CANCER

Minimal functional driver gene heterogeneity among untreated metastases

Johannes G. Reiter1,2,†, Alvin P. Makohon-Moore3,†, Jeffrey M. Gerold2,†, Alexander Heyde2, Marc A. Attiyeh2, Zachary A. Kohutek3, Collin J. Tokheim3, Alexia Brown3, Rayne M. DeBlasio3, Juliana Niyazov3, Amanda Zucker3, Rachel Karchin3,4, Kenneth W. Kinzler3,7,9, Christine A. Iacobuzio-Donahue5,10, Bert Vogelstein7,8,9,11, Martin A. Nowak2,12,‡

Metastases are responsible for the majority of cancer-related deaths. Although genomically heterogeneous within primary tumors is associated with relapse, heterogeneity among treatment-naïve metastases has not been comprehensively assessed. We analyzed sequencing data for 76 untreated metastases from 20 patients and inferred cancer phylogenies for breast, colorectal, endometrial, gastric, lung, melanoma, pancreatic, and prostate cancers. We found that within individual patients, a large majority of driver gene mutations are common to all metastases. Further analysis revealed that the driver gene mutations that were not shared by all metastases are unlikely to have functional consequences. A mathematical model of tumor evolution and metastasis formation provides an explanation for the observed driver gene homogeneity. Thus, single biopsies capture most of the functionally important mutations in metastases and therefore provide essential information for therapeutic decision-making.

The clonal evolution model of cancer proposes that cells accrue advantageous mutations and clonally expand so that these mutations are eventually present in all tumor cells (1–4). Recent studies reported mutations in putative driver genes that were only present in subpopulations of tumor cells (5, 6). The extent to which the acquisition of advantageous mutations continues after the initiation of the primary tumor (7 or during metastasis formation is unknown (8, 9). The growing list of putative driver genes and the increased sensitivity of next-generation sequencing have facilitated the discovery of subclonal driver gene mutations within a tumor (5, 10). Nevertheless, the evolutionary dynamics and the clinical importance of driver gene mutation heterogeneity in solid tumors are not fully understood.

Cells acquire a few mutations during each division because of imperfect DNA replication; hence, any population of cells is genetically heterogeneous (11). Because cancer cells continue to divide after cancer initiation, many new mutations are expected to be present in tumor subpopulations. However, to assess functional heterogeneity, advantageous mutations in putative driver genes must be distinguished from neutral replication errors in those genes. For example, within oncogenes, only few recurrently mutated positions are functional, and therefore, many mutations—even in driver genes—may not have important functional consequences. Moreover, although metastatic disease is responsible for most cancer-related deaths, the heterogeneity of driver gene mutations has predominantly been evaluated in primary tumors. Biopsies of metastatic lesions are not readily available and typically are acquired after exposure to toxic and mutagenic chemotherapies. These treatments can induce selective bottlenecks and confound the interpretation of genetic alterations. Because driver gene mutations increasingly inform clinical treatment decisions, undetected driver heterogeneity among metastases poses a barrier to the success of this precision medicine approach (12). If the founding cells of different metastases carry distinct driver gene mutations, disease progression and treatment could be fundamentally more complex than expected from a primary tumor biopsy alone. Additional driver gene mutations might be present in all or in a subset of metastases (Fig. 1). In both scenarios, more biopsies would be necessary for accurate diagnosis and optimal treatment. Here, we comprehensively analyzed the evidence for driver gene mutation heterogeneity among untreated metastases across cancer types. We also developed a mathematical model to determine the evolutionary mechanisms that give rise to intermetastatic driver mutation heterogeneity.

We analyzed data from 20 cancer patients for whom genome- or exome-wide sequencing was performed for at least two distinct treatment-naïve metastases (13–19). In total, we studied 115 samples, including 76 untreated metastases samples from diverse tissues (mean of 3.8 and median of 3 metastases per patient) (fig. S1 and table S1). We assessed somatic mutations of patients with pancreatic, endometrial, colorectal, breast, gastric, lung, melanoma, and prostate cancer (Fig. 2A). We classified nonsynonymous variants into putative driver and passengers mutations according to the The Cancer Genome Atlas consensus list of 299 putative driver genes (10). To allow for a consistent interpretation of driver gene mutation heterogeneity, we excluded two hypermutated subjects with more than 1000 nonsynonymous mutations and focused on the remaining 18 subjects. In these subjects, we found a median of 4.5 mutated driver genes (range 2 to 18) (Fig. 2A).

To determine the evolutionary timing of somatic mutations, we inferred cancer phylogenies and mapped all variants onto evolutionary trees (supplementary materials, materials and methods, and fig. S2) (20). We classified mutations into those present in all metastases (MetTrunk; hereafter referred to as “trunk”) and those present in a subset of metastases (MetBranch; hereafter referred to as “branch”) (Fig. 2B). We observed similar numbers of nonsynonymous or splice-site variants (hereafter referred to as nonsynonymous) in both categories (Fig. 2A). By contrast, trunks exhibited a twofold enrichment of the ratio of driver gene mutations to nonsynonymous mutations compared with branches (9.1 versus 4.0%; two-sided paired t test, \( P = 0.004 \)) (Fig. 3A). Nevertheless, we observed mutations in driver genes that were heterogeneous among metastases for 12 of 18 subjects.

To investigate whether heterogeneous mutations in putative driver genes were likely to be functional, we used a variety of approaches. We found that a large proportion of nonsynonymous variants in driver genes along trunks were previously detected at least once in other cancers [Catalogue Of Somatic Mutations In Cancer (COSMIC), 37.8%, 31 of 82], whereas a much smaller proportion along branches was present in COSMIC (15.6%, 5 of 32; two-sided Fisher exact test, \( P = 0.025 \)) (Fig. 3B). The fraction of driver gene mutations in branches in COSMIC was in fact similar to that of passenger gene mutations in either trunks or branches (14.1%, 128 of 905, and 12.5%, 89 of 712). Because mutations that are true drivers are often recurrent, we investigated how frequently identical

---

1Canary Center for Cancer Early Detection, Department of Radiology, Stanford University School of Medicine, Palo Alto, CA 94305, USA. 2Program for Evolutionary Dynamics, Harvard University, Cambridge, MA 02138, USA. 3David M. Rubenstein Center for Pancreatic Cancer Research, Human Oncology and Pathogenesis Program, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. 4Department of Radiation Oncology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA. 5Department of Biomedical Engineering, Institute for Computational Medicine, Johns Hopkins University, Baltimore, MD 21218, USA. 6Department of Oncology, Johns Hopkins University School of Medicine, Baltimore, MD 21224, USA. 7Ludwig Center, Johns Hopkins University School of Medicine, Baltimore, MD 21224, USA. 8Sidney Kimmel Cancer Center, The Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. 9Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY 10066, USA. 10Howard Hughes Medical Institute, The Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA. 11Department of Organismic and Evolutionary Biology and Department of Mathematics, Harvard University, Cambridge, MA 02138, USA. 12These authors contributed equally to this work.
nonsynonymous variants were found in COSMIC. Whereas variants in driver genes along trunks on average occurred in 0.32% COSMIC samples (occurrence mean of 82.0 in 25,516 COSMIC samples), driver gene mutations acquired along branches occurred more than 100-fold less frequently (0.0016%; Wilcoxon rank-sum test, \( P = 0.008 \)) (Fig. 3C).

We then used several methods to predict the functional impact of 1755 nonsynonymous variants along trunks and branches. We found that driver gene mutations acquired along trunks were more likely to have predicted functional consequences (Fig. 3, D to F, and fig. S3). Variants with the most likely protein-changing effects (mutation consequences with high impact, such as frameshift or nonsense mutations) were frequently observed in driver genes along trunks but rarely observed along branches (30.5 versus 6.3%; two-sided Fisher’s exact test, \( P = 0.006 \)) (Fig. 3D). The frequency of high-impact variants in driver genes along branches was no higher than that in passenger genes. FATHMM (27) predicted significantly stronger functional effects for driver gene mutations along trunks than along branches (mean scores of -2.1 versus 1.0; scores below -0.75 indicate likely driver mutation; two-sided Wilcoxon rank-sum test, \( P < 0.001 \)) (Fig. 3E). Similarly, CHASMplus (22) predicted significantly higher gene-weighted scores for driver gene mutations along trunks than along branches (mean scores 0.47 versus 0.16; higher values indicate likely functional effects; two-sided Wilcoxon rank-sum test, \( P < 0.001 \)) (Fig. 3F).

To identify the evolutionary determinants of intermetastatic heterogeneity, we developed a mathematical framework in order to assess how rates of growth, mutation, and dissemination give rise to driver gene mutation heterogeneity (supplementary text) (23, 24). The original clone in the primary tumor grows with a rate of \( r_0 = b_0 - d_0 \) per day (birth rate is \( b_i \) and death rate is \( d_i \) for each clone \( i \) and disseminates cells to distant sites with rate \( q_0 \) per day (Fig. 4A). When a cell divides, a daughter cell can acquire an additional driver mutation with probability \( u \). This model produces intermetastatic heterogeneity if not all detectable metastases were seeded from the same subclone in the primary tumor.

Following previously measured growth and selection parameters, we assume a growth rate of \( r_0 = 1.24\% \) per day and a relative growth advantage of a driver gene mutation of \( s = 0.4\% (s = b_i/b_0 - 1) \) (25, 26). To mimic the composition of our cohort, we considered the first four metastases that reach a detectable size of \( 10^6 \) cells (~1 cm³). We found that the probability of intermetastatic driver heterogeneity is 10.5% (\( d = 0.2475, q = 10^{-3} \)) (Fig. 4). The original founding clone of the primary tumor most likely seeds all detectable metastases (Fig. 1A, green cells). The increased growth rate conferred by a new driver mutation is insufficient to compensate for the time spent waiting for the driver mutation to occur (figs. S4 and S5).

The model reveals that the probability of observing intermetastatic driver heterogeneity increases when the primary tumor grows very slowly before metastases are seeded, the average growth advantage of additional driver mutations is very large, and the driver gene mutation rate is high (fig. S6C). By contrast, a high dissemination rate produces less intermetastatic heterogeneity because metastases are established before driver subclones greatly expand (Fig. 4E and fig. S7C). For very high driver growth advantages but slowly growing cancers, another scenario is possible: that all metastases are seeded from the same highly advantageous subclone (Fig. 1B). Last, if driver mutations instead increase the dissemination rate, an almost 10-fold increase is required to produce intermetastatic driver heterogeneity (Fig. 4F and fig. S8).

In real patients, we expect less intermetastatic heterogeneity for several reasons. First, driver gene mutations may not confer the same advantage in the microenvironment of the primary tumor and of a distant site, reducing the probability of heterogeneity (fig. S9). Second, primary tumor growth may slow down because of space or nutrient constraints or surgical removal, also reducing the expected intermetastatic heterogeneity (fig. S10). Third, advanced cancer cells have already acquired multiple driver gene mutations in various pathways, possibly reducing the number of additionally available driver gene mutations that confer a substantial selective advantage (fig. S6B).

Overall, we observed a depletion of heterogeneous mutations in putative driver genes among metastases (Fig. 3). Moreover, the majority of those that were observed had only weak or no predicted functional effects. These results are compatible with multiple recent studies on neurally evolving cancers after transformation (7, 27, 28). However, the mathematical framework demonstrates that a lack of intermetastatic driver heterogeneity does not imply neutral evolution but can also be explained by various other factors,
Fig. 2. Most mutations in putative driver genes occur on the trunk of metastases. (A) Twenty patients with 76 untreated metastases. Thirteen patients acquired mutations in putative driver genes along the MetBranch (MB), whereas seven did not. (B) Inferred phylogeny of a colorectal cancer exhibits intermetastatic driver mutation heterogeneity. Nonsynonymous mutations in driver genes are denoted in orange. Percentages denote branch confidence. Integers denote number of point mutations per branch. Table shows predicted functional effects of mutations in driver genes. Heterogeneous driver mutations were predicted to have no functional effect or were likely sequencing artifacts [low coverage and low variant allele frequency (VAF) across all sites]. MetTrunk (MT) denotes that variant was acquired on the trunk of all metastases. Sample origins, rectum, PT1-5; liver, Met1-6.

Fig. 3. Predicted functional mutations in putative driver genes are strongly enriched along metastases trunks. (A) Ratio of driver gene mutations to nonsynonymous mutations is enriched by twofold along trunks compared with branches. Orange diamond denotes mean, and black bar denotes median (two-sided paired t test, \( P = 0.004 \)). (B) Fraction of nonsynonymous variants in driver genes along MetTrunk in COSMIC was 38% compared with 16% along MetBranch (two-sided Fisher’s exact test, \( P = 0.025 \)). (C) Relative occurrence of variants in driver genes along MetTrunk in individual COSMIC samples was 0.32% compared with 0.0016% along MetBranch (two-sided Wilcoxon rank-sum test, \( P = 0.008 \)). (D) Variant Effect Predictor (VEP) inferred that 30% and 6% of driver gene mutations were of high impact along MetTrunk and MetBranch, respectively (two-sided Fisher’s exact test, \( P = 0.006 \)). (E and F) FATHMM (value below −0.75 indicates likely driver mutation) and CHASMPplus predicted increased functional consequences for variants in driver genes in MetTrunk. Two-sided Wilcoxon rank-sum tests were used. Thick black bars denote 90% confidence interval. No other statistically significant differences were observed. Numbers in brackets denote number of variants in each group. * \( P < 0.05 \); ** \( P < 0.01 \); *** \( P < 0.001 \).
Fig. 4. Mathematical analysis provides an explanation for intermetastatic driver gene mutation homogeneity or heterogeneity.

(A) Primary tumor expands stochastically from a single advanced cancer cell and seeds metastases. Cells of original clone (green) divide at rate \( b_0 \) and die at rate \( d \) per day. Additional driver mutations increase the birth rate to \( b_1 = b_0(1 + s) \), where \( s \) denotes the relative driver advantage \( b_1 \geq b_0, q_0 = q \); (B) to (E), or increase the dissemination rate \( q_i \geq q \). (B) Representative model realizations for typical parameter values. Growth rate \( r_0 = 1.24\% \) per day, \( s = 0.4\% \), and dissemination rate \( q_0 = 10^{-7} \) per cell per day.

including primary tumor growth dynamics (Fig. 4). Furthermore, growth rates may saturate and fitness gains of additional driver gene mutations become smaller because available resources (such as nutrients and oxygen) are already almost optimally utilized—a phenomenon that is observed in bacterial evolution (29).

Several limitations of this study should be noted. First, we exclusively focused on single-nucleotide variants and small insertions and deletions because their functionality can be predicted with multiple methods, and their heterogeneity has immediate clinical consequences for therapy selection (22). We did not assess recurrent noncoding, copy-number, or epigenetic alterations because functional prediction methods for them are not yet available. Second, we cannot exclude the possibility that mutations in yet-undiscovered driver genes of metastases are heterogeneous. Third, we could not evaluate micrometastases that are not visible clinically.

Because therapy selection and treatment success of previously untreated patients increasingly depends on the identification of genetic alterations, it will be critical to extend this analysis to larger cohorts and more cancer types in order to investigate whether minimal driver gene mutation heterogeneity is a general phenomenon of advanced disease. This pan-cancer analysis of untreated metastases suggests that a single biopsy accurately represents the driver gene mutations of a patient’s metastases.

REFERENCES AND NOTES

Acknowledgments

Funding: This work was supported by the National Institutes of Health grants K99CA229991 (J.G.R.), CA179991 (C.A.I.-D.), F33CA180682 (A.P.M.-M.), T32 CA60001-06 (A.P.M.-M.), F33CA200266 (C.J.T.), U24CA204887 (R.K.), CA43460 (B.V.), as well as by the Lustgarten Foundation for Pancreatic Cancer Research, The Virginia and D. K. Ludwig Fund for Cancer Research, an Erwin Schrödinger fellowship (J.G.R.; Austrian Science Fund FWF J-3996), a Landry Cancer Biology fellowship (J.M.G.), and the Office of Naval Research grant N00034-16-1-2914.


Competing interests: K.W.K. and B.V. are founders of Personal Genome Diagnostics. B.V. and K.W.K. are on the Scientific Advisory Board of Exelixis GP. These companies and others have licensed technologies from Johns Hopkins, and K.W.K. and B.V. receive equity or royalties from these licenses. The terms of these arrangements are being managed by Johns Hopkins University in accordance with its conflict of interest policies.

Data and materials availability: Accession nos. for the raw sequencing data are available in the original publications (13–18). Data of Brown et al. (18), Hong et al. (14), and Makohon-Moore et al. (17) as well as of subjects MSK1 and MSK2 are deposited at the European Genome-Phenome Archive (www.ebi.ac.uk/ega) and are available under accession nos. EGAS00001000760, EGAS00001000542, EGAS00001002186, and EGAS00001002777, respectively. Data of Gibson et al. (16) and Sanborn et al. (13) are deposited to the database of Genotypes and Phenotypes (dbGaP) at the National Center for Biotechnology Information (NCBI) under accession codes phs001127.v1.p1 and phs000941.v1.p1, respectively. Data of Kim et al. (15) are deposited to the Sequence Read Archive (SRA) at the NCBI under the project ID of PRJNA271316.

Supplementary materials

www.sciencemag.org/content/361/6406/1033/suppl/DC1

Materials and Methods
Supplementary Text
Figs. S1 to S32
Table S1
References (30–56)
27 March 2018; accepted 2 August 2018
10.1126/science.aat7171
Minimal functional driver gene heterogeneity among untreated metastases


Science 361 (6406), 1033-1037.
DOI: 10.1126/science.aat7171

Metastatic drivers same as primary
Treatment decisions for cancer patients are increasingly guided by analysis of the gene mutations that drive primary tumor growth. Relatively little is known about driver gene mutations in metastases, which cause most cancer-related deaths. Reiter et al. explored whether the growth of different metastatic lesions within an individual patient is fueled by the same or distinct gene mutations. In a study of 76 untreated metastases from 20 patients with different types of cancer, all metastases within a patient shared the same functional driver gene mutations. Thus, analysis of a single biopsy could help oncologists select the optimal therapy for patients with widespread metastatic disease. Science; this issue p. 1033