D9S925) lesion. F, a p53 mutation appears as a partial expanse within the expanse of a p16 LOH (D9S925) lesion. G, a p16 null (mutation and LOH in D9S925), p53 null (mutation and LOH in TP53) clone has expanded to fixation. H, the only example of fixation of an aneuploid, p53 null (TP53 LOH and mutation) clone including a shift in D9S301 on 9q but without p16 lesions. A potential progenitor diploid clone carrying the same shift is observed at 29 cm. p16 lesions are marked as mutation/methylation, and p53 lesions are marked as LOH/mutation. There seems to be a competing diploid clone with a p16 mutation (p16-/-) but it only comprises 6% of the proliferating cells in the segment.

expected by the background rate ($P < 0.001$) based on the Binomial p_b of the observed number of fixations among the informative patients. We then repeated the estimation of the background fixation rate excluding the lesions that went to fixation significantly above the background rate and again excluded any new outliers based on the new estimated background fixation rate. This continued until no additional outliers could be identified.

Bootstrapping (65) was used to estimate the distribution of possible back-

ground fixation rates. We generated alternative data sets by sampling patients from the cohort 211 times, with replacement. The above method was used to estimate the background fixation rates for 1000 bootstrap data sets, and a 95% confidence interval was calculated from the results. The median value from the bootstrap results was finally used as an estimate of the true background fixation rate, and any lesions that were fixed more often than expected at $P < 0.001$ were then identified as potentially advantageous lesions. The same

Fig. 4. Frequency of fixations for lesions in each locus. In both *A* and *B*, microsatellite shifts are in *white*, loss of heterozygosity (LOH) in *gray*, methylation and aneuploidy in *dark gray*, and sequence mutations in *black*. *A*, the vertical axis is the frequency of fixations (number of fixations divided by number of patients informative for the locus) for all of the loci labeled on the horizontal axis. LOH and shift lesions are separated on the horizontal axis but otherwise the loci are ordered from P to Q arms along chromosomes 17 and 9. A lesion was considered fixed if it was detected in at least 90% of the proliferating cells from a Barrett's segment. The gray horizontal line extending across all markers represents the median estimated background fixation frequency, obtained from the bootstrap analysis, representing the chance that a lesion is fixed at random. *B*, frequency of fixations that did not occur in a Barrett's segment with a fixation of a p16 lesion (p16 mutation, p16 methylation or LOH in D9S925 or D9S932). The fixations in *B* occurred in a single patient in our cohort.



methods were used to evaluate putative hitchhiker fixations, which were defined as any lesion that went to fixation in a patient that also had a p16 mutation, p16 methylation, or p16 LOH (LOH in D9S925 or D9S932) lesion at fixation as described above (Fig. 2). p16 mutation and methylation lesions were only considered potential second hits if they expanded in a clone that also had a p16 LOH lesion.

A permutation test (66) was performed on the microsatellite shift data by randomly permuting the shift data and testing if multiple shifts tended to occur in the same locus in the same patient in the permuted data as frequently as they do in the real data. The permuted data, by construction, has the same frequency of microsatellite shifts as the real data, as well as the same distribution of flow sorted fractions per patient. Thus, it represents the null hypothesis that microsatellite shifts are occurring randomly and independently in the different flow sorted fractions and do not represent clonal markers. A permutation was generated by first reassigning the observed shifts to loci in flow sorted fractions chosen randomly with uniform probability across all of the informative loci in our data set. If the same locus in a flow sorted fraction was chosen twice, then that choice was rejected and another random locus was chosen from our data set. The randomized shifts were then measured for the percentage of loci that included more than one shift within the same participant relative to the total number of loci that included at least one shift. This statistic was then compared with the observed frequency of loci with multiple shifts. There were 10,000 such permutations generated to estimate a *P*.

The association between shift patterns and p16 lesions was tested with a χ^2 test of the 2×2 table counting the number of samples with the possible

combinations of p16 lesion/wild-type and the presence or absence of a shift in a microsatellite marker.

RESULTS

We observed a variety of different relationships among expanses of p16 lesions, p53 lesions, and microsatellite shifts (Fig. 3). Clones with

Table 1 Frequency of fixations

The frequency of fixations (expansion > 90%) of p16 and p53 lesions in the cohort. Some patients only had a single sorted fraction that was informative for a p16 or p53 lesion and so were excluded from the fixation frequency estimates for those lesions.

Lesion	Number of fixations (%)	Number of patients assayed	Number of biopsies assayed
p16 LOH ^a (D9S925 or D9S932)	73 (35.6%)	205	768
p16 methylation	18 (25.4%)	71	285
p16 mutation	10 (5.3%)	189	722
Any p16 lesion	91 (43.5%)	209	778
p53 LOH (TP53 or TP53Di)	10 (6.5%)	153	570
p53 mutation	10 (6.2%)	162	611
Any p53 lesion	15 (7.5%)	200	743
Combined p16 and p53 lesion	11 (5.5%)	199	741

^a LOH, loss of heterozygosity.

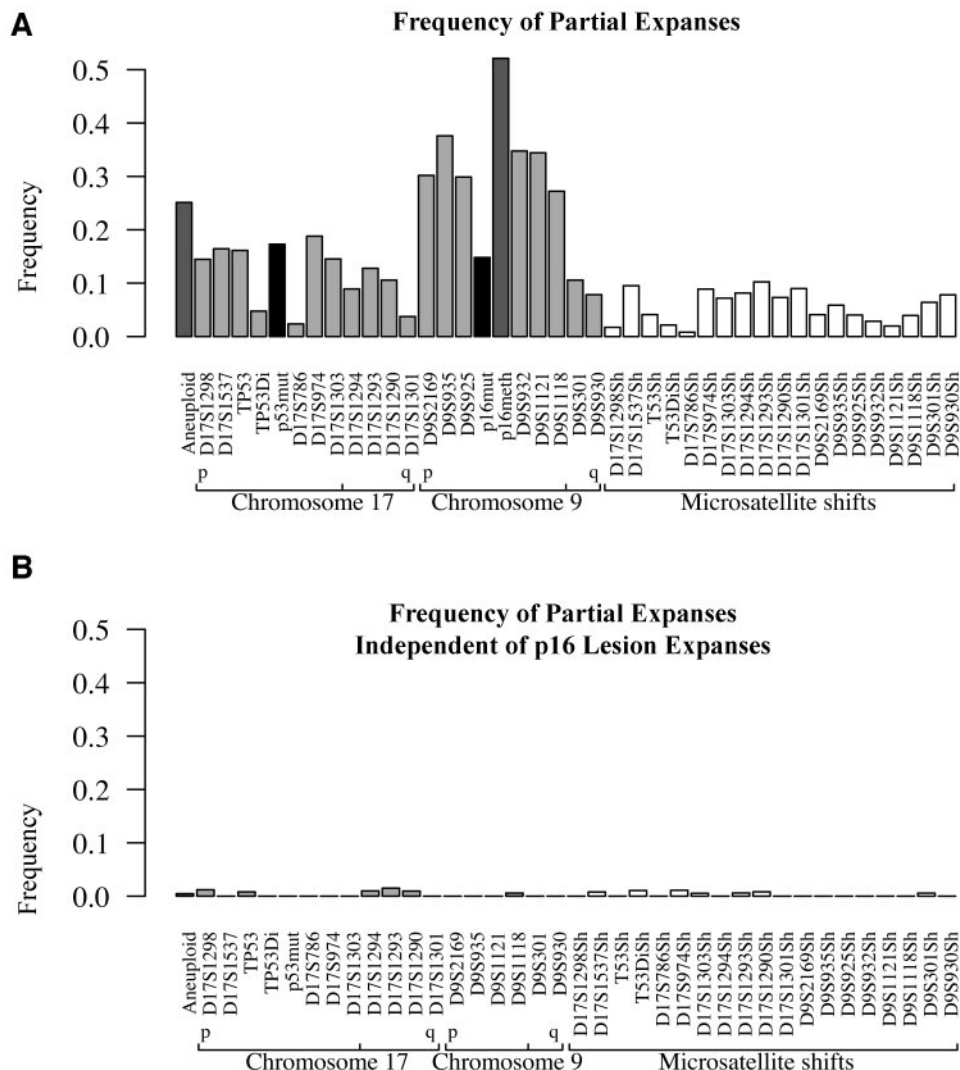


Fig. 5. Frequency of partial expanses for lesions in each locus. A partial expansion is a lesion that is detected in <90% of the proliferating cells from a Barrett's segment. In both A and B, microsatellite shifts are in white, loss of heterozygosity in gray, methylation and aneuploidy in dark gray, and sequence mutations in black. A, the vertical axis is the frequency of partial expanses (number of partial expanses divided by number of patients informative for the locus) for all of the loci labeled on the horizontal axis. B, frequency of partial expanses that are not contained within expanses of a p16 lesion (p16 mutation, p16 methylation or LOH in D9S925 or D9S932). These expanses are independent of the p16 lesion expanses and so cannot be hitchhikers on a p16 selective sweep.

p16 lesions have gone to fixation more than any other lesions in the cohort (Fig. 3, A–G). In all but 1 case in our cohort, p53 lesions occur within or are coextensive with p16 lesion expanses.

LOH at D9S925, which is linked to p16, was the lesion that went to fixation in the most patients in the cohort. Thus, it was the first putative advantageous lesion identified by the algorithm (Fig. 2). Subsequent iterations of the algorithm identified methylation of the p16 promoter, p16 sequence mutations, and LOH at D9S932 (linked to p16) as advantageous lesions. The number of fixations for lesions in each locus, normalized by the number of patients informative for each locus, is shown in Fig. 4A. The frequencies of partial expanses (<90% of the proliferating cells) are shown in Fig. 5A. Fig. 4B and Fig. 5B show the frequencies of fixations and partial expanses after the algorithm has filtered out expanses that may be explained by p16 lesions (LOH in D9S925, p16 methylation, p16 mutations, and LOH in D9S932). Of all expanses, 97.5% (1113 of 1141) can be explained by p16 lesion expanses. The full data for the number of observed expanses of different sizes and the filtering steps of the algorithm appear in the appendix.

Lesions in p16 tend to expand to fixation more frequently (43.5% of the cohort) than any other locus evaluated (Table 1; Fig. 4). When we drop the restriction on the cohort that they must have an informative locus flanking p16, chromosome 9p (p16) LOH, mutation, and methylation remain the most frequently fixed lesions (data not

Table 2 Probability of observed fixations due to chance

The probability of the observed number of fixations of lesions in loci based on the exact binomial probability with $p_b = 0.011$. A fixation is defined as a lesion that is observed in greater than 90% of the sampled proliferating cells in a Barrett's segment. These data suggest that lesions in p16 and p53 go to fixation more frequently than would be expected if they were neutral mutations being fixed as hitchhikers. Note that fixations of the LOH events that inactivate p53 may be distributed among any of the microsatellites linked to p53 and so some linked loci show more significant signs of selection than other loci. In addition, LOH in some microsatellites near p53 such as TP53Di are probably not detected as selected loci because they are frequently uninformative and are particularly sensitive to low levels of normal cell contamination, which obscures LOH in these loci.

Fixation of lesion	Binomial probability that fixations are hitchhikers
p16 methylation	<0.001
p16 mutation	<0.001
9p (p16) LOH ^a (D9S2169, D9S935, D9S925, D9S932, D9S1121, D9S1118)	<0.001
p53 mutation	<0.001
17p (p53) LOH	
TP53, D17S1303, D17S1298	<0.001
D17S1537, D17S974	<0.01
D17S786	<0.05
9q LOH	
D9S930	<0.01
D9S301	≅0.05
17q LOH (D17S1294, D17S1293, D17S1290, D17S1301) and TP53Di	≅0.05
Aneuploidy and tetraploidy	≅0.05
All microsatellite shifts	≅0.05

^a LOH, loss of heterozygosity.

shown), and so they are selected by the algorithm as the most likely to be advantageous among the loci analyzed. The only difference is that LOH at the D9S1118 locus (closely linked but not adjacent to p16) is chosen first as the lesion that most frequently went to fixation. This is probably because D9S1118 has the highest frequency of heterozygosity among the microsatellites on chromosome 9p in our cohort.

Only one Barrett's segment has a p53 lesion at fixation without apparent fixation of a p16 lesion (Fig. 3H). In this case, an aneuploid clone with p53 lesions has reached fixation (94% of cells) in a Barrett's segment that also contains a separate, small (6% of cells) diploid clone with p16 mutations. Three additional cases each show a p16 lesion in all of the sorted fractions that have a p53 lesion, but with 1 additional fraction having no p16 lesion and no data for p53. These 3 cases are thus categorized as p53 fixations but not as p16 fixations in Table 1, but there is no evidence that p53 lesions arose before p16 lesions in these cases. p16 and p53 fixations do not appear to be independent of each other ($\chi^2 = 14.5$; $P < 0.001$). If our observation that p16 lesions go to fixation in 43.5% of Barrett's segments is representative of the true rate of p16 lesion fixation, the probability that 14 of 15 p53 lesion fixations would co-occur with p16 lesions by chance is < 0.001 (Fisher's exact test).

From our bootstrap analysis of the iterated binomial regression, we estimated the background (hitchhiker) rate of fixation as a binomial distribution with $p_b = 0.0113$ (95% confidence interval, 0.0067–0.0168). Fixation of p16 methylation, p16 mutations, and 9p (p16) LOH occurred more often than would be expected based on the median bootstrap background rate of $p_b = 0.0113$ (exact Binomial $P < 0.001$). Fixation of 17p (p53) LOH and p53 mutations also occurred more often than expected (exact Binomial $P < 0.001$; Table 2). There was no evidence that aneuploid nor tetraploid (4N fraction $> 6\%$) clones tended to expand to fixation more often than expected (exact Binomial $P > 0.05$).

Microsatellite shifts were found at a frequency of 0.029 across all genotypes, and 50% (106 of 211) of patients had at least one shift among the 19 microsatellites evaluated. Of all microsatellite shifts, 70% (168 of 240) appeared in more than one sorted fraction from a patient. Shifts occurred in multiple samples in an individual at a significantly higher frequency than expected if the shifts were random (Permutation test, $P < 0.001$). On the basis of samples for which we have complete data for p16 lesions, 94% (17 of 18) of patients have shift expanses that can be explained as hitchhikers on the expansion of p16 lesions (e.g., Fig. 3, D and E), with the exception of the patient in Fig. 3H. Shift patterns were detected in association with a p16 lesion more than would be expected at random ($\chi^2 = 14.6$; 1 degree of freedom; $P < 0.001$, including samples with incomplete p16 data.).

Expanses removed by filtering (illustrated in Figs. 4 and 5) may be analyzed to detect lesions that have frequently gone to fixation in tandem with or in the genetic background of a p16 lesion. The algorithm identifies p53 sequence mutations as potential lesions associated with a later selective sweep. There were 9 cases in which a p53 mutation spread to fixation in the genetic background of a p16 lesion. After further filtering of expanses that may be hitchhikers on this potential late p53 mutant selective sweep, few expanses remain (see Appendix).

Among the expanses that were identified as potential hitchhikers on p16 lesions, bootstrap estimates of the binomial regression gave a background fixation rate of $p_b = 0.0122$ (95% confidence interval, 0.0076–0.0174). This is the estimate of the probability that a lesion will hitchhike on the first selective sweep. Given the median background fixation estimate of $p_b = 0.0122$ among the potential hitchhikers, fixations of p16 mutations, p16 methylation, p53 mutations, and 17p (p53) LOH occurred more often than would be expected (Table 3; exact Binomial $P < 0.01$).

Table 3 Probability of observed fixations in a p16 lesion background due to chance

Lesions observed in the same cells with a p16 lesion may either be neutral mutations that are hitchhikers on the selective sweep driven by the p16 lesion, or they may themselves be advantageous and associated with selective sweeps. Because p16 LOH was used as the first putative advantageous lesion to identify the potential hitchhikers, we exclude other 9p loci with LOH as probably occurring as part of the same event that generated the p16 LOH. The background rate of fixations, $p_b = 0.012$, is our best estimate for the probability that a lesion will hitchhike on the selective sweep of a p16 lesion and reach fixation by chance. The exact binomial probability of the observed rate of fixation based on this background binomial probability indicates that some lesions go to fixation more frequently than would be expected by chance. These results suggest that second hits in p16 and lesions in p53 reach fixation more frequently than would be expected if they were simply hitchhikers on the first p16 lesion.

Fixation of lesion coextensive with a p16 lesion	Binomial probability that fixations are hitchhikers
p16 methylation	< 0.01
p16 mutation	< 0.01
p53 mutation	< 0.001
17p (p53) LOH ^a	
TP53, D17S1303, D17S1298	< 0.001
D17S974	< 0.01
D17S1537	< 0.05
D17S786	≥ 0.05
9q LOH	
D9S930	< 0.01
D9S301	≥ 0.05
17q LOH (D17S1294, D17S1293, D17S1290, D17S1301) and TP53Di	≥ 0.05
Aneuploidy and tetraploidy	≥ 0.05
All microsatellite shifts	≥ 0.05

^a LOH, loss of heterozygosity.

There is evidence that both p16+/- and p16-/- clones expand. We have 4 assays for lesions in p16, including LOH at the flanking microsatellite markers D9S932 and D9S925, sequence mutations, and methylation of the promoter. Using these 4 assays, 91 patients have a p16 lesion that has swept to fixation. Ten patients represent clear examples of p16-/- clones that have gone to fixation. Fifty-eight patients are missing data for the second allele throughout the segment although a p16 lesion in the other allele has reached fixation. In the remaining 23 patients, the p16 LOH lesion that has gone to fixation has not acquired a lesion in the second allele (p16+/-) in at least some of the sorted fractions. The rest of the Barrett's segment in these 23 patients is either p16-/- or p16+/-; missing data makes it impossible to say if the entire segment is purely p16+/- or if an additional lesion has generated a p16-/- subclone.

DISCUSSION

A central task for evolutionary biology is to distinguish the regularities of selected advantageous genetic lesions from the historical accidents of selectively neutral lesions. By the time a neoplasm evolves malignancy, it typically has acquired a large number of genetic lesions (67, 68). Stoler *et al.* (67) estimate an average of $> 10^4$ lesions in a colorectal carcinoma cell. Allelotyping studies in esophageal adenocarcinomas have detected LOH on virtually every chromosome arm (21, 22, 46, 47). How could all of these lesions undergo clonal expansion? We hypothesize that the vast majority of genetic lesions detected in cancers have little, if any, selective advantage that promotes progression to malignancy but are instead historical accidents, hitchhikers on selectively advantageous lesions that arose during progression. Because we take multiple biopsies from each segment, our algorithm effectively carries out a clonal ordering analysis (2, 13, 16) on each of the 211 Barrett's segments. In BE, we report that p16 lesions frequently spread to fixation early in progression, and p53 lesions expand only in the background of p16 lesion expanses, with a single exception. In cases that have both p16 and p53 lesions at fixation, clonal ordering cannot determine whether one preceded the other in two selective sweeps or both lesions developed as early

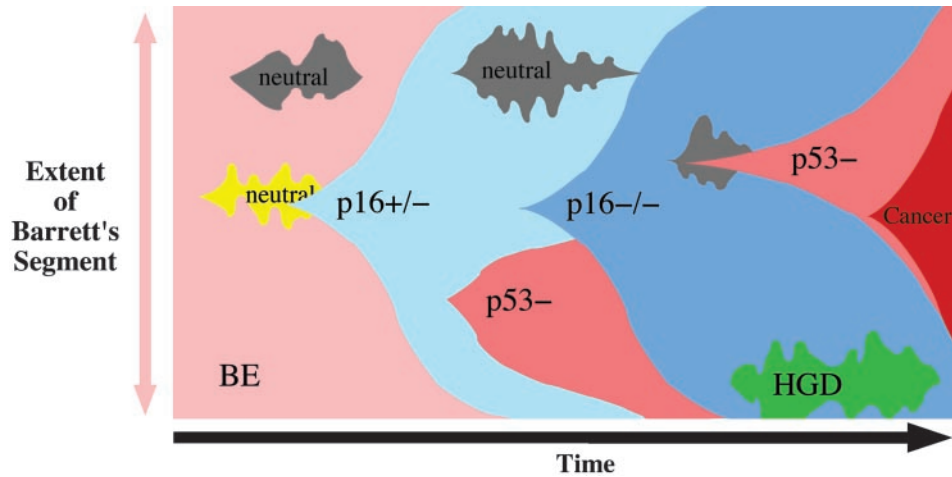


Fig. 6. A depiction of clonal evolution in a hypothetical case of Barrett's esophagus (BE). The vertical axis represents the vertical extent of the Barrett's segment. The horizontal axis represents time. Different colors signify different clones identified by their genetic lesions. A cell with a new lesion that arises within the expanse of a previous lesion carries both lesions. In this example, a neutral mutation might arise early in BE. Later a cell with that neutral mutation also acquires a lesion in one allele of p16, which then drives a selective sweep. The neutral mutation is thus carried to fixation as a hitchhiker. During this sweep, a clone with a p53 lesion arises and hitchhikes on the later part of the p16^{+/-} selective sweep. This is driven extinct by a selective sweep of a p16^{-/-} clone. Later, one of the cells with the p16^{-/-} lesions acquires another p53 lesion, which eventually leads to cancer. In the mean time, other neutral mutations, and perhaps even a clone with high-grade dysplasia, might appear and then go extinct. This view of progression was inspired by asexual population genetics (92); HGD, high-grade dysplasia.

events in a single sweep. However, fixation of a p53 lesion does not appear to be independent of p16 lesion fixations ($\chi^2 P < 0.001$). Our data suggest that both p53 lesions and second hits in p16 are advantageous and drive selective sweeps because they occur more frequently than expected by chance. All of the other lesions we have evaluated can be explained as hitchhikers on the selective sweeps of p16 lesions. These results may be tested by future prospective investigations that track expansion of clones in patients with BE to determine the temporal dynamics of clonal evolution.

There appears to be strong selection for p16 lesions in BE because p16 lesions go to fixation more frequently than any other locus we have assessed, and previous studies have reported p16 lesions in up to 90% of Barrett's segments of all histological grades (25, 69). Both p16^{+/-} and p16^{-/-} genotypes seem to be selected, and second hits in p16 go to fixation more frequently than would be expected if they were hitchhikers ($P < 0.01$), suggesting two selective sweeps, as illustrated in Fig. 6. Our results are consistent with other studies showing expansion of p16 lesions in other neoplastic conditions (49, 70, 71). If BE patients included subpopulations characterized by different genetic pathways to cancer, we would have observed hitchhiking expanses that could not be explained by p16 lesions. The algorithm would also have identified other selectively advantageous lesions on chromosomes 9 and 17 after it had filtered out the expanses explainable by p16 lesions. The fact that clones with p16 methylation are found at fixation, analogous to clones with p16 LOH and sequence mutations, suggests that in BE p16 methylation is acting as a clonal marker and that p14ARF is not the target of early alterations in BE. p14ARF is not hypermethylated in most cases and 91% (30 of 33) of cases with p16 point mutations left p14ARF unaltered or caused a conservative amino acid change.⁵ Although it is possible that p16 lesions are themselves hitchhikers on an as yet unidentified advantageous lesion elsewhere in the genome, this hypothesis is not well supported by the data. If p16 lesions are hitchhikers on another advantageous lesion, then either different p16 lesions would have to be so ubiquitous as a mosaic in a Barrett's segment that the advantageous lesion nearly always arose in a p16 lesion background or the

advantageous lesion nearly always acquired a p16 lesion early in progression. It seems unlikely that the p16 locus would be highly susceptible to unselected background lesions generated by three different mechanisms (mutation, methylation, and LOH). Our data suggest that during BE progression loss of each of the two p16 alleles predisposes to a selective sweep. The prediction that p16^{-/-} clones will expand in the background of p16 heterozygous (+/-) clones can be tested in prospective studies.

Others have noted apparent selection of what appear to be neutral lesions during human neoplastic progression and have speculated that these may occur in the background of a selectively advantageous mutation (44, 45, 48, 49). However, methods have not been available previously to distinguish advantageous and neutral mutations. Evolutionarily neutral lesions arising in a p16-deficient clone will tend to spread throughout the segment as hitchhikers on the p16 selective sweep. The most obvious cases of such hitchhiking are the large expanses of microsatellite shifts. A shift may expand to fixation if it occurs before a selective sweep (e.g., Fig. 3D). Alternatively, it may occur during the sweep and so appear as a partial expanse (e.g., Fig. 3E). It is impossible to exclude the possibility that expanses of some lesions may be generated by hitchhiking on a selective sweep driven by an unmeasured lesion elsewhere in the genome. Only ~90% of BE segments have p16 lesions, and a lesion in another as yet unidentified locus or pathway may drive selective sweeps in the remaining 10% of the cases. For example, we found a single patient with a microsatellite shift but no p16 lesions in the same sorted fraction. This is the same patient with the only p53 lesion expanse that cannot be explained by a p16 lesion. However, our data show that the vast majority of expanses in our cohort can be explained as hitchhikers on p16-mediated sweeps, without hypothesizing an unmeasured selective sweep.

p53 lesions appear to be selectively advantageous later in BE progression in the background of p16 lesions. We commonly observe p16 lesion expansions alone (e.g., Fig. 3, A–E), p53 lesion expansions within p16 lesion expansions (e.g., Fig. 3F), and the combination of both p16 lesions and p53 lesions at fixation (e.g., Fig. 3G). However, we observed p53 lesion expansion alone in only a single case, and no p16 expansions within p53 lesion expansions (Figs. 4 and 5). In the single exception with p53 expansion in a wild-type p16 background

⁵ T. G. Paulson, G. Longton, P. C. Galipeau, C. C. Maley, L. J. Prevo, H. Kissel, P. L. Blount, and B. J. Reid. p16 alterations allow evasion of the tumor suppressor effects of intestinal epithelium by clonal spread in Barrett's esophagus, manuscript in preparation.

Table A1 *Expanses of genetic lesions in each locus for 211 patients*

The entries in the table indicate the number of patients with a lesion (LOH, mutation or methylation) in a marker that has expanded to a particular proportion of their Barrett's segment. Entries in parentheses indicate the expanses of microsatellite shifts. Expanses have been grouped by 10% deciles. Size 0 expanses have been excluded for clarity. The number of patients with a fixation of a lesion in a particular locus is listed in the rows labeled with 0.9-1 expanses and the cells with the greatest number of fixations for an iteration are highlighted in grey. The columns correspond to the iterations of the filtering algorithm depicted in Figure 2. Each column shows the remaining expanses that cannot be explained by hitchhiking on clonal expansions driven by lesions in the locus indicated at the column head or any of the loci from previous column heads. Thus the last column shows the expanses that cannot be explained by hitchhiking on D9S925 LOH, p16 methylation, D9S932 LOH or p16 mutations. Before any filtering, LOH lesions in microsatellite D9S925, adjacent to p16, most frequently expand to fixation (in 53 patients). The tie with D9S1118 LOH is broken by choosing the marker with the most intermediate expanses. After filtering out expanses that may be explained as hitchhikers on D9S925 LOH, p16 methylation (p16meth) is the next most commonly fixed lesion. This is followed by LOH lesions in D9S932 and sequence mutations in p16 (p16mut). The absence of expanses in the final column shows that the vast majority of the expanses may be explained as hitchhikers on p16 lesions.

A. Chromosome 9						
Marker	Expanse	Unfiltered→	D9S925→	p16meth→	D9S932→	p16mut
D9S2169	0-0.1	1	0	0	0	0
5.19Mb	0.1-0.2	4 (2)	1 (1)	0	0	0
	0.2-0.3	4 (3)	1 (2)	0	0	0
	0.3-0.4	7	1	0	0	0
	0.4-0.5	11	1	0	0	0
	0.5-0.6	2	0	0	0	0
	0.6-0.7	1	0	0	0	0
	0.7-0.8	0	0	0	0	0
	0.8-0.9	2 (2)	0	0	0	0
	0.9-1.0	35 (1)	2 (1)	1 (1)	0	0
	D9S935	0-0.1	3 (1)	0	0	0
5.19Mb	0.1-0.2	5 (3)	1 (1)	0	0	0
	0.2-0.3	7 (2)	0 (2)	0	0	0
	0.3-0.4	7	1	0	0	0
	0.4-0.5	9 (2)	0	0	0	0
	0.5-0.6	6	1	0	0	0
	0.6-0.7	4 (1)	0	0	0	0
	0.7-0.8	0	0	0	0	0
	0.8-0.9	6 (3)	0	0	0	0
	0.9-1.0	36	2	1	0	0
	D9S925	0-0.1	1	1	1	1
18.28Mb	0.1-0.2	6 (4)	6	6	6	6
	0.2-0.3	2 (1)	2	2	2	2
	0.3-0.4	6	6	6	6	6
	0.4-0.5	7 (2)	7	7	7	7
	0.5-0.6	6 (1)	6	6	6	6
	0.6-0.7	9	9	9	9	9
	0.7-0.8	4	4	4	4	4
	0.8-0.9	8	8	8	8	8
	0.9-1.0	53 (1)	53	53	53	53
	p16mut	0-0.1	3	2	1	1
22.2Mb	0.1-0.2	1	1	0	0	0
	0.2-0.3	11	4	1	1	1
	0.3-0.4	4	2	0	0	0
	0.4-0.5	2	1	0	0	0
	0.5-0.6	1	0	0	0	0
	0.6-0.7	3	1	0	0	0
	0.7-0.8	0	0	0	0	0
	0.8-0.9	3	1	0	0	0
	0.9-1.0	10	4	2	2	2
	p16meth	0-0.1	1	1	1	1
22.3Mb	0.1-0.2	5	2	2	2	2
	0.2-0.3	3	0	0	0	0
	0.3-0.4	4	1	1	1	1
	0.4-0.5	8	5	5	5	5
	0.5-0.6	10	5	5	5	5
	0.6-0.7	3	1	1	1	1
	0.7-0.8	1	0	0	0	0
	0.8-0.9	2	1	1	1	1
	0.9-1.0	18	13	13	13	13
	D9S932	0-0.1	4 (1)	1	0	0
24.42Mb	0.1-0.2	7 (1)	0	0	0	0
	0.2-0.3	2 (1)	1	0	0	0
	0.3-0.4	3 (1)	0	0	0	0
	0.4-0.5	17 (2)	0	0	0	0
	0.5-0.6	4	1	0	0	0
	0.6-0.7	7	1	1	1	1
	0.7-0.8	2	0	0	0	0
	0.8-0.9	3	0	0	0	0
	0.9-1.0	47 (1)	7	3	3	3
	D9S1121	0-0.1	4	0	0	0
25.39Mb	0.1-0.2	3 (1)	0	0	0	0

Table A1 *Continued*

A. Chromosome 9							
Marker	Expanse	Unfiltered→	D9S925→	p16meth→	D9S932→	p16mut	
D9S1118	0.2-0.3	3	0	0	0	0	
	0.3-0.4	2 (1)	0	0	0	0	
	0.4-0.5	13	2	0	0	0	
	0.5-0.6	4	0	0	0	0	
	0.6-0.7	10 (1)	2 (1)	1	0	0	
	0.7-0.8	1	1	0	0	0	
	0.8-0.9	2 (1)	0	0	0	0	
	0.9-1.0	42	5	0	0	0	
	D9S1118	0-0.1	4 (1)	1	0	0	0
	31.92Mb	0.1-0.2	4 (2)	0	0	0	0
		0.2-0.3	4	0	0	0	0
		0.3-0.4	5 (1)	0	0	0	0
		0.4-0.5	12 (2)	1	0	0	0
		0.5-0.6	6 (1)	0	0	0	0
		0.6-0.7	5	1	0	0	0
0.7-0.8		1 (1)	1 (1)	0	0	0	
0.8-0.9		5	1	1	1	1	
0.9-1.0		53 (1)	4	1	0	0	
D9S301		0-0.1	2 (1)	0 (1)	0	0	0
67.37Mb		0.1-0.2	0 (1)	0	0	0	0
		0.2-0.3	3 (2)	2	0	0	0
		0.3-0.4	0 (5)	0 (2)	0	0	0
		0.4-0.5	3	0	0	0	0
		0.5-0.6	2	0	0	0	0
	0.6-0.7	1	0	0	0	0	
	0.7-0.8	0 (2)	0 (2)	0 (2)	0 (2)	0 (1)	
	0.8-0.9	2	0	0	0	0	
	0.9-1.0	2 (3)	0 (1)	0	0	0	
	D9S930	0-0.1	3 (2)	1	0	0	0
	108.69Mb	0.1-0.2	1 (2)	1	0	0	0
		0.2-0.3	3 (1)	1	0	0	0
		0.3-0.4	3 (2)	0 (1)	0	0	0
		0.4-0.5	2 (7)	0 (3)	0	0	0
		0.5-0.6	0 (1)	0	0	0	0
0.6-0.7		0	0	0	0	0	
0.7-0.8		0	0	0	0	0	
0.8-0.9		0	0	0	0	0	
0.9-1.0		6 (3)	1 (1)	1	1	1	
Aneuploid		0-0.1	18	1	0	0	0
0.1-0.2		11	2	0	0	0	
0.2-0.3		7	3	0	0	0	
0.3-0.4		2	1	0	0	0	
0.4-0.5		7	0	0	0	0	
0.5-0.6		3	1	0	0	0	
0.6-0.7	1	0	0	0	0		
0.7-0.8	2	1	1	1	1		
0.8-0.9	2	0	0	0	0		
0.9-1.0	1	0	0	0	0		

B. Chromosome 17							
Marker	Expanse	Unfiltered→	D9S925→	p16meth→	D9S932→	p16mut	
D17S1298	0-0.1	1 (1)	0 (1)	0	0	0	
	0.1-0.2	0 (1)	0 (1)	0	0	0	
	0.2-0.3	4 (1)	1 (1)	0 (1)	0	0	
	0.3-0.4	0	0	0	0	0	
	0.4-0.5	1	0	0	0	0	
	0.5-0.6	1	0	0	0	0	
	0.6-0.7	2	0	0	0	0	
	0.7-0.8	0	0	0	0	0	
	0.8-0.9	3	1	1	1	1	
	0.9-1.0	8 (2)	2	0	0	0	
	D17S1537	0-0.1	1	0	0	0	0
	5.85Mb	0.1-0.2	1 (1)	1 (1)	0	0	0
		0.2-0.3	1 (4)	0 (1)	0 (1)	0	0
		0.3-0.4	1 (3)	0 (1)	0	0	0
		0.4-0.5	2 (2)	2 (1)	0 (1)	0 (1)	0 (1)
0.5-0.6		1 (1)	0	0	0	0	
0.6-0.7		2	0	0	0	0	
0.7-0.8		1 (1)	1 (1)	0	0	0	
0.8-0.9		2	0	0	0	0	
0.9-1.0		4 (2)	1 (1)	1	1	1	
TP53		0-0.1	3 (2)	1 (1)	0	0	0
7.51Mb		0.1-0.2	2 (1)	0	0	0	0
		0.2-0.3	2 (1)	0	0	0	0
		0.3-0.4	6 (1)	3	0	0	0
		0.4-0.5	2	0	0	0	0
		0.5-0.6	0 (2)	0	0	0	0
	0.6-0.7	2	1	0	0	0	
	0.7-0.8	0 (1)	0 (1)	0	0	0	
	0.8-0.9	3	2	1	1	1	

Table A1 *Continued*

B. Chromosome 17							
Marker	Expanse	Unfiltered→	D9S925→	p16meth→	D9S932→	p16mut	
TP53Di 7.51Mb	0.9–1.0	8	2	0	0	0	
	0–0.1	0	0	0	0	0	
	0.1–0.2	0	0	0	0	0	
	0.2–0.3	1 (2)	0 (2)	0 (1)	0 (1)	0 (1)	
	0.3–0.4	1	0	0	0	0	
	0.4–0.5	1	1	0	0	0	
	0.5–0.6	0	0	0	0	0	
	0.6–0.7	0	0	0	0	0	
	0.7–0.8	0	0	0	0	0	
	0.8–0.9	0	0	0	0	0	
	0.9–1.0	2	0	0	0	0	
	p53mut 8.8Mb	0–0.1	2	0	0	0	0
		0.1–0.2	6	0	0	0	0
		0.2–0.3	2	0	0	0	0
0.3–0.4		9	4	0	0	0	
0.4–0.5		3	0	0	0	0	
0.5–0.6		0	0	0	0	0	
0.6–0.7		3	1	0	0	0	
0.7–0.8		1	1	0	0	0	
0.8–0.9		2	0	0	0	0	
0.9–1.0		10	4	2	1	1	
D17S786 8.75Mb		0–0.1	0	0	0	0	0
		0.1–0.2	0	0	0	0	0
		0.2–0.3	0	0	0	0	0
		0.3–0.4	0	0	0	0	0
	0.4–0.5	1	1	0	0	0	
	0.5–0.6	1	0	0	0	0	
	0.6–0.7	0	0	0	0	0	
	0.7–0.8	0	0	0	0	0	
	0.8–0.9	0 (1)	0	0	0	0	
	0.9–1.0	4 (2)	1 (2)	1 (1)	1 (1)	1	
	D17S974 10.46Mb	0–0.1	2	0	0	0	0
		0.1–0.2	3 (2)	0	0	0	0
		0.2–0.3	3 (6)	0 (3)	0 (2)	0 (1)	0 (1)
		0.3–0.4	1 (1)	1 (1)	0	0	0
0.4–0.5		5 (7)	1 (1)	0 (1)	0 (1)	0 (1)	
0.5–0.6		1	0	0	0	0	
0.6–0.7		2	1	0	0	0	
0.7–0.8		3	0	0	0	0	
0.8–0.9		2	0	0	0	0	
0.9–1.0		6 (2)	2	0	0	0	
D17S1303 10.8Mb		0–0.1	0	0	0	0	0
		0.1–0.2	2 (2)	0 (1)	0	0	0
		0.2–0.3	3 (1)	0 (1)	0 (1)	0	0
		0.3–0.4	2 (3)	0 (1)	0	0	0
	0.4–0.5	6 (2)	0 (1)	0 (1)	0 (1)	0 (1)	
	0.5–0.6	1 (1)	0	0	0	0	
	0.6–0.7	1 (3)	0 (1)	0	0	0	
	0.7–0.8	2	0	0	0	0	
	0.8–0.9	0 (1)	0	0	0	0	
	0.9–1.0	9	5	1	1	1	
	D17S1294 28.23Mb	0–0.1	2 (3)	0 (1)	0	0	0
		0.1–0.2	1 (6)	0 (1)	0	0	0
		0.2–0.3	0 (3)	0 (1)	0 (1)	0	0
		0.3–0.4	4	1	0	0	0
0.4–0.5		1 (2)	0	0	0	0	
0.5–0.6		0	0	0	0	0	
0.6–0.7		0	0	0	0	0	
0.7–0.8		0	0	0	0	0	
0.8–0.9		1	1	1	1	1	
0.9–1.0		2	2	0	0	0	
D17S1293 32.41Mb		0–0.1	5	0	0	0	0
		0.1–0.2	2 (5)	1 (2)	0	0	0
		0.2–0.3	1 (4)	0 (2)	0 (1)	0 (1)	0 (1)
		0.3–0.4	2 (2)	1	0	0	0
	0.4–0.5	2 (4)	1 (1)	0	0	0	
	0.5–0.6	0	0	0	0	0	
	0.6–0.7	0 (1)	0 (1)	0	0	0	
	0.7–0.8	3	3	1	1	1	
	0.8–0.9	2 (1)	1 (1)	1	1	1	
	0.9–1.0	2 (2)	1 (1)	0	0	0	
	D17S1290 56.67Mb	0–0.1	3 (3)	1 (1)	0	0	0
		0.1–0.2	2 (1)	0	0	0	0
		0.2–0.3	3 (2)	0 (2)	0 (2)	0 (1)	0 (1)
		0.3–0.4	1 (1)	1	0	0	0
0.4–0.5		0 (1)	0	0	0	0	
0.5–0.6		1 (1)	0	0	0	0	
0.6–0.7		0	0	0	0	0	
0.7–0.8		1	1	1	1	1	
0.8–0.9		0	0	0	0	0	
0.9–1.0		1 (1)	0 (1)	0	0	0	

Table A1 *Continued*

B. Chromosome 17						
Marker	Expanse	Unfiltered→	D9S925→	p16meth→	D9S932→	p16mut
D17S1301 73.14Mb	0–0.1	0 (1)	0 (1)	0	0	0
	0.1–0.2	0 (3)	0 (2)	0 (1)	0 (1)	0
	0.2–0.3	0 (5)	0 (4)	0	0	0
	0.3–0.4	1 (1)	0 (1)	0 (1)	0	0
	0.4–0.5	0 (3)	0 (2)	0	0	0
	0.5–0.6	0	0	0	0	0
	0.6–0.7	1 (3)	0 (3)	0 (1)	0	0
	0.7–0.8	1	0	0	0	0
	0.8–0.9	1	0	0	0	0
	0.9–1.0	1 (2)	0 (1)	0	0	0

(Fig. 3H), both alleles were inactivated (p53^{-/-}), and it is possible that rarely, two p53 lesions or a dominant-negative mutation can develop conferring a selective advantage without a p16-mediated clonal expansion. The p53 mutation observed in the exceptional case was a 10-bp deletion that created a frame shift truncation and has not been reported previously in the p53 mutation databases (72). Alternatively, lesions elsewhere in the genome may be advantageous for clonal expansion in the 10% of Barrett’s segments without p16 lesions, and this may be an example of a less common pathway to cancer that bypassed the *p16* tumor suppressor gene. Due to the stochastic nature of evolution and the potential for modifier and other selective loci, it is likely that progression in some individuals will not conform to the predictions of the model. However, we may then focus on those patients to discover alternative pathways in neoplastic progression.

One explanation of our results is that there is little or no selective pressure to lose p53 until later in progression when telomere shortening or other mechanisms of DNA damage lead to p53-mediated cell cycle arrest or apoptosis (73). Alternatively, the primary phenotype of loss of p53 in BE may be genomic instability (74–76). If this is correct, then there would typically be a delay between the loss of p53 and the generation of a selectively advantageous lesion that would carry the p53 lesion to fixation as a hitchhiker. A third hypothesis is that loss of p53 may be a selectively neutral event caused by a decrease in the length of the short telomere on the p arm of chromosome 17 (77) where p53 resides, but it is hard to imagine that loss of p53 could be entirely neutral given its roles in cell cycle control and apoptosis (73). Our data are consistent with a later selective sweep driven by p53 loss as seen in 15 cases of sweeps of p53 lesions (mutations and LOH). These sweeps occur more frequently than would be expected if they were hitchhikers ($P < 0.001$). One hypothesis that has been proposed for such observations is that later genetic lesions, such as those in p53, can only reach fixation in the background of an earlier, “gatekeeper” lesion, such as those in p16 (78). For example, a p53 lesion may be selectively advantageous within the stem cell compartment of a Barrett’s crypt but may not be able to spread and establish other mutant crypts in the absence of p16 lesions. In this way, clones with p53 lesions would remain too small for us to detect until they acquired a p16 lesion, and we would not observe them early in progression. In BE, prospective studies should be able to distinguish between cases where a clone with a p53 lesion is actively expanding, remaining a constant size, or perhaps even contracting. For example, we predict that clones with both p16 and p53 lesions will tend to expand in competition with p53 wild-type clones. In contrast, we predict that p53 mutant clones will not typically expand in a p16 wild-type genetic background.

BE is defined by intestinal metaplasia in the esophagus (15). However, BE also fulfills all of the criteria for the medical definition of a benign neoplasm (79) because it typically consists of large, clonal populations (25), is hyperproliferative relative to normal esophageal

Table A2 The expanses of genetic lesions that occur within the background of a p16 lesion

These expanses were filtered out of the data as potential hitchhikers on p16 lesions (Table A). Amongst these, p53 mutations go to fixation in more patients than any other lesion (highlighted in grey), potentially indicating a later selective sweep that occurs after the sweeps driven by p16 lesions. Microsatellite loci linked to p16 on the p arm of chromosome 9 have been excluded from this analysis because they are likely to occur as part of the same event that generates LOH in p16.

A. Chromosome 9			
Marker	Expanse	Unfiltered→	p53mut
D9S2169 5.19 Mb	0-0.1	0	0
	0.1-0.2	0 (2)	0
	0.2-0.3	0 (3)	0 (2)
	0.3-0.4	0	0
	0.4-0.5	0	0
	0.5-0.6	0	0
	0.6-0.7	0	0
	0.7-0.8	0	0
	0.8-0.9	0 (2)	0 (2)
	0.9-1.0	0 (1)	0 (1)
D9S935 5.19 Mb	0-0.1	0 (1)	0 (1)
	0.1-0.2	0 (3)	0 (2)
	0.2-0.3	0 (2)	0 (1)
	0.3-0.4	0	0
	0.4-0.5	0 (2)	0 (1)
	0.5-0.6	0	0
	0.6-0.7	0 (1)	0 (1)
	0.7-0.8	0	0
	0.8-0.9	0 (3)	0 (3)
	0.9-1.0	0	0
D9S925 18.28 Mb	0-0.1	0	0
	0.1-0.2	0 (4)	0 (3)
	0.2-0.3	0 (1)	0 (1)
	0.3-0.4	0	0
	0.4-0.5	0 (2)	0 (2)
	0.5-0.6	0 (1)	0
	0.6-0.7	0	0
	0.7-0.8	0	0
	0.8-0.9	0	0
	0.9-1.0	0 (1)	0 (1)
p16mut 22.2 Mb	0-0.1	2	2
	0.1-0.2	1	0
	0.2-0.3	10	7
	0.3-0.4	4	3
	0.4-0.5	2	2
	0.5-0.6	1	1
	0.6-0.7	3	3
	0.7-0.8	0	0
	0.8-0.9	3	1
	0.9-1.0	8	6
p16meth 22.3 Mb	0-0.1	0	0
	0.1-0.2	3	1
	0.2-0.3	3	2
	0.3-0.4	3	1
	0.4-0.5	3	2
	0.5-0.6	5	3
	0.6-0.7	2	1
	0.7-0.8	1	1
	0.8-0.9	1	1
	0.9-1.0	5	2
D9S932 24.42 Mb	0-0.1	0 (1)	0
	0.1-0.2	0 (1)	0
	0.2-0.3	0 (1)	0 (1)
	0.3-0.4	0 (1)	0
	0.4-0.5	0 (2)	0 (1)
	0.5-0.6	0	0
	0.6-0.7	0	0
	0.7-0.8	0	0
	0.8-0.9	0	0
	0.9-1.0	0 (1)	0 (1)
D9S1121 25.39 Mb	0-0.1	0	0
	0.1-0.2	0 (1)	0
	0.2-0.3	0	0
	0.3-0.4	0 (1)	0 (1)
	0.4-0.5	0	0
	0.5-0.6	0	0
	0.6-0.7	0 (1)	0
	0.7-0.8	0	0
	0.8-0.9	0 (1)	0 (1)
	0.9-1.0	0	0
D9S1118 31.92 Mb	0-0.1	0 (1)	0
	0.1-0.2	0 (2)	0 (2)
	0.2-0.3	0	0

Table A2 Continued

A. Chromosome 9			
Marker	Expanse	Unfiltered→	p53mut
D9S301 67.37 Mb	0.3-0.4	0 (1)	0 (1)
	0.4-0.5	0 (2)	0 (2)
	0.5-0.6	0 (1)	0 (1)
	0.6-0.7	0	0
	0.7-0.8	0 (1)	0 (1)
	0.8-0.9	0	0
	0.9-1.0	0 (1)	0 (1)
	0-0.1	2 (1)	0 (1)
	0.1-0.2	0 (1)	0 (1)
	0.2-0.3	3 (2)	1
D9S930 108.69 Mb	0.3-0.4	0 (5)	0 (4)
	0.4-0.5	3	0
	0.5-0.6	2	1
	0.6-0.7	1	1
	0.7-0.8	0 (1)	0 (1)
	0.8-0.9	2	1
	0.9-1.0	2 (3)	2 (2)
	0-0.1	3 (2)	0 (1)
	0.1-0.2	1 (2)	1 (2)
	0.2-0.3	3 (1)	2 (1)
Aneuploid	0.3-0.4	3 (2)	1 (2)
	0.4-0.5	2 (7)	1 (6)
	0.5-0.6	0 (1)	0
	0.6-0.7	0	0
	0.7-0.8	0	0
	0.8-0.9	0	0
	0.9-1.0	5 (3)	4 (3)
	0-0.1	18	6
	0.1-0.2	11	3
	0.2-0.3	7	2
B. Chromosome 17	0.3-0.4	2	0
	0.4-0.5	7	1
	0.5-0.6	3	1
	0.6-0.7	1	0
	0.7-0.8	1	0
	0.8-0.9	2	1
	0.9-1.0	1	0
	0-0.1	1	0
	0.1-0.2	1 (1)	0
	0.2-0.3	4 (1)	3 (1)
D17S1298 3.62 Mb	0.3-0.4	0	0
	0.4-0.5	1	0
	0.5-0.6	1	0
	0.6-0.7	2	1
	0.7-0.8	0	0
	0.8-0.9	2	0
	0.9-1.0	8 (2)	1 (1)
	0-0.1	1	0
	0.1-0.2	1 (1)	1
	0.2-0.3	1 (4)	1 (2)
D17S1537 5.85 Mb	0.3-0.4	1 (3)	0 (2)
	0.4-0.5	2 (1)	1 (1)
	0.5-0.6	1 (1)	0
	0.6-0.7	2	0
	0.7-0.8	1 (1)	0 (1)
	0.8-0.9	2	0
	0.9-1.0	3 (2)	1 (2)
	0-0.1	3 (2)	0
	0.1-0.2	2 (1)	1
	0.2-0.3	2 (1)	0
TP53 7.51 Mb	0.3-0.4	6 (1)	2
	0.4-0.5	2	1
	0.5-0.6	0 (2)	0 (1)
	0.6-0.7	2	0
	0.7-0.8	0 (1)	0 (1)
	0.8-0.9	2	1
	0.9-1.0	8	3
	0-0.1	0	0
	0.1-0.2	0	0
	0.2-0.3	1 (1)	0
TP53Di 7.51 Mb	0.3-0.4	1	1
	0.4-0.5	1	1
	0.5-0.6	0	0
	0.6-0.7	0	0
	0.7-0.8	0	0
	0.8-0.9	0	0
	0.9-1.0	2	0

Table A2 *Continued*

A. Chromosome 9			
Marker	Expanse	Unfiltered→	p53mut
p53mut	0-0.1	2	2
8.8 Mb	0.1-0.2	6	6
	0.2-0.3	2	2
	0.3-0.4	9	9
	0.4-0.5	3	3
	0.5-0.6	0	0
	0.6-0.7	3	3
	0.7-0.8	1	1
	0.8-0.9	2	2
	0.9-1.0	9	9
D17S786 8.75 Mb	0-0.1	0	0
	0.1-0.2	0	0
	0.2-0.3	0	0
	0.3-0.4	0	0
	0.4-0.5	1	0
	0.5-0.6	1	0
	0.6-0.7	0	0
	0.7-0.8	0	0
	0.8-0.9	0 (1)	0 (1)
D17S974 10.46 Mb	0.9-1.0	3 (2)	1 (2)
	0-0.1	2	0
	0.1-0.2	3 (2)	0 (1)
	0.2-0.3	3 (5)	1 (2)
	0.3-0.4	1 (1)	0 (1)
	0.4-0.5	5 (6)	2 (5)
	0.5-0.6	1	0
	0.6-0.7	2	0
	0.7-0.8	3	2
D17S1303 10.8 Mb	0.8-0.9	2	1
	0.9-1.0	6 (2)	2 (2)
	0-0.1	0	0
	0.1-0.2	2 (2)	0 (2)
	0.2-0.3	3 (1)	1 (1)
	0.3-0.4	2 (3)	0 (1)
	0.4-0.5	6 (1)	2
	0.5-0.6	1 (1)	0 (1)
	0.6-0.7	1 (3)	0 (3)
D17S1294 28.23 Mb	0.7-0.8	2	0
	0.8-0.9	0 (1)	0
	0.9-1.0	8	1
	0-0.1	2 (3)	0 (1)
	0.1-0.2	1 (6)	1 (3)
	0.2-0.3	0 (3)	0 (1)
	0.3-0.4	4	2
	0.4-0.5	1 (2)	0 (1)
	0.5-0.6	0	0
D17S1293 32.41 Mb	0.6-0.7	0	0
	0.7-0.8	0	0
	0.8-0.9	0	0
	0.9-1.0	2	1
	0-0.1	5	1
	0.1-0.2	2 (5)	2 (2)
	0.2-0.3	1 (3)	0 (2)
	0.3-0.4	2 (2)	0 (2)
	0.4-0.5	2 (4)	1 (3)
D17S1290 56.67 Mb	0.5-0.6	0	0
	0.6-0.7	0 (1)	0 (1)
	0.7-0.8	2	2
	0.8-0.9	1 (1)	1 (1)
	0.9-1.0	2 (2)	0
	0-0.1	3 (3)	1 (1)
	0.1-0.2	2 (1)	0 (1)
	0.2-0.3	3 (1)	1 (1)
	0.3-0.4	1 (1)	0 (1)
D17S1301 73.14 Mb	0.4-0.5	0 (1)	0
	0.5-0.6	1 (1)	0 (1)
	0.6-0.7	0	0
	0.7-0.8	0	0
	0.8-0.9	0	0
	0.9-1.0	1 (1)	0 (1)
	0-0.1	0 (1)	0
	0.1-0.2	0 (3)	0 (1)
	0.2-0.3	0 (5)	0 (3)
0.3-0.4	1 (1)	0 (1)	
0.4-0.5	0 (3)	0 (2)	
0.5-0.6	0	0	
0.6-0.7	1 (3)	1 (3)	
0.7-0.8	1	0	
0.8-0.9	1	1	
0.9-1.0	1 (2)	1 (2)	

and gastric epithelium (58, 80, 81), and is progressive (2). The evolution of clones that we report here is also consistent with neoplasia, and recognition of BE as a benign neoplasm may improve our understanding of the basis for its precancerous potential. The study of evolution requires measurement of changes in allele frequencies, and we did not assess the relationship between genotypes and dysplastic phenotypes because the two were generated on different samples that could not be directly compared quantitatively, although the relationships among p16, p53, and cytometric abnormalities and dysplasia have been published previously by us and others (2, 17, 18, 25, 69, 82-86).

Here, we describe a method by which advantageous mutations may be distinguished from hitchhikers during multistep human neoplastic progression. Our present analyses extend our previous work by systematically measuring regularities in the data across 211 research patients without cancer, 782 biopsies, and 1467 flow-sorted populations. This method of analysis can be generalized to many other neoplasms in which clonal fields can be evaluated spatially, including head and neck, bladder, and lung cancer among others. In fact, these techniques can be applied to any neoplastic condition in which relatively pure clones can be isolated, their expanses measured, and many neoplasms can be assessed. We achieve this in BE through acquisition of multiple biopsies per segment and flow cytometric purification of cells in the biopsies. If only one biopsy is available per neoplasm, it is still possible to make some inferences of advantageous lesions by identifying those lesions that are detected in more neoplasms than one would expect if the lesions were evolutionarily neutral (87, 88). Evaluation of precancerous neoplasms, rather than the more genetically abnormal cancers, can reduce the number of hitchhiking neutral mutations that must be filtered to identify advantageous lesions. Bladder cancer is another excellent example, and previous studies have assayed multiple samples from resection specimens for genome-wide LOH analysis, identifying 33 regions associated with progression and cancer (89-91). These analyses, which have extensive spatial and genomic mapping, show clonal expanses with LOH in chromosomes 9p and 17p and are consistent with our results. We have emphasized analysis of a large number of patients to facilitate the distinction between advantageous and hitchhiker lesions.

In summary, we have shown that p16 lesions predispose to as many as two selective sweeps within a Barrett's segment as each allele is inactivated. We have also found evidence for subsequent sweeps by clones that have acquired p53 lesions in a p16-deficient genetic background. These observations imply that competition exists among different clones within the Barrett's segment, and our data thus provide direct evidence for the concept articulated more than 25 years ago that evolution drives human neoplastic progression (Fig. 6; Ref. 1). The predictions of our model may be investigated in longitudinal studies to determine how frequently the regularities we inferred from these cross-sectional data also occur temporally.

APPENDIX

We use a filtering algorithm to identify lesions that may be selectively advantageous for clones in BE and then remove expanses of other lesions that may be explained as hitchhikers on those advantageous lesions. In four iterations, the algorithm picks out LOH at both flanking microsatellites of p16 (D9S925 and D9S932), methylation of the p16 promoter, and p16 sequence mutations as advantageous lesions. The first locus identified as a potentially advantageous lesion is LOH in D9S925, which is the most closely linked microsatellite on the telomeric side of p16. This locus was found to reach fixation most frequently, as shown in the "Unfiltered" column of Table A1. Once expanses in the other loci that could be explained as hitchhikers had

been filtered out, as shown in the “D9S925” column of Table A1, p16 promoter methylation was the next most common lesion assayed at fixation. After filtering out expanses that can be explained by p16 methylation as shown in the “p16 meth” column, the next most common lesion reaching fixation was LOH at D9S932, the most closely linked microsatellite on the centromeric side of p16. After filtering on LOH at D9S932, p16 sequence mutations (“p16mut”) were the most common lesions reaching fixation, as shown in the remaining columns of Table A1.

We exclude any loci from the filtering that were identified as loci with advantageous lesions in previous rounds of filtering. Thus, the rows of data for loci identified in the column headers of Tables A1 and A2 do not change in subsequent columns.

The expanses filtered out of Table A1 as potential hitchhikers may be analyzed with the same algorithm to search for lesions that frequently go to fixation in the genetic background of a p16 lesion (Table A2). The algorithm identifies p53 mutations as a potentially advantageous lesion in clones that also have a p16 lesion. The last column of Table A2 shows the expanses within p16 expanses that cannot be explained as hitchhikers on a later selective sweep of a p53 mutant clone.

REFERENCES

- Nowell PC. The clonal evolution of tumor cell populations. *Science* 1976;194:23–8.
- Barrett MT, Sanchez CA, Prevo LJ, et al. Evolution of neoplastic cell lineages in Barrett oesophagus. *Nat Genet* 1999;22:106–9.
- Gonzalez-Garcia I, Sole RV, Costa J. Metapopulation dynamics and spatial heterogeneity in cancer. *Proc Natl Acad Sci USA* 2002;99:13085–9.
- Nicolson GL. Tumor cell instability, diversification, and progression to the metastatic phenotype: from oncogene to oncofetal expression. *Cancer Res* 1987;47:1473–87.
- Fujii H, Marsh C, Cairns P, Sidransky D, Gabrielson E. Genetic divergence in the clonal evolution of breast cancer. *Cancer Res* 1996;56:1493–7.
- Poste G, Greig R. On the genesis and regulation of cellular heterogeneity in malignant tumors. *Invasion and Metastasis* 1982;2:137–76.
- Endler JA. *Natural Selection in the Wild*, Vol. 21. Princeton: Princeton University Press, 1986.
- Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;100:57–70.
- Vogelstein B, Fearon ER, Hamilton SR, et al. Genetic alterations during colorectal-tumor development. *N Eng J Med* 1988;319:525–32.
- Maitra A, Gazdar AF, Moore TO, Moore AY. Loss of heterozygosity analysis of cutaneous melanoma and benign melanocytic nevi: laser capture microdissection demonstrates clonal genetic changes in acquired nevocellular nevi. *Hum Pathol* 2002;33:191–7.
- Gazdar AF, Bader S, Hung J, et al. Molecular genetic changes found in human lung cancer and its precursor lesions. *Cold Spring Harb Symp Quant Biol* 1994;59:565–72.
- O’Connell P, Pekkel V, Fuqua SA, Osborne CK, Clark GM, Allred DC. Analysis of loss of heterozygosity in 399 premalignant breast lesions at 15 genetic loci. *J Natl Cancer Inst* 1998;90:697–703.
- Neshat K, Sanchez CA, Galipeau PC, et al. Barrett’s esophagus: a model of human neoplastic progression. *Cold Spring Harbor Symp Quant Biol* 1994;59:577–83.
- Levine DS, Reid BJ. Endoscopic Diagnosis of Esophageal Neoplasms. *Gastrointestinal Tumors* 1992;2:395–413.
- Sampliner RE. Updated guidelines for the diagnosis, surveillance, and therapy of Barrett’s esophagus. *Am J Gastroenterol* 2002;97:1888–95.
- Blount PL, Meltzer SJ, Yin J, Huang Y, Krasna MJ, Reid BJ. Clonal ordering of 17p and 5q allelic losses in Barrett dysplasia and adenocarcinoma. *Proc Natl Acad Sci USA* 1993;90:3221–5.
- Reid BJ, Levine DS, Longton G, Blount PL, Rabinovitch PS. Predictors of progression to cancer in Barrett’s esophagus: baseline histology and flow cytometry identify low- and high-risk patient subsets. *Am J Gastroenterol* 2000;95:1669–76.
- Reid BJ, Prevo LJ, Galipeau PC, et al. Predictors of Progression in Barrett’s Esophagus II: Baseline 17p (p53) Loss of Heterozygosity Identifies a Patient Subset at Increased Risk for Neoplastic Progression. *Am J Gastroenterol* 2001;96:2839–48.
- Rabinovitch PS, Longton G, Blount PL, Levine DS, Reid BJ. Predictors of progression in Barrett’s esophagus III: baseline flow cytometric variables. *Am J Gastroenterol* 2001;96:3071–83.
- Wild CP, Hardie LJ. Reflux, Barrett’s Oesophagus and adenocarcinoma: Burning questions. *Nat Rev Cancer* 2003;3:676–85.
- Barrett MT, Galipeau PC, Sanchez CA, Emond MJ, Reid BJ. Determination of the frequency of loss of heterozygosity in esophageal adenocarcinoma by cell sorting, whole genome amplification and microsatellite polymorphisms. *Oncogene* 1996;12:1873–8.
- Gleeson CM, Sloan JM, McGuigan JA, Ritchie AJ, Weber JL, Russell SE. Barrett’s oesophagus: microsatellite analysis provides evidence to support the proposed metaplasia-dysplasia-carcinoma sequence. *Genes, Chromosomes Cancer* 1998;21:49–60.
- Wong DJ, Barrett MT, Stoger R, Emond MJ, Reid BJ. p16^{INK4a} promoter is hypermethylated at a high frequency in esophageal adenocarcinomas. *Cancer Res* 1997;57:2619–22.
- Riegman PH, Vissers KJ, Alers JC, et al. Genomic alterations in malignant transformation of Barrett’s esophagus. *Cancer Res* 2001;61:3164–70.
- Wong DJ, Paulson TG, Prevo LJ, et al. p16 INK4a lesions are common, early abnormalities that undergo clonal expansion in Barrett’s metaplastic epithelium. *Cancer Res* 2001;61:8284–9.
- Braakhuis BJ, Tabor MP, Kummer JA, Leemans CR, Brakenhoff RH. A genetic explanation of Slaughter’s concept of field cancerization: evidence and clinical implications. *Cancer Res* 2003;63:1727–30.
- Tomlinson IPM, Novelli MR, Bodmer WF. The Mutation Rate and Cancer. *Proc Natl Acad Sci USA* 1996;93:14800–3.
- Luebeck EG, Moolgavkar SH. Simulating the Process Malig. *Trans Math Biosci* 1994;123:127–46.
- Sherman CD, Portier CJ. Stochastic simulation of a multistage model of carcinogenesis. *Math Biosci* 1996;134:35–50.
- Maley CC, Forrest S. Exploring the relationship between neutral and selective mutations in cancer. *Artificial Life* 2000;6:325–45.
- Hartl DL, Clark AG. *Principles of Population Genetics*. Sunderland, MA: Sinauer Associates, Inc., 1997.
- Darwin C. *On the Origin of Species*. London: John Murray, 1859.
- Fisher RA. *The Genetical Theory of Natural Selection*. Oxford, UK: Clarendon Press, 1930.
- Kimura M. Evolutionary rate at the molecular level. *Nature* 1968;217:624–6.
- Kimura M. Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles. *Genet Res* 1968;11:247–69.
- Smith JM, Haigh J. The hitch-hiking effect of a favourable gene. *Genet Res* 1974;23:23–35.
- Renan MJ. How many mutations are required for tumorigenesis? implications from human cancer data. *Mol Carcinog* 1993;7:139–46.
- Vogelstein B, Kinzler KW. The multistep nature of cancer. *Trends Genet* 1993;9:138–41.
- Hahn WC, Counter CM, Lundberg AS, Beijersbergen RL, Brooks MW, Weinberg RA. Creation of human tumour cells with defined genetic elements. *Nature* 1999;400:464–8.
- Meltzer SJ, Yin J, Manin B, et al. Microsatellite instability occurs frequently and in both diploid and aneuploid cell populations of Barrett’s-associated esophageal adenocarcinomas. *Cancer Res* 1994;54:3379–82.
- Gleeson CM, Sloan JM, McGuigan JA, Ritchie AJ, Weber JL, Russell SE. Ubiquitous somatic alterations at microsatellite alleles occur infrequently in Barrett’s-associated esophageal adenocarcinoma. *Cancer Res* 1996;56:259–63.
- Kulke MH, Thakore KS, Thomas G, et al. Microsatellite instability and hMLH1/hMSH2 expression in Barrett esophagus-associated adenocarcinoma. *Cancer* 2001;91:1451–7.
- Mao L, Lee DJ, Tockman MS, Erozan YS, Askin F, Sidransky D. Microsatellite alterations as clonal markers for the detection of human cancer. *Proc Natl Acad Sci USA* 1994;91:9871–5.
- Ahrendt SA, Decker PA, Doffek K, et al. Microsatellite instability at selected tetranucleotide repeats is associated with p53 mutations in non-small cell lung cancer. *Cancer Res* 2000;60:2488–91.
- Danaee H, Nelson HH, Karagas MR, et al. Microsatellite instability at tetranucleotide repeats in skin and bladder cancer. *Oncogene* 2002;21:4894–9.
- Dolan K, Garde J, Gosney J, et al. Allelotyping analysis of oesophageal adenocarcinoma: loss of heterozygosity occurs at multiple sites. *Br J Cancer* 1998;78:950–7.
- Hammoud ZT, Kaleem Z, Cooper JD, Sundaresan RS, Patterson GA, Goodfellow PJ. Allelotyping analysis of esophageal adenocarcinomas: evidence for the involvement of sequences on the long arm of chromosome 4. *Cancer Res* 1996;56:4499–502.
- Vogelstein B, Fearon ER, Kern SE, et al. Allelotyping of colorectal carcinomas. *Science* 1989;244:207–11.
- Bedi GC, Westra WH, Gabrielson E, Koch W, Sidransky D. Multiple head and neck tumors: evidence for a common clonal origin. *Cancer Res* 1996;56:2484–7.
- Montgomery E, Bronner MP, Goldblum JR, et al. Reproducibility of the diagnosis of dysplasia in Barrett esophagus: a reaffirmation. *Hum Pathol* 2001;32:368–78.
- Ormsby AH, Perras RE, Henricks WH, et al. Interobserver variation in Barrett’s (BE) related high-grade dysplasia (HGD) and superficial adenocarcinoma: can it be improved using uniform pathologic criteria? *Gastroenterology* 2000;118:A3764.
- Alikhan M, Rex D, Khan A, Rahmani E, Cummings O, Ulbright TM. Variable pathologic interpretation of columnar lined esophagus by general pathologists in community practice. *Gastrointestinal Endoscopy* 1999;50:23–6.
- Reid BJ, Haggitt RC, Rubin CE, et al. Observer variation in the diagnosis of dysplasia in Barrett’s esophagus. *Hum Pathol* 1988;19:166–78.
- Weston AP, Sharma P, Topalovskii M, Cherian R, Dixon A. Low-grade dysplasia in Barrett’s esophagus: variable fate during long-term prospective follow-up. *Gastroenterology* 1999;116:A352.
- Conio M, Bianchi S, Lapertosa G, et al. Long-term endoscopic surveillance of patients with Barrett’s esophagus. Incidence of dysplasia and adenocarcinoma: a prospective study. *Am J Gastroenterol* 2003;98:1931–9.
- Sharma P, Falk G, Weston A, Reker D, Johnston M, Sampliner R. Natural history of low grade dysplasia - an infrequent finding which usually regresses - preliminary results from the Barrett’s esophagus study. *Gastroenterology* 2002;122:A20.
- Paulson TG, Galipeau PC, Reid BJ. Loss of heterozygosity analysis using whole genome amplification, cell sorting, and fluorescence-based PCR. *Genome Res* 1999;9:482–91.

58. Reid BJ, Sanchez CA, Blount PL, Levine DS. Barrett's esophagus: cell cycle abnormalities in advancing stages of neoplastic progression. *Gastroenterology* 1993; 105:119–29.
59. Blount PL, Galipeau PC, Sanchez CA, et al. 17p allelic losses in diploid cells of patients with Barrett's esophagus who develop aneuploidy. *Cancer Res* 1994;54: 2292–5.
60. Rabinovitch PS. DNA content histogram and cell-cycle analysis. *Methods Cell Biol* 1994;41:263–96.
61. Barrett MT, Reid BJ, Joslyn G. Genotypic analysis of multiple loci in somatic cells by whole genome amplification. *Nucleic Acids Res* 1995;23:3488–92.
62. Galipeau PC, Prevo LJ, Sanchez CA, Longton GM, Reid BJ. Clonal expansion and loss of heterozygosity at chromosomes 9p and 17p in premalignant esophageal (Barrett's) tissue. *J Natl Cancer Inst* 1999;91:2087–95.
63. Wong DJ, Foster SA, Galloway DA, Reid BJ. Progressive region-specific de novo methylation of the p16 CpG island in primary human mammary epithelial cell strains during escape from M(0) growth arrest. *Mol Cell Biol* 1999;19:5642–51.
64. Prevo LJ, Sanchez CA, Galipeau PC, Reid BJ. p53-mutant clones and field effects in Barrett's esophagus. *Cancer Res* 1999;59:4784–7.
65. Efron B, Tibshirani R. The bootstrap (with discussion). *Stat Sci* 1986;1:54–77.
66. Good PI. *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*. New York: Springer-Verlag, 2000.
67. Stoler DL, Chen N, Basik M, et al. The onset and extent of genomic instability in sporadic colorectal tumor progression. *Proc Natl Acad Sci USA* 1999;96:15121–6.
68. Wiseman SM, Loree TR, Rigual NR, et al. Papillary thyroid cancer: High inter-(simple sequence repeat) genomic instability in a typically indolent cancer. *Head Neck* 2003;25:825–32.
69. Eads CA, Lord RV, Kurumboor SK, et al. Fields of Aberrant CpG island hypermethylation in Barrett's esophagus and associated adenocarcinoma. *Cancer Res* 2000; 60:5021–6.
70. Califano J, van der Riet P, Westra W, et al. Genetic progression model for head and neck cancer: implications for field cancerization. *Cancer Res* 1996;56:2488–92.
71. Califano J, Westra WH, Meininger G, Corio R, Koch WM, Sidransky D. Genetic progression and clonal relationship of recurrent premalignant head and neck lesions. *Clin Cancer Res* 2000;6:347–52.
72. Hollstein M, Shomer B, Greenblatt M, et al. Somatic point mutations in the p53 gene of human tumors and cell lines: updated compilation. *Nucleic Acids Res* 1996;24: 141–6.
73. Ko LJ, Prives C. p53: puzzle and paradigm. *Genes Dev* 1996;10:1054–72.
74. Shao C, Deng L, Henegariu O, Liang L, Stambrook PJ, Tischfield JA. Chromosome instability contributes to loss of heterozygosity in mice lacking p53. *Proc Natl Acad Sci USA* 2000;97:7405–10.
75. Xia F, Wang X, Wang YH, Tsang NM, Yandell DW, Kelsey KT, Liber HL. Altered p53 status correlates with differences in sensitivity to radiation-induced mutation and apoptosis in two closely related human lymphoblast lines. *Cancer Res* 1995;55:12–5.
76. Honma M, Hayashi M, Sofuni T. Cytotoxic and mutagenic responses to X-rays and chemical mutagens in normal and p53-mutated human lymphoblastoid cells. *Mutat Res* 1997;374:89–98.
77. Martens UM, Zijlmans JM, Poon SS, et al. Short telomeres on human chromosome 17p. *Nat Genet* 1998;18:76–80.
78. Kinzler KW, Vogelstein B. Lessons from hereditary colorectal cancer. *Cell* 1996;87: 159–70.
79. Newman WA. *Dorland's Illustrated Medical Dictionary*, 29 edition. Philadelphia: W. B. Saunders Co., 2000.
80. Gulizia JM, Wang H, Antonioli D, et al. Proliferative characteristics of intestinalized mucosa in the distal esophagus and gastroesophageal junction (short-segment Barrett's esophagus): a case control study. *Hum Pathol* 1999;30:412–8.
81. Gray MR, Hall PA, Nash J, Ansari B, Lane DP, Kingsnorth AN. Epithelial proliferation in Barrett's esophagus by proliferating cell nuclear antigen immunolocalization. *Gastroenterology* 1992;103:1769–76.
82. Suspiro A, Pereira AD, Afonso A, et al. Losses of heterozygosity on chromosomes 9p and 17p are frequent events in Barrett's metaplasia not associated with dysplasia or adenocarcinoma. *Am J Gastroenterol* 2003;98:728–34.
83. Bian YS, Osterheld MC, Fontollet C, Bosman FT, Benhattar J. p16 inactivation by methylation of the CDKN2A promoter occurs early during neoplastic progression in Barrett's esophagus. *Gastroenterology* 2002;122:1113–21.
84. Bian Y, Osterheld M, Bosman F, Benhattar J, Fontollet C. p53 gene mutation and protein accumulation during neoplastic progression in Barrett's esophagus. *Modern Pathology* 2001;14:397–403.
85. Gimenez A, Minguela A, Parrilla P, et al. Flow cytometric DNA analysis and p53 protein expression show a good correlation with histologic findings in patients with Barrett's esophagus. *Cancer* 1998;83:641–51.
86. Montgomery EA, Hartmann DP, Carr NJ, Holterman DA, Sobin LH, Azumi N. Barrett esophagus with dysplasia. Flow cytometric DNA analysis of routine, paraffin-embedded mucosal biopsies. *Am J Clin Pathol* 1996;106:298–304.
87. Newton MA. Discovering combinations of genomic alterations associated with cancer. *J Am Stat Assoc* 2002;931–42.
88. Newton MA, Yang H, Gorman P, Tomlinson I, Roylance RA. Statistical approach to modeling genomic aberrations in cancer cells (with discussion). In: Bernardo JM, Bayarri MJ, Berger JO, Dawid AP, Hecherman D, Smith AFM, and West M, editors. *Bayesian Statistics 7*: Oxford University Press, 2003.
89. Chaturvedi V, Li L, Hodges S, et al. Superimposed histologic and genetic mapping of chromosome 17 alterations in human urinary bladder neoplasia. *Oncogene* 1997;14: 2059–70.
90. Czerniak B, Chaturvedi V, Li L, et al. Superimposed histologic and genetic mapping of chromosome 9 in progression of human urinary bladder neoplasia: implications for a genetic model of multistep urothelial carcinogenesis and early detection of urinary bladder cancer. *Oncogene* 1999;18:1185–96.
91. Czerniak B, Li L, Chaturvedi V, Ro JY, Johnston DA, Hodges S, Benedict WF. Genetic modeling of human urinary bladder carcinogenesis. *Genes Chromosomes Cancer* 2000;27:392–402.
92. Crow JF, Kimura M. *An Introduction to Population Genetic Theory*. New York: Harper and Row, 1970.