HHS Public Access

Author manuscript

J Pathol. Author manuscript; available in PMC 2016 November 01.

Published in final edited form as:

J Pathol. 2015 November ; 237(3): 355–362. doi:10.1002/path.4581.

Many Private Mutations Originate From The First Few Divisions Of A Human Colorectal Adenoma

Haeyoun Kang[1], **Matthew P. Salomon**[2], **Andrea Sottoriva**[3], **Junsong Zhao**[2], **Morgan Toy**[4], **Michael F. Press**[4], **Christina Curtis**[5], **Paul Marjoram**[2], **Kimberly Siegmund**[2], and **Darryl Shibata**[4]

[1] Department of Pathology, CHA University, Seongnam-si, South Korea

[2] Department of Preventive Medicine, Keck School of Medicine of the University of Southern California, Los Angeles, California, USA

[3] Center for Evolution and Cancer, The Institute of Cancer Research, London, UK

[4] Department of Pathology, Keck School of Medicine of the University of Southern California, Los Angeles, California, USA

[5] Department of Medicine, Stanford University, Stanford, California, USA

Abstract

Intratumoral mutational heterogeneity (ITH) or the presence of different private mutations in different parts of the same tumor is commonly observed in human tumors. The mechanisms generating such ITH are uncertain. Here we find ITH can be remarkably well-structured by measuring point mutations, chromosome copy numbers and DNA passenger methylation from opposite sides and individual glands of a 6 cm human colorectal adenoma. ITH was present between tumor sides and individual glands, but the private mutations were side specific and subdivided the adenoma into two major subclones. Furthermore, ITH disappeared within individual glands because the glands were clonal populations composed of cells with identical mutant genotypes. Despite mutation clonality, the glands were relatively old, diverse populations when their individual cells were compared for passenger methylation and by FISH. These observations can be organized into an expanding star-like ancestral tree with co-clonal expansion, where many private mutations and multiple related clones arise during the first few divisions. As a consequence, most detectable mutational ITH in the final tumor originates from the first few divisions. Much of the early history of a tumor, especially the first few divisions, may be embedded within the detectable ITH of tumor genomes.

Keywords

colorectal; adenoma; intratumoral heterogeneity; mutation topography; Big Bang tumorigenesis

Corresponding Author: Darryl Shibata, University of Southern California Keck School of Medicine, 1441 Eastlake Avenue, NOR2424, Los Angeles, California, USA. dshibata@usc.edu.

Conflicts of Interest: None

Contributions: The data were generated by HK, MT, MP, and DS. The data were analyzed by MPS, AS, JZ, CC, PM, KS, and DS. The manuscript was written by HK, AS, PM, KS, and DS.

Introduction

What happens during the first few divisions that start the final growth of a visible human tumor is uncertain. The first few divisions may resemble most cell divisions, or may be unique, characterized perhaps by genomic instability with mutation bursts such as kataegis [1], chromothripsis [2], polyploidization [3], or centrosome duplications [4]. Very early growing human tumors are difficult to detect and it is uncertain if any removed small tumor would have actually progressed. However, it is possible to reconstruct the start of tumorigenesis by sampling genomes from different parts of the same tumor [5], potentially reconstructing even the first few tumor cell divisions of the last clonal expansion (Fig 1).

A recent study inferred that many human colorectal tumors are single co-clonal expansions (Big Bang tumorigenesis [6]) where most of the mutational intratumor heterogeneity (ITH) commonly observed in human tumors [7,8] originates from the first few divisions of growth. Here we illustrate with a human adenoma how extensive multi-regional sampling can reconstruct many details of an expanding population from the final topographic distributions of public (acquired before growth) and private (acquired during growth) mutations. Many "detectable" private mutations in the final tumor originate from the first few divisions because a point mutation must have a frequency >10% to be reliably detected by current exome sequencing [9]. For a simple exponential expansion at a diploid locus, frequencies in the final tumor are 50% for public mutations, 25% when a private mutation occurs during the first division, 12.5% during the second division and 6.25% during the third division.

The ancestral tree of an expanding population (Fig 1) predicts the adenoma will be composed of two distinct halves, each originating from one daughter of the first tumor cell division. Many related co-clones should be generated and scattered during growth, resulting in a star-like phylogeny. In the simplest scenario, tumor cells do not intermix, and any initial gland ITH created during the first few divisions would be dispersed during growth and gland fission [10] such that early private mutations will either be completely absent or present within all ~10,000 cells of the final glands. Alternatively, cell intermixing may be extensive, resulting in gland cells with different early private mutations (Fig S1). Most tumor mutations appear to be selectively neutral "passengers" [11] and in this analysis they are employed as lineage fate-markers of their co-clones. Here we demonstrate that the first few divisions of a human adenoma were unlike any other, with evidence of a mutation burst, polyploidization and the absence of cell intermixing.

Materials and Methods

Specimen

Two ~0.5 cm^3 bulk samples (left side or A, and right side or B) from opposite sides of a fresh 6 cm adenoma (Adenoma K in Reference 6) were incubated in an EDTA-solution to release tumor glands [5]. Individual glands/fragments of ~10,000 cells were placed in 500 ul microfuge tube and frozen (−20C) until analysis [6]. Bulk specimens were collections (hundreds) of these glands. Normal colon was obtained more than 10 cm away from the tumor. The study received local institutional review board approval.

SNP microarrays

DNAs from single glands or bulk samples were analyzed with SNP microarrays (Human OmniExpress, ~700,000 SNPs) using standard Illumina protocols and GenomeStudio with a quality threshold of 0.15. Glands with call rates <90% were not analyzed. Data were quantile-normalized and filtered using normal adjacent colon for a patient specific reference [12]. This algorithm converts BAF to mBAF (a diploid heterozygous site (AB) has an mBAF of 0, and a homozygous site has an mBAF of 1.0). Chromosome breakpoints were called by identifying segments of >100 SNPs with mBAF values >0.2. A linear transformation converted the Illumina CN data into tumor CN calls. Each segment was assigned a tentative CN: for mBAF>0.8 we set CN=1; mBAF <0.2, CN= 2; mBAF~0.3, CN= 3; and mBAF <0.2 but high raw CN values, CN= 4. A specimen specific linear regression (raw CN versus tentative CN) was used to call the tumor CN for each segment. Data are accessible via the ArrayExpress database under accession E-MTAB-2140

Exome sequencing, targeted resquencing and FISH

Bulk sequencing coverages were 24X for KA, 25X for KB and 40X for normal colon (Illumina TruSeq Exome Enrichment Kit, HiSeq 2000 platform, 100 bp paired end reads). Mutations were called using MuTect [9] at standard high confidence settings. Further filtering removed 76 mutations found within 200 bp of each other because they were commonly found in spuriously mutated genes such as CDC27 [13], and 9 of 9 such clustered mutations could not be validated by Sanger sequencing.

Targeted resequencing (AmpliSeq) was performed with PCR and IonTorrent sequencing. Gland DNA was the amount (10ng) recommended by the manufacturer. Poor quality calls (Q <20) were discarded. Mutation frequencies were the number of mutant calls divided by total calls. A threshold of > 5% was used to call a specimen mutation positive, and is based on an average “mutation” frequency at non-mutated bases of ~3% in our data.

Three sets of centromere (CH6, 7, 17) and gene probes (GOPC CH6q22.1, KDELR2 CH7p22.1, HER2 CH17q12) were used on fixed tissue sections [6]. 20 nuclei from each gland were counted to calculate average gland CNs and diversities (Shannon Index [14]).

DNA passenger methylation studies

Bisulfite treated gland DNA was amplified at two X-chromosome passenger loci (BRS and LOC) as previously described [5]. A total of 24 epialleles were analyzed at each locus per gland. The mitotic age of the adenoma was estimated using epiallele pairwise distances and age curves previously generated for 11 human colorectal cancers, which had mitotic ages (divisions since initiation) from 250 to 1,130 divisions [5].

Topographic mutation studies

Tumor regions (~2-5 glands) were microdissected using ink-blot protection [15] from five paraffin adenoma blocks. PCR with individual targeted resequencing primers was followed by Sanger sequencing. A peak threshold of 5% was used to call a region mutation positive.

Results

The adenoma, with a 0.2 cm focus of intramucosal carcinoma, was 6 cm in diameter and removed from a 50 year old male. Two bulk specimens (~0.5 cm^3) were sampled from opposite sides, and tumor glands/fragments (~10,000 cells) were isolated by EDTA-washout [5]. Mutations/alterations are sub-classified based on when they were acquired (Fig 1). Public mutations are acquired between the zygote and the first tumor cell and appear in all final tumor cells. Early private mutations are acquired during the first few divisions after tumor initiation and appear in all cells on one tumor side. Late private mutations appear in only some cells on one tumor side.

Chromosome Copy Numbers (CNs)

Chromosome CNs were measured with SNP microarrays from the bulk samples (left side = “A”, and right side = “B”) and 8 individual glands (4 left and 4 right). The left side was predominantly triploid and the right side was predominantly diploid (Fig 2 with details in Fig S2). Therefore, CN ITH is present between adenoma sides but minimal within a side. Most chromosomes were intact with 26 segments (20 whole chromosomes). There were two public (5q and 20p) LOH regions and a single late private LOH region (in 17p) in one left side gland. By contrast, except for a 5q LOH region (for which both sides had CN = 1), CNs were different for every chromosome between the left and right glands (Fig 2C). This massive CN change was essentially a haploid reduplication (22 of 24 segments) with the addition of two other chromosome segments on the left tumor side.

Point Mutations

Bulk exome sequencing revealed 227 left side mutations and 233 right side mutations, of which 70 (30%) were found on both sides. Subsets of these mutations (N= 55) were resequenced (AmpliSeq) from 12 different glands (6 left and 6 right) and both bulk specimens to an average depth of 754 reads (minimum of 20 reads). There were 13 public mutations (including a stop mutation in APC, Q789*) and 42 private mutations, with 24 private mutations on the left tumor side and 18 private mutations on the right side. Similar to the CN data and consistent with a lack of adenoma cell intermixing, private mutations were side-specific and never found on both adenoma sides (Fig 3A). Six left side early private mutations were on two of three chromosome copies, consistent with mutation acquisition followed by chromosome reduplication (Fig 3A).

The gland mutation data (Fig 3A) construct a parsimonious tree without homoplasy (Fig 3B). The significance of this tree is uncertain because the real adenoma tree has a single common ancestor and billions of tips representing the billions of present day cells. We propose that this tree approximates the first few adenoma divisions, where branches represent the genotypes of the first few adenoma cells. This hypothesis was further tested by examining two predictions. If the mutations were acquired during the first few divisions, after growth they should be disseminated throughout the adenoma and become clonal within individual tumor glands.

Detectable Mutations Are Widespread

The bulk regions (~0.5 cm^3) each represent only ~2% of the volume of the 6 cm adenoma. To test whether the private mutations were localized or disseminated, select mutations (3 public and 15 private) were Sanger sequenced from multiple microdissected tumor spots from five additional independently obtained tissue sections (Fig 3C, with details in Fig S3). The three public mutations were detected in nearly all (>90%) sampled tumor regions, and the left and right private mutations separated into distinct regional patches. Right private mutations were present throughout four tissue sections. The last tissue section was divided into two distinct halves, equally occupied by either left or right private mutations. Based on this seven region sample (five tissue sections and two bulk regions), the left and right early adenoma tree branches respectively formed ~20% and 80% of the final tumor (Fig 3C).

If the three tested early right mutations (CREB5, FBXO1, SF1) in Fig 3B were acquired in the same mitosis, they should co-localize in the tissue sections (see SOM). However, patch-like regions with three distinct genotypes (CREB5+,SF1+, FBXO1+, FHAD1+, CHD7+ like the B1 gland genotype; CREB5+, SF1+, FBXO1+ like the B2 gland genotype; and CREB5+, SF1+ or a new “B0” genotype) were observed, subdividing the first right branch in Fig 3A into at least two divisions (CREB5 and SF1 mutation in one division, and then FBXO1). This sequence of mutation accumulation (Fig 3D) represents co-clonal expansion and not sequential clonal evolution because progeny of these first few divisions occupied large portions of the final tumor (Fig 3C), indicating a growing star-like phylogeny (Fig 1). The detection of the private mutations throughout the seven tumor regions suggests that the first adenoma cell was correctly inferred (see supplementary online material).

Adenoma Glands Are Clonal For Detectable Point Mutations

In an expanding population, the earliest private mutations will become clonal in the final tumor glands because early tumor cells have millions of progeny. Gland cell composition can be estimated from mutation frequency (MF) and CN data. Gland MFs may vary from clonal frequencies because of many factors including ITH from cell intermixing, normal cell contamination and allele amplification/sequencing biases. Indeed public mutations, which should be clonal, had “bell-shaped” MF distributions centered about their expected clonal values. To control for these and other uncertainties, public and private MFs were compared (Fig 4). For the bulk specimens, private MFs were generally lower than expected compared to public MFs, consistent with private mutations in only some tumor cells. By contrast, both private and public average gland MFs were close to their expected clonal frequencies and not significantly different between private and public mutations. Therefore, as with the CN data, the point mutation data are consistent with clonal or “quantum” glands filled with cells of the same genotypes. Each gland has a well-defined genotype where a detectable private mutation is either absent or present in all gland cells. Mutation ITH is present between adenoma sides and between some neighboring glands, but disappears within a gland.

Adenoma Glands Are Diverse Cell Populations

Although glands will inevitably be clonal for early private mutations due to growth, ongoing sequential selection can also cause clonality. A difference between these scenarios is the “age” of the glands. With a single expansion, glands are created early and therefore will be

uniformly old or diverse populations and clonal for the early private mutations. With ongoing selection, glands can also be clonal for the private mutations but will be relatively young, minimally diverse populations because of recent bottlenecks. Gland diversity can be estimated by measuring differences between individual gland cells at highly unstable genomic features such as chromosome CN and passenger methylation.

Prior DNA passenger methylation studies demonstrated that most CRCs are single expansions with relatively old and diverse glands [5]. Similarly, the adenoma glands were diverse populations with multiple unique epialleles (Fig 5A) and, consistent with a single expansion, adenoma diversity appeared relatively uniform because gland diversities (defined by the number of unique gland patterns and gland pairwise distances) were not significantly different between opposite sides. The passenger DNA methylation diversity of this adenoma is comparable to that found in other CRCs, whose estimated mitotic ages ranged between 250 and 1,100 divisions since tumor initiation [5].

Consistent with old and diverse gland populations, and chromosomal instability [16], CNs were different between adjacent adenoma cells within a single gland, with higher Shannon Diversity Index values compared to normal colon crypts for 5 of 6 FISH probes (Fig 5B). The observation that average gland ploidy is identical at most chromosome gland fragments within a side (Fig 2C) yet different between cells within a gland is more consistent with a single clonal expansion, with similar average integer gland CNs reflecting ancestry from an early side progenitor.

Private and Public Mutation Spectra Differ

Point mutation spectra can infer underlying mutational mechanisms [1]. Most public point mutations were acquired long ago in normal colon [17], whereas the private mutations arose in the early tumor microenvironment. The public and private mutation spectra differed (Fig 5C), with private mutations having significantly more C>A transversions and significantly fewer A>G transitions. Consistent with mutations arising by the same mechanisms and within the same early microenvironment, there were no significant differences between the left and right private mutation spectra.

Reconstruction of the First Few Divisions

Using the early ancestral adenoma tree as a guide (Fig 3B), the genotypes of the first adenoma cell and the first “side” bulk progenitors can be inferred (Fig 6). Public mutations, accumulated over ~50 years between the zygote and tumor initiation, define the genotype of the first adenoma cell. A minority of the point mutations accumulated before tumor initiation, with two small regions of LOH. The only high frequency recurrent colorectal tumor driver mutation [11] was a truncating APC mutation with LOH.

Early private mutations found in all glands of one bulk sample define the genotype of the side bulk progenitor. Assuming a star-like phylogeny with exponential expansion and no cell intermixing, the last possible common ancestor for this tumor region (~ 2% each of the 6 cm adenoma) could arise by the fifth or sixth cell division after tumor initiation as $(½)^6 = 0.015$ and $(½)^5 = 0.031$ (see SOM for details). During the up to six divisions between the first adenoma cell and a bulk side progenitor, there appeared to be a single coordinated haploid

duplication leading to near triploidy in the first progenitor left side cell (Fig 2). Six left side early private mutations were on two chromosome copies (Fig 3A), indicating they occurred before the haploid duplication. A scenario (see Discussion) that could generate both side progenitors after a few divisions includes an early point mutation burst (30 of the 42 resequenced private mutations) and unscheduled whole-genome duplication with uneven cytokinesis (Fig 6). A mutation burst could also explain mutation spectra differences between the public and private mutations (Fig 5C).

Discussion

Tumors expand from single cells. Although tumor ancestry can be inferred from single specimens [18], here we illustrate how this process is facilitated by comparing genomes from multiple tumor regions. The ITH of this adenoma was well-structured, with detectable mutational ITH disappearing within glands and private mutations segregated by tumor side. Adenoma glands were clonal with respect to average chromosome CNs and detectable private mutations. Moreover, the glands were uniformly old cell populations with diverse passenger methylation patterns and CN differences between individual cells.

This structured ITH can be explained by a relatively simple star-like co-clonal expansion where most detectable mutational ITH essentially originates all at once during the first few adenoma cell divisions. If the first adenoma cell acquired a selective proliferative advantage, its immediate daughter cells should also be similarly capable, with the co-clonal expansion producing a star-like ancestral tree (Fig 1) and a spherical tumor. The left and right side progenitors had different genotypes and slightly different proliferation, with progeny of the left being predominantly triploid and producing about 20% of the adenoma. The patch-like subclone distribution in the final adenoma (Fig 3C) is reminiscent of normal growth where cells do not intermix, exemplified by distinct, contiguous G6PD colon crypt patches generated by X-linked inactivation during development in appropriate heterozygotes [19].

Stepwise clonal evolution could also sequentially generate multiple subclones within the same tumor. However, private mutations that arise after the initial expansion must undergo subsequent clonal expansion to become detectable. Stepwise clonal evolution becomes increasingly less plausible as multi-regional sampling uncovers increasingly more subclones because of the paucity of distinguishable driver mutations. Indeed, only a single canonical driver mutation in APC was detected in this adenoma. In addition, each new selective sweep would have to be large enough to confer detectability but still be limited enough to maintain ITH. A more parsimonious scenario is a single star-like expansion where many co-clones and their hitchhiking private passenger mutations arise early and inherently become detectable from growth. Selection occurs once at the start of tumorigenesis, which can help explain the paucity of driver mutations in this adenoma and most tumors [11]. Although not sampled, potentially stepwise progression could explain the small focus of carcinoma in this adenoma.

The generation of multiple distinct co-clones during the first few adenoma cell divisions can be explained by an unusual burst of private mutations and massive chromosomal missegregation. With a normal mutation rate ($\sim 10^{-9}$ per base per division), at most one new

exome mutation (~40 MB target) should arise per division, yet at least 30 new verified resequenced private mutations were inferred to arise during the first five to six divisions minimally required to form each small bulk tumor region. This mutation burst is not kategnesis [1] because the mutations were not clustered. The mutation spectra differed between the public and private mutations, with significantly fewer A>G transitions and greater C>A transversions, but no obvious mutation mechanism was evident.

A precedent for this type of division is the idea that polyploidization may be an important early step in tumorigenesis [3,4]. Tetraploidization has been observed in APC mutant cells [20,21] and an APC mutation with LOH was present in this first adenoma cell. DNA damage, telomere shortening and replication stress can also lead to unscheduled DNA duplication [3]. Multiple centrosome duplications and subsequent unequal cytokinesis could lead to a single division resulting in essentially a haploid gain in the left “triploid” side progenitor.

This analysis illustrates how growth embeds early human tumor histories within genomes taken from its different parts. Growth is a major signature embedded within ITH, where most detectable mutations occurred long ago [6]. The early tumor history is robust to many changes (mutations, small subclonal expansions) that occur later in tumorigenesis because all cells still retain early private mutations in their genomes. Tumors may evolve uniquely, and comprehensive, systematic sampling is imperative for understanding how individual tumors arose [22-26]. Further ancestral reconstructions of more human tumors should illuminate whether they are also single co-clonal expansions and if their first few divisions are also unusually chaotic, which may help inform better efforts for early cancer detection and prevention.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by grants from the NIH (R21 CA185016, P30CA014089)

References

1. Nik-Zainal S, Alexandrov LB, Wedge DC, et al. Mutational processes molding the genomes of 21 breast cancers. Cell. 2012; 149:979–993. [PubMed: 22608084]
2. Stephens PJ, Greenman CD, Fu B, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell. 2011; 144:27–40. [PubMed: 21215367]
3. Davoli T, de Lange T. The causes and consequences of polyploidy in normal development and cancer. Ann Rev Cell Dev Biology. 2011; 27:585–610.
4. Pihan GA. Centrosome dysfunction contributes to chromosome instability, chromoanagenesis, and genome reprograming in cancer. Front Oncol. 2013; 3:277. [PubMed: 24282781]
5. Siegmund KD, Marjoram P, Woo YJ, et al. Inferring clonal expansion and cancer stem cell dynamics from DNA methylation patterns in colorectal cancers. Proc Natl Acad Sci USA. 2009; 106:4828–4833. [PubMed: 19261858]
6. Sottoriva A, Kang H, Ma Z, et al. A Big Bang model of human colorectal tumor growth. Nat Genet. 2015; 47:209–216. [PubMed: 25665006]

7. Shibata D. Cancer. Heterogeneity and tumor history. Science. 2012; 336:304–305. [PubMed: 22517848]
8. Burrell RA, McGranahan N, Bartek J, et al. The causes and consequences of genetic heterogeneity in cancer evolution. Nature. 2013; 501:338–345. [PubMed: 24048066]
9. Cibulskis K, Lawrence MS, Carter SL, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol. 2013; 31:213–219. [PubMed: 23396013]
10. Humphries A, Wright NA. Colonic crypt organization and tumorigenesis. Nat Rev Cancer. 2008; 8:415–424. [PubMed: 18480839]
11. Vogelstein B, Papadopoulos N, Velculescu VE, et al. Cancer genome landscapes. Science. 2013; 339:1546–1558. [PubMed: 23539594]
12. Olshen AB, Bengtsson H, Neuvial P, et al. Parent-specific copy number in paired tumor-normal studies using circular binary segmentation. Bioinformatics. 2011; 27:2038–2046. [PubMed: 21666266]
13. Jia P, Li F, Xia J, et al. Consensus rules in variant detection from next-generation sequencing data. PLoS One. 2012; 7:e38470. [PubMed: 22715385]
14. Park SY, Gonen M, Kim HJ, et al. Cellular and genetic diversity in the progression of in situ human breast carcinomas to an invasive phenotype. J Clin Invest. 2010; 120:636–644. [PubMed: 20101094]
15. Shibata D, Hawes D, Li ZH, et al. Specific genetic analysis of microscopic tissue after selective ultraviolet radiation fractionation and the polymerase chain reaction. Am J Pathol. 1992; 141:539–543. [PubMed: 1325739]
16. Lengauer C, Kinzler KW, Vogelstein B. Genetic instability in colorectal cancers. Nature. 1997; 386:623–627. [PubMed: 9121588]
17. Tomasetti C, Vogelstein B, Parmigiani G. Half or more of the somatic mutations in cancers of self-renewing tissues originate prior to tumor initiation. Proc Natl Acad Sci USA. 2013; 110:1999–2004. [PubMed: 23345422]
18. Nik-Zainal S, Van Loo P, Wedge DC, et al. The life history of 21 breast cancers. Cell. 2012; 149:994–1007. [PubMed: 22608083]
19. Novelli M, Cossu A, Oukrif D, et al. X-inactivation patch size in human female tissue confounds the assessment of tumor clonality. Proc Natl Acad Sci USA. 2003; 100:3311–3314. [PubMed: 12610207]
20. Fodde R, Smits R, Clevers H. APC, signal transduction and genetic instability in colorectal cancer. Nat Rev Cancer. 2001; 1:55–67. [PubMed: 11900252]
21. Dikovskaya D, Schiffmann D, Newton IP, et al. Loss of APC induces polyploidy as a result of a combination of defects in mitosis and apoptosis. J Cell Biol. 2007; 176:183–195. [PubMed: 17227893]
22. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. New Engl J Med. 2012; 366:883–892. [PubMed: 22397650]
23. Johnson BE, Mazor T, Hong C, et al. Mutational analysis reveals the origin and therapy-driven evolution of recurrent glioma. Science. 2014; 343:189–193. [PubMed: 24336570]
24. Zhang J, Fujimoto J, Zhang J, et al. Intratumor heterogeneity in localized lung adenocarcinomas delineated by multiregion sequencing. Science. 2014; 346:256–259. [PubMed: 25301631]
25. de Bruin EC, McGranahan N, Mitter R, et al. Spatial and temporal diversity in genomic instability processes defines lung cancer evolution. Science. 2014; 346:251–256. [PubMed: 25301630]
26. Humphries A, Cereser B, Gay LJ, et al. Lineage tracing reveals multipotent stem cells maintain human adenomas and the pattern of clonal expansion in tumor evolution. Proc Natl Acad Sci U S A. 2013; 110:2490–2499. [PubMed: 23335629]

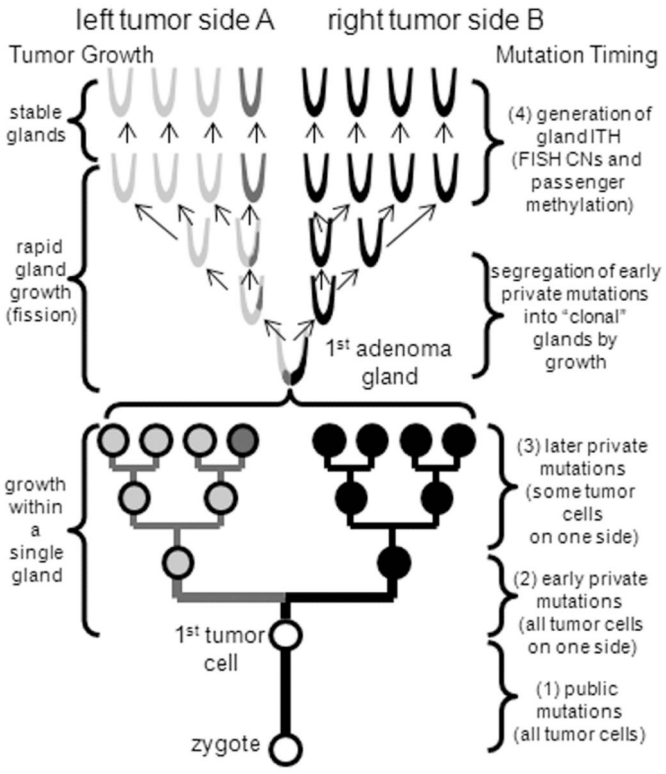

Figure 1. Reconstructing the first few adenoma cell divisions

Measurements of multiple glands from opposite adenoma sides allow reconstruction of this co-clonal expansion.

	left side				right side			
	KA1	KA2	KA6	KA7	KB1	KB3	KB4	KB5
17p LOH		2						
chr1	3	3	3	3	2	2	2	2
chr2	3	3	3	3	2	2	2	2
chr3	3	3	3	3	2	2	2	2
chr4	3	3	3	3	2	2	2	2
chr5p	3	3	3	3	2	2	2	2
chr5q	3	3	3	3	2	2	2	2
chr6	3	3	3	3	2	2	2	2
chr7	4	4	4	4	3	3	3	3
chr8	3	3	3	3	2	2	2	2
chr9	4	4	4	4	2	2	2	2
chr10	3	3	3	3	2	2	2	2
chr11	3	3	3	3	2	2	2	2
chr12	3	3	3	3	2	2	2	2
chr13	4	4	4	4	3	3	3	3
chr14	3	3	3	3	2	2	2	2
chr15	3	3	3	3	2	2	2	2
chr16	3	3	3	3	2	2	2	2
chr17	3	3	3	3	2	2	2	2
chr18	3	3	3	3	2	2	2	2
chr19	3	3	3	3	2	2	2	2
chr20q	3	3	3	3	2	2	2	2
chr21	4	4	4	4	2	2	2	2
chr22	3	3	3	3	2	2	2	2
5q LOH	1	1	1	1	1	1	1	1
20p LOH	2	2	2	2	1	1	1	1

Figure 2. SNP data

Heat map summarizing the SNP microarray data illustrating that the left and right adenoma sides are distinct, but CNs (adjusted to nearest integer values) are identical within a side (additional details in Fig S2).

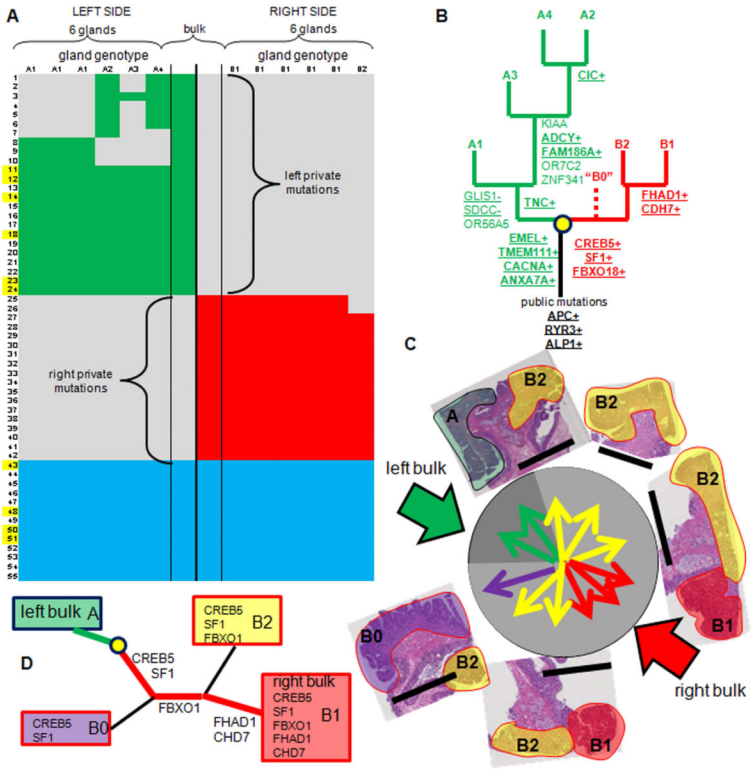

Figure 3. Point mutation analysis

A: Mutation resequencing data illustrate that public mutations (blue) were in all glands. Private mutations (green, left, or red, right) were isolated to each side. The glands are be classified into "genotypes" based on their mutations, with 4 and 2 different gland genotypes on the left and right sides. Yellow loci indicate mutations on more than one chromosome copy. Mutation frequency analysis (Fig 4) indicated a quantum clonal private mutation distribution -a private mutation is either absent from a gland or present in all cells of a gland. A resequenced mutation list is provided in Supplementary Table S1. **B:** Parsimonious ancestral tree manually inferred from the 6 indentified bulk gland genotypes (Fig 3A). Underlined loci indicate mutations further examined in five additional independent microscopic tissues sections. Most tested private mutations are also present in other parts of the adenoma. An additional early branch leading to a B0 subclone was inferred when sampling additional tissue sections (see below). **C:** Topographical distributions of 3 public and 15 private mutations tested with microdissection and Sanger sequencing in five additional tissue sections (details in Fig S3). Tumor section locations are inferred from the mutation data. The three public mutations were detected throughout the sections. There was no intermixing of the left and right private mutations in the tissue sections, with patch-like mutation topographies observed with the right side private mutations, which segregated into three distinct subclones (BO, B1, B2, see Fig 3D). The subdivision into left and right subclones was captured in one tissue slide. The left branch appeared to occupy about 20% of the final adenoma. Bars indicate 1 cm. **D:** The bulk specimen ancestral tree (Fig 3B) was further resolved with the five additional tissue sections, which identified one earlier subclone (B0), which lacked the FBXO1 private mutation. There were no tumor regions

without private mutations, suggesting that the first adenoma cell (yellow dot) was correctly inferred.

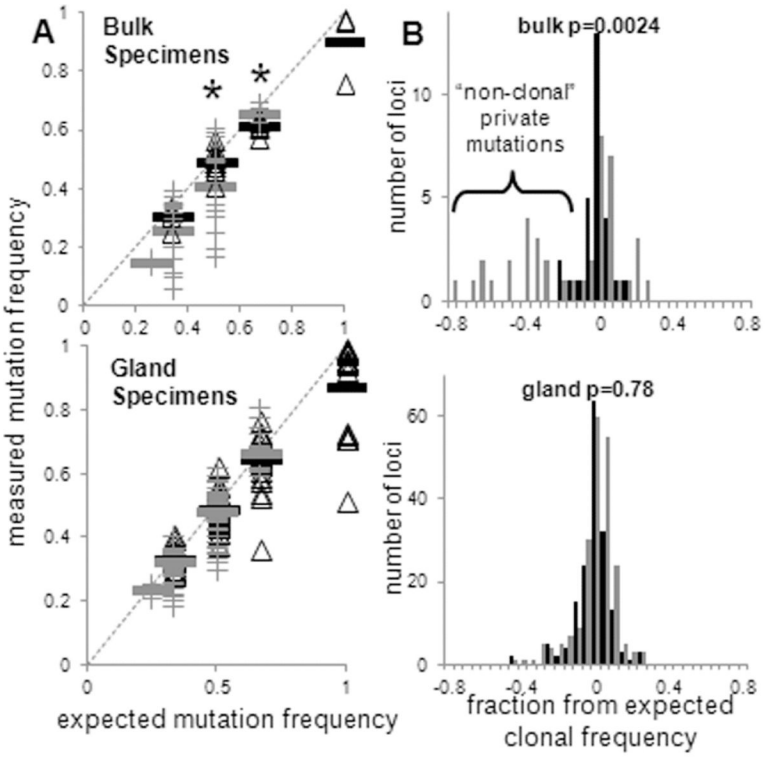

Figure 4. Clonal adenoma glands

A: Public mutations (black triangles) illustrate variations between measured and expected clonal MF values. Comparisons between public and private (grey crosses) MFs in the bulk specimens illustrate that private mutations were significantly less frequent (t-test) at the triploid and diploid loci (*), indicating that the private mutations are in only some tumor cells. However, individual gland public and private MFs were not statistically different (t-test) indicating that both public and private mutations (when present) were in all gland cells. **B:** Data plotted with respect to differences between expected and measured MFs at all loci reveal "bell shaped" distributions. In the bulk specimens, private mutations (grey bars) had significantly lower frequencies than public mutations (black bars), indicating that the private mutations are not in all tumor cells. In the gland specimens, public and private mutation frequencies are statistically indistinguishable, indicating that both public and private mutations (when present) are in all cells of a gland.

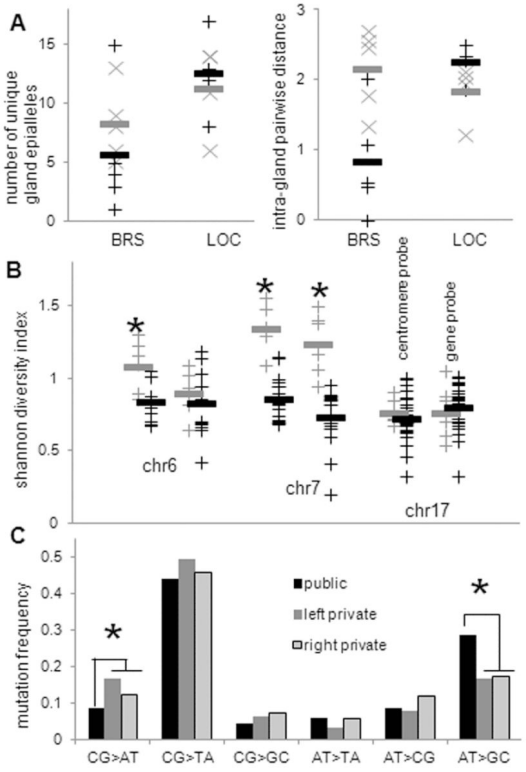

Figure 5. Diverse adenoma glands

A: Numbers of unique epiallele patterns (among 24 sampled epialleles) per gland were not significantly different (t-test) between the sides (grey left adenoma side, black right adenoma side). Intra-gland pairwise distances were not significantly different between the sides (t-test). **B:** FISH revealed chromosome CNs differed between cells within the same gland. Adenoma Shannon diversity index values (grey) were higher (more diverse) than normal colon crypt values (black) for five of the six FISH probes and significantly higher (t-test) for three probes (*). **C:** The mutation spectra of public exome mutations were significantly different from the private exome mutations ($p= 2.0\times10^{-7}$, chi-square test) with more C>A transversions and fewer T>C transitions. The differences between left and right private mutations were not significant (p=0.24).

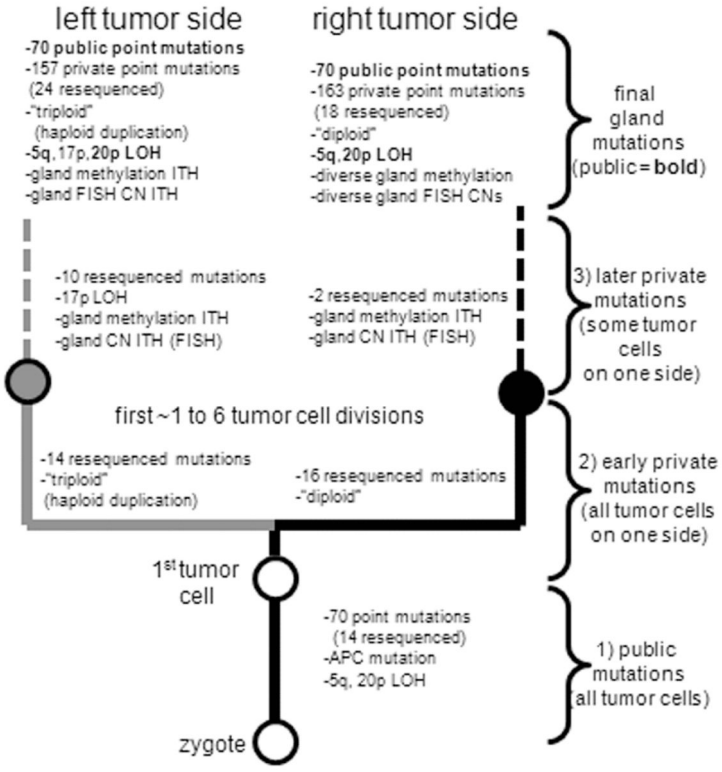

Figure 6. The first few inferred adenoma cell divisions

Potentially all of the early private mutations arose during the first few divisions, with genomic instability and a haploid duplication leading predominantly to triploidy on the left side.