Genome organization in proximity to the BAP1 locus appears to play a pivotal role in a variety of cancers

Amit Sharma1 | Hongde Liu2 | Fabian Tobar-Tosse3 | Angela Noll4 | Tikam Chand Dakal5 | Huamei Li2 | Frank G. Holz1 | Karin U. Loeffler1 | Martina C. Herwig-Carl1

Abstract
Cancer studies primarily focus on the characterization of the key driver genes and the underlying pathways. However, the contribution of other cancer-associated genes located in the genomic neighborhood of the driver genes could help to understand further aspects of cancer progression. Given the frequent involvement of chromosome 3 in multiple human cancers, in particular in the form of the prognostically highly relevant monosomy 3 in uveal melanoma (UM), we investigated the cumulative impact of cancer-associated genes on chromosome 3. Our analysis showed that these genes are enriched with repetitive elements with genes surrounded by distinctive repeats (MIR, hAT-Charlie, ERVL-MaLR, LINE-2, and simple/low complexity) in the promoter being more precisely associated with cancer-related pathways than the ones with major transposable elements (SINE/Alu and LINE-1). Additionally, these genes showed strong intrachromosomal chromatin interactions in 3D nuclear organization. Further investigations revealed a genomic hotspot in the vicinity of BAP1 locus, which is affected in 27 types of different cancers and contains abundant noncoding RNAs that are often expressed in a tissue-specific manner. The cross-species comparison of these cancer-associated genes revealed mostly a shared synteny in closer primates. However, near to the BAP1 locus signs of chromosomal inversions were observed during the course of evolution. To our knowledge, this is the first study to characterize the entire genomic neighborhood of cancer-associated genes located on any single chromosome. Based on our results, we hypothesize that monosomy of chromosome 3 will have important clinical and molecular consequences in the respective diseases and in particular in UM.

Keywords
cancer, evolution, monosomy 3, repetitive sequence, uveal melanoma
INTRODUCTION

Over the years, it has been established that the accumulation of genetic and epigenetic alterations contribute significantly to cancer progression. The extensive sequencing efforts and subsequent computational methods have led to the discovery of several cancer-associated driver genes. However, most studies seem to agree that the cancer genome is not limited to single deletions, mutations, or chromosomal rearrangements but reflects the cumulative impact of multiple dysregulated pathways. Although tremendous progress has been made to define the cancer genome, the significance of cancer-associated genes in the genomic neighborhood of clinically relevant genes has not been discussed in detail yet. Given the frequent involvement of (the short arm of) chromosome 3 in multiple human cancers, it has been proposed that some regions could harbor multiple tumor suppressor genes. The contribution of monosomy 3 in uveal melanoma (UM), the most frequent primary intraocular tumor in Caucasian adults, was recognized nearly three decades ago and represents, along with other factors, a predictive factor for the development of systemic metastasis (reviewed in Sharma et al). Interestingly, BAP1 (BRCA1-associated protein 1) is frequently mutated in metastatic UM, but also plays a significant role in pleural mesothelioma, lung adenocarcinoma, and renal cell carcinoma. Apart from BAP1, other genes at chromosome 3, such as SETD2 (renal cell carcinoma), ADAMTS9 (esophageal squamous cell carcinoma), SOX2, ECT2, PRKCI, PIK3CA (lung squamous cell carcinoma), and ROBO1 (breast cancer and colorectal cancer) show genetic alterations in cancer.

Considering the contribution of chromosome 3 in multiple types of cancer, herein, we investigated all known cancer-associated genes on this particular chromosome. We provide insights about the repetitive elements surrounding their promoter regions and defined the evolutionary and clinical aspect of the hotspot region that was identified in the vicinity of the BAP1 locus.

MATERIAL AND METHODS

The information with regard to cancer genes was retrieved from a publically available database (www.cancer-genetics.org). The selection criteria were based on the information of known mutation types (somatic/germline) and the methylation status while deletions and translocations were excluded. Based on the number of citations supporting each gene from the database (http://bionlp.bcgsc.ca/cancermine/), we could identify few genes as driver genes (n = 6), oncogenes (n = 27), and tumor suppressor genes (n = 22). To generate significant pathways the web tool Metascape was applied. Human 3D chromatin interaction data were downloaded from 4Dgenome, (https://4dgenome.research.chop.edu/Tables/4DGenome_HomoSapiens_hg19.txt). Briefly, the number of chromatin interactions linked between genes were counted and plotted with tool circos. In the plot, line color darkness is proportional to the number of interactions.

3 | RESULTS

3.1 | Chromatin interaction network in 3D nuclear genome

Our database search resulted in the identification of 130 cancer-associated genes dispersed over chromosome 3 (Figure 1A and Table S1). The experimentally determined dataset showed an extensive interaction network between these genes, which suggests that these genes might transcribe synchronously (Figure 1B). Specifically, genes such as RASSF1, SEMA3F, NPR2, HYAL1/2, GNA12, and TUSC2 showed extensive chromatin interactions (Figure 1B). These genes were found to be located together in genomic hotspot region 2 of chromosome 3 (introduced below). However, the majority of other cancer-related genes showed weaker chromatin interactions, suggesting a mutual control of epigenetic regulation. The enrichment analysis for all cancer genes revealed a significant involvement of...
pathways related to tumorigenesis, cancer development, and cancer metastasis, which is expected considering their direct role in cancers (Figure 1C).

3.2 Configuration of repetitive elements surrounding genes

We identified specific repeats and configurations in the promoter regions of cancer-associated genes on chromosome 3 (Figure 2A and Table S2). Among them, 63 of 124 analyzed genes showed at least 2 repetitive types of elements in the promoter while 32 genes showed single/unique repetitive elements. We also identified 29 genes that lacked any kind of repeat sequence (Figure 2B). We investigated the genes without repeats with GO categories and found that 3 genes (CTNNB1, PIK3CA, and MLH1) were highly coassociated with “cancer”, whereas genes like SEMA3B, ROBO1, and EPHA3 were significantly coassociated with “axon development” (Figure 2C and Table S2). Likewise, the analysis of the genes with repeats in the promoter showed an association not only with cancer-associated pathways but also with the categories of metabolism, development, and immune system pathways (Figure 2D). Briefly, this analysis indicates that, although these genes harbor cancer potential, the repetitive sequences around their promoter might influence their role in other significant pathways that are not necessarily associated with cancer.

3.3 Characterization of hotspot regions and associated long noncoding RNAs

We identified chromosomal region 3p21-3p21.31 as a hotspot for cancer-related genes, including BAP1. Based on the information from respective databases, the genes located in the hotspot region of chromosome 3 were apparently found to be associated with 27 (out
of 72 identified) different types of cancer (Figure 3A). Interestingly, most of the genes were found to be involved in the regulation of breast and lung cancer. Based on the genomic distances between them, we could further define 3 subgroups within this major hotspot region. Interestingly, most of the genes (9/10) from hotspot region 2 showed high protein-protein interaction affinity in the STRING analysis (Figure S1). As mentioned above, this region contains the genes RASSF1, SEMA3F, NPRL2, HYAL1/2, GNAI2, and TUSC2, which showed extensive chromatin interactions. With regard to hotspot 3, the three chromatin regulators BAP1 (histone deubiquitinating enzyme), SETD1 (methyltransferase for H3K36me3), and PBRM1 (subunit of the SWI/SNF chromatin remodeling complex) also showed closed interactions in the STRING analysis (Figure S1). We further investigated and localized several long noncoding RNAs (lncRNA) in this hotspot region (Figure 3B and Table S3). However, their potential role to modulate the activity of these genes is unclear.

3.4 Orthologues in closer primates and inversions near to BAP1

The investigations about the evolutionary spectrum of these cancer-associated genes revealed orthologues in closer primates such as chimpanzee, gorilla, orangutan, rhesus macaque, and baboon (Figures 3C and S2, Table S4). Although orthologous genes revealed a shared synteny, we observed inversions in the vicinity of hotspot regions in lower hominoids (Figure 3C). Specifically, some genes on hotspot region 1 (CCR1 and CCR3) in chimpanzee and gorilla were found to be inverted in comparison to the human. In addition, the CCK gene in gorilla was found at another chromosome. In baboons, 7 of these cancer-related genes were found to be located at another chromosome (chr11). In orangutan, all genes were at the same chromosome, but blocks of them did not share the same order as in human (also described in Tsend-Ayush et al26). Importantly, BAP1 was found within a bigger block, which was found to be disturbed by the integration of 4 cancer-related genes (CDCP1, CD86, CDC25A, and CCR3) (Figures 3C and S2). Rhesus macaque and baboon also showed inversions and rearrangements for these genes; moreover, their variations were comparable due to their phylogenetic relationship.

A shared synteny is a sign of shared gene order, which also codes for the positional orthology in different species.27 Herein, we used the GOC scores provided by Ensembl (www.ensembl.org), which were calculated based on the available pairwise whole-genome alignments to compute the percentage of genes lying in synteny blocks between human and primates. The GOC score indicates how many of the 4 closest neighboring genes shows orthologs compared.
to another species (https://www.ensembl.org/info/genome/compara/Ortholog_qc_manual.html). In our analysis, we found all percentages lying in a similar range among human and primate species with approximately 10% of orthologous genes showing no synteny in at least one human-primate comparison (Table S5). To determine how often carcinoma diseases correlate with evolutionary disrupted synteny blocks, we further investigated all those genes (showing no synteny in at least 4 human/primate comparisons) for enrichment analysis (www.webgestalt.org). We identified 4 cancer-related enriched disease pathways in the top significant hits (Table S6). Hence, we could speculate that the disruption of synteny could play a role in the development of cancer.

4 | DISCUSSION

The alteration of gene expression levels in sex chromosomal aneuploidies has previously been determined. Likewise, in autosomes, the complete/partial loss of chromosome 3 (such as in UM and the 3p deletion syndrome) might also reflect a similar response to gene dosage. In other words, a lower dosage of the chromosome 3-associated 1085 protein coding genes can be assumed in UM under monosomy 3 conditions. The importance of monosomy 3 in UM pathogenesis is also emphasized by the observation that patients with a BAP1 tumor predisposition syndrome who have a high risk to develop several types of cancer, show typically in addition to the germline BAP1 mutation a loss of 1 copy of chromosome 3 in UM. To understand the impact of this genomic imbalance, we first investigated the genomic location of 130 cancer-associated genes at chromosome 3. As the chromosomes are spatially distributed in a nonrandom manner, we also checked the spatial positioning patterns of these genes based on 3D genome organization studies. The analysis showed a chromatin interaction network between these genes.

Next, we checked the prevalence of repetitive DNA sequence “repeats” in their promotor regions of these genes. Interestingly, the genes with repeats like MIR, hAT, ERVL, LINE/L2, simple repeat, and low complexity repeats showed a more precise association to cancer-related pathways than SINE/Alu and LINE/L1 repeats. Hence, the differentially identified signature of repetitive sequences suggests a shared and unique molecular feature associated with repetitive elements and the cancer genome. The prevalence of repetitive sequences in the functional parts of genomes and their association with human diseases remains undisputed. Whether the absence of these repeats in monosomy of chromosome 3 in UM tumors or the 3p deletion syndrome is compensated by additional factors or alternative genomic regions needs further elucidation. Whether the loss of other constituents, such as microRNAs (miRNAs) or lncRNAs, at chromosome 3 also affects the chromatin conformation, which could in turn contribute to or against the cancer development, is yet to be determined.
In our analysis, we highlighted the chromosomal region 3p21-3p21.31 as a hotspot for cancer-related genes, which appears to be involved in 27 different types of cancer. Based on the genomic distances between them, we could further define 3 subgroups within this major hotspot region. Interestingly, hotspot region 2 showed a high protein-protein interaction affinity and extensive chromatin interactions. Previously, 1 study using breast cancer cell lines discussed the potential involvement of epigenetic repression for the downregulation of a few genes located in this hotspot region, hence indicating the possibility of peculiar genome organization around these genes. However, no partial deletion covering the entire hotspot region has been reported to date. Our analysis also suggests that chromatin organization in this hotspot region could play an important role explaining why some genomic sites are prone to get more frequent mutations. This is also evident from the clinical spectrum as BAP1 is the only gene of the hotspot region that is known to be mutated in UM, whereas genes such as BAP1, PBRM1, and SETD2 are frequently involved in renal cell carcinoma. Recently, we showed the altered expression of the BAP1 gene in 29 cancer types. In addition, we showed that C-terminal BAP1 mutations have a stronger intrinsic control to regulate the miRNA network and the binding of subsequent protein complexes.

In the current study, we also investigated the genomic characteristic of these cancer-associated genes in closer primates (chimpanzee, gorilla, orangutan, rhesus macaque, and baboon) and observed a shared synteny (i.e., preserved distribution of genes on chromosomes of different species). Interestingly, inversions in the vicinity of the hotspot region in lower hominoids were also found. Apart from this, some genes were also found to be duplicated during evolution, as previously described. Due to widespread genomic rearrangements within the genome of gibbons, it was cautiously not incorporated into this study. Overall, these results indicate that these genes are always clustered together in closer primates. Although inversions happened in lower hominoids (orangutan, baboon, and rhesus), the genetic closeness remained intact from gorilla and chimpanzee up to humans. Whether these genes are also organized at the same chromosome in distant primates and why speciation did not distribute them on different chromosomes in closer primates need further investigation. Importantly, how the closer primates are protected from cancer, despite having such potentially active orthologue genes as in humans, paves the way for further scientific discussion.

Taken together, we could show that the cancer-associated genes on chromosome 3 are enriched with repetitive elements and show strong intrachromosomal chromatin interactions. The genomic hotspot in the vicinity of the BAP1 locus appears to be involved in multiple cancers. Cross-species comparison revealed a shared synteny of these genes and inversions near the hotspot region in closer primates. These results support the hypothesis that UM in particular (but also other types of cancer) is affected by important clinical and molecular consequences due to monosomy of chromosome 3.

ACKNOWLEDGMENTS
This work was funded by the German Research Foundation (DFG; HE 5775/5-1).

DISCLOSURE
The authors have no conflict of interests.

ORCID
Amit Sharma  https://orcid.org/0000-0002-2216-5389
Martina C. Herwig-Carl  https://orcid.org/0000-0002-5943-7675

REFERENCES
1. Kok K, Nayloy SL, Buys CH. Deletions of the short arm of chromosome 3 in solid tumors and the search for suppressor genes. Adv Cancer Res. 1997;71:27-92.
2. Tschentscher F, Prescher G, Horsman DE, et al. Partial deletions of the long and short arm of chromosome 3 point to two tumor suppressor genes in uveal melanoma. Cancer Res. 2001;61:3439-3442.
3. Zaborovsky ER, Lerman MI, Minja JD. Tumor suppressor genes on chromosome 3p involved in the pathogenesis of lung and other cancers. Oncogene. 2002;21:6915-6935.
4. Hesson LB, Cooper WN, Latif F. Evaluation of the 3p21.3 tumor-suppressor gene cluster. Oncogene. 2007;26:7283-7301.
5. Ji L, Minna JD, Roth JA. 3p21.3 tumor suppressor cluster: prospects for translational applications. Future Oncol. 2005;1:79-92.
6. Horsman DE, Sroka H, Rootman J, White VA. Monosomy 3 and isochromosome 8q in a uveal melanoma. Cancer Genet Cytogenet. 1990;45:249-253.
7. Prescher G, Bornfeld N, Becher R. Nonrandom chromosomal abnormalities in primary uveal melanoma. J Natl Cancer Inst. 1990;82:1765-1769.
8. Sharma A, Stei MM, Frohlich H, Holz FG, Loeffler KU, Herwig-Carl MC. Genetic and epigenetic insights into uveal melanoma. Clin Genet. 2018;93:952-961.
9. Harbour JW, Onken MD, Roberson EDO, et al. Frequent mutation of BAP1 in metastasizing uveal melanomas. Science. 2010;330:1410-1413.
10. Herwig-Carl MC, Sharma A, Moulin A, Strack C, Loeffler KU. BAP1 immunostaining in uveal melanoma: potentials and pitfalls. Ocul Oncol Pathol. 2018;4:297.
11. Herwig-Carl MC, Sharma A, Holler T, Holz FG, Schlitter AM, Loeffler KU. Spatial intratumor heterogeneity in uveal melanoma: tumor cell subtypes with a presumed invasive potential exhibit a particular epigenetic staining reaction. Exp Eye Res. 2019;182:175-181.
12. Melzer C, Sharma A, Peters S, et al. Basal cell carcinomas developing independently from BAP1-tumor predisposition syndrome in a patient with bilateral uveal melanoma: diagnostic challenges to identify patients with BAP1-TPDS. Genes Chromosomes Cancer. 2019;58:357-364.
13. Piva F, Santoni M, Matrana MR, et al. BAP1, PBRM1 and SETD2 in clear-cell renal cell carcinoma: molecular diagnostics and possible targets for personalized therapies. Expert Rev Mol Diagn. 2015;15:1201-1210.
14. Lo PHY, Leung ACC, Kwok CYC, et al. Identification of a tumor suppressive critical region mapping to 3p14.2 in esophageal squamous cell carcinoma and studies of a candidate tumor suppressor gene, ADAMTS9. Oncogene. 2007;26:148-157.
15. Fields AP, Justilien V, Murray NR. The chromosome 3q26 OncCassette: a multigenic driver of human cancer. Adv Biol Regul. 2016;60:47-63.
16. Dallol A, Forgacs E, Martinez A, et al. Tumour specific promoter region methylation of the human homologue of the Drosophila
Roundabout gene DUTT1 (ROBO1) in human cancers. Oncogene. 2002;21:3020-3028.

17. Villacís RAR, Abreu FB, Miranda PM, et al. ROBO1 deletion as a novel germline alteration in breast and colorectal cancer patients. Tumour Biol. 2016;37:3145-3153.

18. Tobar-Tosse F, Velez PE, Ocampo-Toro E, Moreno PA. Structure, clustering and functional insights of repeats configurations in the upstream promoter region of the human coding genes. BMC Genom. 2018;19:862.

19. Durinck S, Moreau Y, Kasprzyk A, et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics. 2005;21:3439-3440.

20. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc. 2009;4:1184-1191.

21. Kent WJ, Sugnet CW, Furey TS, et al. The human genome browser at UCSC. Genome Res. 2002;12:996-1006.

22. Hinrichs AS, Karolchik D, Baertsch R, et al. The UCSC Genome Browser database: update 2006. Nucleic Acids Res. 2006;34:D590-D598.

23. Krzywinski M, Schein J, Birol I, et al. Circos: an information aesthetic for comparative genomics. Genome Res. 2009;19:1639-1645.

24. Guy L, Kultima JR, Andersson SG. genoPlotR: comparative gene and genome visualization in R. Bioinformatics. 2010;26:2334-2335.

25. Smedley D, Haider S, Ballester B, et al. BioMart–biological queries made easy. BMC Genom. 2009;10:22.

26. Tsind-Ayush E, Grützner F, Yue Y, et al. Plasticity of human chromosome 3 during primate evolution. Genomics. 2004;83:193-202.

27. Dewey CN. Positional orthology: putting genomic evolutionary relationships into context. Brief Bioinform. 2011;12:401-412.

28. Sharma A, Jamil MA, Nuesgen N, et al. DNA methylation signature in peripheral blood reveals distinct characteristics of human X chromosome numerical aberrations. Clin Epigenetics. 2015;7:76.

29. Ewens KG, Lalonde E, Richards-Yutz J, Shields CL, Ganguly A. Comparison of germline versus somatic BAP1 mutations for risk of metastasis in uveal melanoma. BMC Cancer. 2018;18:1172.

30. Maerker DA, Zeschnigk M, Nelles J, et al. BAP1 germline mutation in two first grade family members with uveal melanoma. Br J Ophthalmol. 2014;98:224-227.

31. Dekker J, Belmont AS, Guttmann M, et al. The 4D nucleome project. Nature. 2017;549:219-226.

32. Sharma A, Jamil MA, Nuesgen N, et al. Detailed methylation map of LINE-1 5’-promoter region reveals hypomethylated CpG hotspots associated with tumor tissue specificity. Mol Genet Genomic Med. 2019;7:e601.

33. Sharma A, Kaut O, Pavlova A, et al. Skewed X-chromosome inactivation and XIST locus methylation levels do not contribute to the lower prevalence of Parkinson’s disease in females. Neurobiol Aging. 2017;57(248):e1-e5.

34. da Costa PE, Cavalli LR, Rainho CA. Evidence of epigenetic regulation of the tumor suppressor gene cluster flanking RASSF1 in breast cancer cell lines. Epigenetics. 2011;6:1413-1424.

35. Sharma A, Biswas A, Liu H, et al. Mutational landscape of the BAP1 locus reveals an intrinsic control to regulate the miRNA network and the binding of protein complexes in uveal melanoma. Cancers. 2019;11:1600.

36. Carbone L, Alan Harris R, Gnerre S, et al. Gibbon genome and the fast karyotype evolution of small apes. Nature. 2014;513:195-201.

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Sharma A, Liu H, Tobar-Tosse F, et al. Genome organization in proximity to the BAP1 locus appears to play a pivotal role in a variety of cancers. Cancer Sci. 2020;111:1385-1391. https://doi.org/10.1111/cas.14319