

1 Globally altered epigenetic landscape and lagging osteogenic differentiation 2 in H3.3-G34W-mutant giant cell tumor of bone

3
4 Pavlo Lutsik^{1,18}, Annika Baude^{1,18}, Daniela Mancarella^{1,2,18}, Simin Öz^{1,18}, Alexander Kühn^{1,2},
5 Reka Toth¹, Joschka Hey^{1,2}, Umut H. Toprak^{3,4}, Jinyeong Lim⁵, Viet Ha Nguyen⁵, Chao Jiang⁶,
6 Anand Mayakonda^{1,2}, Mark Hartmann⁷, Felix Rosemann¹, Kersten Breuer¹, Dominik
7 Vonficht^{2,12,13}, Florian Grünschlager^{2,12,13}, Suman Lee⁵, Maren Kirstin Schuhmacher⁸, Denis
8 Kusevic⁸, Anna Jauch⁹, Dieter Weichenhan¹, Jozef Zustin¹⁶, Matthias Schlesner⁴, Simon
9 Haas^{12,13}, Joo Hyun Park¹¹, Yoon Jung Park¹¹, Udo Oppermann^{6,14}, Albert Jeltsch⁸, Florian
10 Haller¹⁰, Jörg Fellenberg¹⁵, Anders M. Lindroth^{5,*}, Christoph Plass^{1,17*}

11 Affiliations

- 12 1. Division of Cancer Epigenomics, German Cancer Research Center (DKFZ), 69120
13 Heidelberg, Germany
- 14 2. Faculty of Biosciences, Ruprecht-Karls-University of Heidelberg, 69120 Heidelberg,
15 Germany
- 16 3. Division of Theoretical Bioinformatics, German Cancer Research Center (DKFZ), 69120
17 Heidelberg, Germany
- 18 4. Bioinformatics and Omics Data Analytics, German Cancer Research Center (DKFZ), 69120
19 Heidelberg, Germany
- 20 5. Graduate School of Cancer Science and Policy, Cancer Biomedical Science, National Cancer
21 Center, Goyang-si Gyeonggi-do, 10408, Republic of Korea, Republic of Korea
- 22 6. Botnar Research Centre, Oxford NIHR BRC, Nuffield Department of Orthopaedics,
23 Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford OX3 7LD, UK
- 24 7. Section Translational Cancer Epigenomics, Division of Translational Medical
25 Oncology, National Center for Tumor Diseases (NCT) & German Cancer Research Center
26 (DKFZ), Heidelberg, Germany
- 27 8. Department of Biochemistry, Institute of Biochemistry and Technical Biochemistry,
28 University of Stuttgart, 70049 Stuttgart, Germany
- 29 9. Institute of Human Genetics, Ruprecht-Karls-University of Heidelberg, 69120 Heidelberg,
30 Germany
- 31 10. Institute of Pathology, University Hospital Erlangen, Friedrich-Alexander University
32 Erlangen- Nürnberg, 91054 Erlangen, Germany
- 33 11. Department of Nutritional Science and Food Management, Ewha Womans University, Seoul,
34 03760, Republic of Korea
- 35 12. Division of Stem Cells and Cancer, German Cancer Research Center (DKFZ), 69120
36 Heidelberg, Germany
- 37 13. Heidelberg Institute for Stem Cell Technology and Experimental Medicine – HI-STEM
38 gGmbH, 69120 Heidelberg, Germany
- 39 14. FRIAS- Freiburg Institute of Advanced Studies, University of Freiburg, Freiburg, Germany
- 40 15. Department of Experimental Orthopaedics, Orthopaedic University Hospital Heidelberg,
41 Ruprecht-Karl University of Heidelberg, 69118 Heidelberg, Germany
- 42 16. Institute of Pathology, University Medical Center Hamburg-Eppendorf, 20251 Hamburg,
43 Germany
- 44 17. German Consortium for Translational Cancer Research (DKTK)
- 45 18. These authors contributed equally

46 (*) Corresponding authors

47
48
49 ~~Christoph Plass (plass@dkfz.de), Anders Lindroth (lindroth@oc.mh)~~ reports new research that has not been certified by peer review and should not be used to guide clinical practice.

50 **Abstract**

51 The neoplastic stromal cells of giant cell tumor of bone (GCTB) carry a mutation in *H3F3A*,
52 leading to a mutant histone variant, H3.3-G34W, as a sole recurrent genetic alteration. We show
53 that in patient-derived stromal cells H3.3-G34W is incorporated into the chromatin and
54 associates with massive epigenetic alterations on the DNA methylation, chromatin accessibility
55 and histone modification level that can be partially recapitulated in an orthogonal cell line
56 system by the introduction of H3.3-G34W. These epigenetic alterations affect mainly
57 heterochromatic and bivalent regions and provide possible explanations for the genomic
58 instability as well as the osteolytic phenotype of GCTB. The mutation occurs in differentiating
59 mesenchymal stem cells and associates with an impaired osteogenic differentiation. We
60 propose that the observed epigenetic alterations reflect distinct differentiation stages of H3.3
61 WT and H3.3 MUT stromal cells and add to H3.3-G34W associated changes.

62 **Important abbreviations**

63 H3.3-G34W, mutated histone variant; H3.3 MUT, stromal cells expressing H3.3-G34W;
64 H3.3 WT, stromal cells expressing wildtype H3.3

65 Introduction

66 The discovery of mutated histone genes in aggressive cancers raised a lot of interest in the
67 cancer research community due to their ability to globally alter the epigenomic landscape (1).
68 A frequently mutated histone is the non-canonical histone variant H3.3 (2,3). In contrast to
69 canonical H3.1 and H3.2, the incorporation of histone variant H3.3 is replication-independent
70 and its turnover occurs throughout the cell cycle (4). Deposition occurs either via the HIRA
71 chaperone complex at sites of gene activation or through ATRX-DAXX into heterochromatic
72 regions (5) and the silent allele of imprinted genes (6). Mouse embryonic stem cells require
73 H3.3 for correct establishment of H3K27me3 patterns at bivalent promoters of developmentally
74 regulated genes (7,8). While two human genes, *H3F3A* and *H3F3B*, encode for an identical
75 H3.3 protein, oncogenic mutations occur gene-specifically in different tumor types. Lysine-27-
76 to-methionine (H3.3-K27M) and glycine-to-arginine or valine substitution (H3.3-G34R/V) in
77 pediatric gliomas (9), as well as glycine-34-to-tryptophan or leucine (H3.3-G34W/L)
78 substitutions in giant cell tumor of bone (GCTB) (10) have been described, all due to mutations
79 in *H3F3A*. For *H3F3B*, mutations leading to a lysine -36-to-methionine (H3.3-K36M)
80 substitution were reported in chondroblastomas (10,11). The molecular consequence of the
81 H3.3-K27M mutation is a global loss of the repressive chromatin mark H3K27me3 through
82 inactivation of the PRC2 complex (12-14). Similarly, the H3.3-K36M mutation suppresses the
83 deposition of the H3K36me3 mark by interference with histone methyltransferases NSD2 and
84 SETD2 (11,15). Recent findings suggested reduced levels of H3K36me3 and increased levels
85 of H3K27me3 *in cis* in HeLa cells overexpressing H3.3-G34W (16). However, the detailed
86 effects of this mutant histone variant on the epigenome are yet to be determined and to be
87 analyzed in patients. GCTB, where this mutation was shown as the sole alteration, offers a
88 unique system to study these effects in primary patient material.

89 GCTB is a rare locally aggressive bone neoplasm, which typically affects the meta-epiphyseal
90 regions of long bones in young adults (17). These tumors consist of three major cell types:
91 stromal cells originating from mesenchymal stem cells (MSC), multinuclear giant cells and
92 mononuclear histiocytic cells (18). The GCTB stromal cells show incidence of H3.3-G34W in
93 more than 90% of cases and display markers of both MSC and pre-osteoblast cell populations
94 (10,19). The neoplastic stromal cell population secretes high levels of the receptor activator of
95 NF- κ B ligand (RANKL) and reduced levels of its decoy receptor, osteoprotegerin (OPG),
96 thereby attracting and activating surrounding monocytes. Upon activation, the recruited

97 monocytes fuse to form multinucleated giant cells, which resemble osteoclasts and lead to
98 massive bone destruction (17).

99 Here we investigate the effects of H3.3-G34W on global epigenomic patterns in patient samples
100 from four different centers. We find epigenetic distortions that contribute to the phenotypes of
101 GCTB, stochastic genomic instability and increased osteolysis. Furthermore, we demonstrate
102 that neoplastic and non-neoplastic GCTB stromal cells represent distinct stages of osteogenic
103 differentiation. Differentiation-related epigenetic differences add to the overall picture of H3.3-
104 G34W-associated global epigenetic alterations, whereas the differentiation delay is potentially
105 driven by the direct effects of H3.3-G34W. Our findings collectively suggest that the single
106 alteration of G34W induces epigenomic changes with implications for the development of
107 stromal cells and the tumorigenic process.

108 **Results**

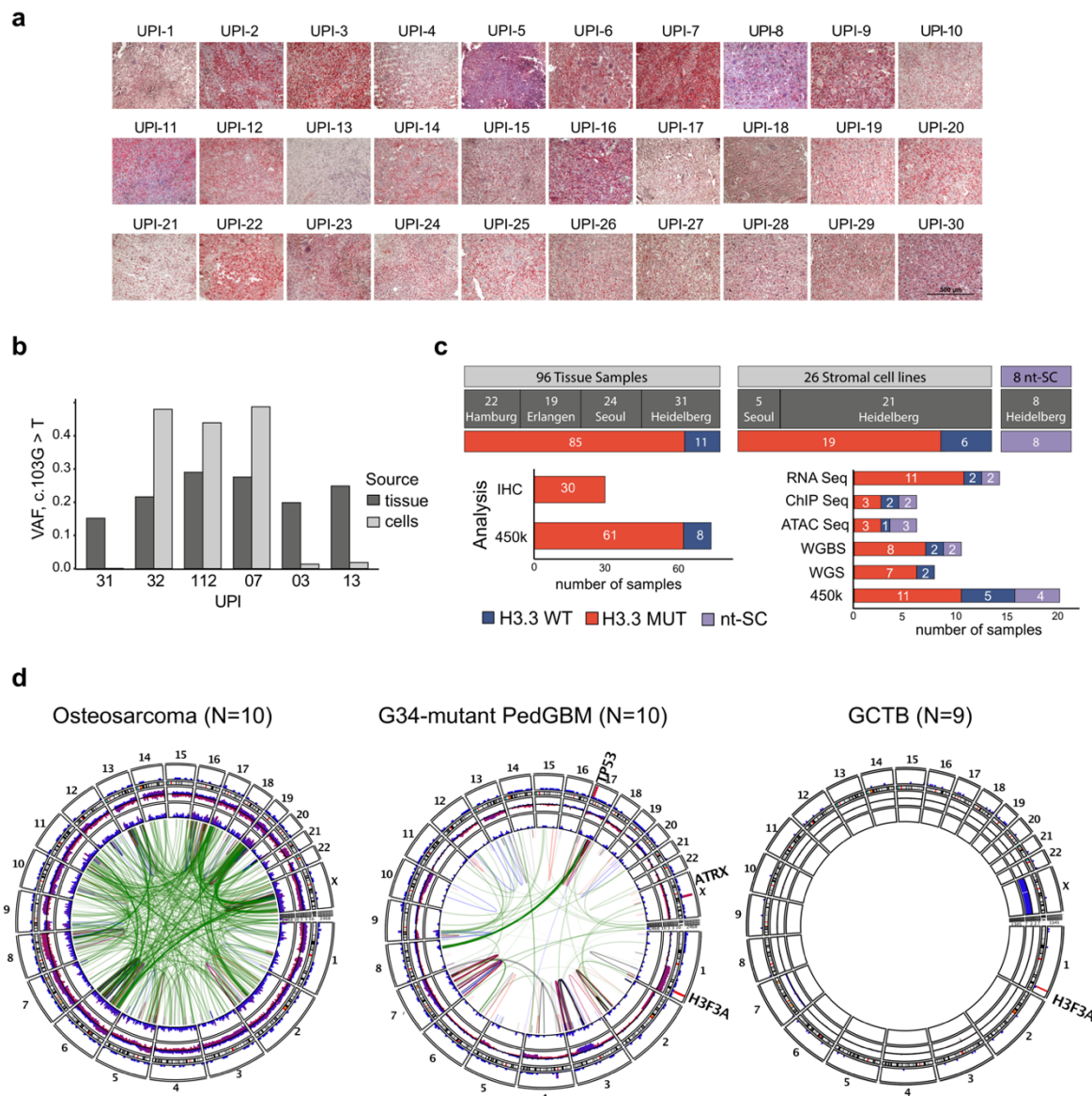
109 ***H3K36 methylation levels are not changed in H3.3-G34W-expressing stromal cells from*** 110 ***GCTB patients***

111 Recent biochemical studies have shown that G34 substitutions in H3.3, including G34W, inhibit
112 the activity of the histone methyltransferase SETD2, which is responsible for H3K36me₃, *in*
113 *cis* (16). We verified these effects in HEK293 cells stably overexpressing H3.3-G34W or wild-
114 type H3.3 as a control and confirmed *in cis* effects on H3.3-G34W K36 trimethylation levels.
115 We did not observe any *in trans* effects on endogenous H3 modifications as found for other
116 mutant histones such as H3-K36M (**Figure S1a**).

117 To specifically test if these biochemical findings apply to patient samples, we obtained access
118 to GCTB biopsies from four different cohorts (**Table S1**). For the initial characterization of 30
119 GCTB samples (**Table S1**), we performed immunohistochemical analysis with a H3.3-G34W-
120 specific antibody. Positive staining was observed and validated in 29 of 30 cases (**Fig. 1a** and
121 **Fig. S1b**). The H3.3-G34W-negative case (unified patient identifier, UPI-13) carried a *H3F3A*
122 (c.103_104GG>TT) mutation encoding a H3.3-G34L substitution that has already been
123 described for GCTB (10) (**Fig. 1b, S1c**). We established both, neoplastic, H3.3 G34W-
124 expressing (H3.3 MUT) and non-neoplastic, H3.3-G34W negative (H3.3 WT) stromal cell lines
125 (**Fig. 1b; Table S1**), from primary tissue. Cell type differences were ruled out by flow
126 cytometric analyses which revealed a high expression of MSC markers and low expression of
127 hematopoietic markers in both H3.3 WT and H3.3 MUT cell lines, suggesting that both are of

128 mesenchymal origin (**Fig. S1d**). In total, we collected 96 tissue samples from 95 different
129 GCTB patients from four different cohorts (**Table S1**). We were able to establish 26 stromal
130 cell lines from 24 different GCTB patients from two cohorts (**Fig. 1c**). In addition to H3.3 WT
131 cells, we analyzed bone marrow-derived primary MSCs from non-GCTB patients from here on
132 referred to as nontumoral stromal cells (nt-SC) (**Table S1**). We verified the mutational status
133 of the cells using Sanger resequencing, ultra-deep resequencing on the MiSeq platform or
134 whole-genome sequencing (**Table S1, Fig. 1b, S1c, S1e**). In addition to the common mutation
135 leading to the G34W alteration in H3.3, whole-genome sequencing of seven patient-derived
136 primary H3.3 MUT cell lines (**Table S1**) did not reveal any recurrent genetic alterations (**Fig.**
137 **S1e**). In contrast to other malignancies carrying *H3F3A* mutations, as for example pediatric
138 glioblastoma with co-occurring mutations in TP53 and ATRX (9,20), and to other bone tumors,
139 e.g. osteosarcoma, GCTB showed an extremely low overall mutation frequency for H3.3 MUT
140 and H3.3 WT cells (**Fig. 1d, Fig. S1e**). This places GCTB in a unique position to uncover
141 epigenomic alterations linked to H3.3-G34W.

142 To understand how H3.3-G34W exerts its function in tumor cells, we analyzed protein fractions
143 by Western blot and found H3.3-G34W incorporated into chromatin (**Fig. S2a**). To analyze *in*
144 *cis* effects of H3.3-G34W in patient derived cells, we used a specific and verified antibody to
145 map H3.3-G34W and identify sites of H3.3-G34W enrichment in two independent patient cell
146 lines. We identified high-confidence genomic regions showing H3.3-G34W enrichment (**Fig.**
147 **S2b, c, Table S2**) and profiled H3K36me3 along with several other histone modifications using
148 ChIP-mentation. Changes of H3K36me3 or H3K27me3 between H3.3 WT and H3.3 MUT
149 stromal cells at sites of H3.3-G34W enrichment were minute and did not recapitulate the *in cis*
150 results observed in HEK293 cells (**Fig. S2b, c**). Furthermore, using Western blot analysis of
151 whole cell lysates we did not reveal any changes in the total amount of K36 methylated histone
152 H3 indicating that the G34W substitution does not affect methylation of K36 on other histones
153 lacking the mutation. (**Fig. S2d**).



154

155

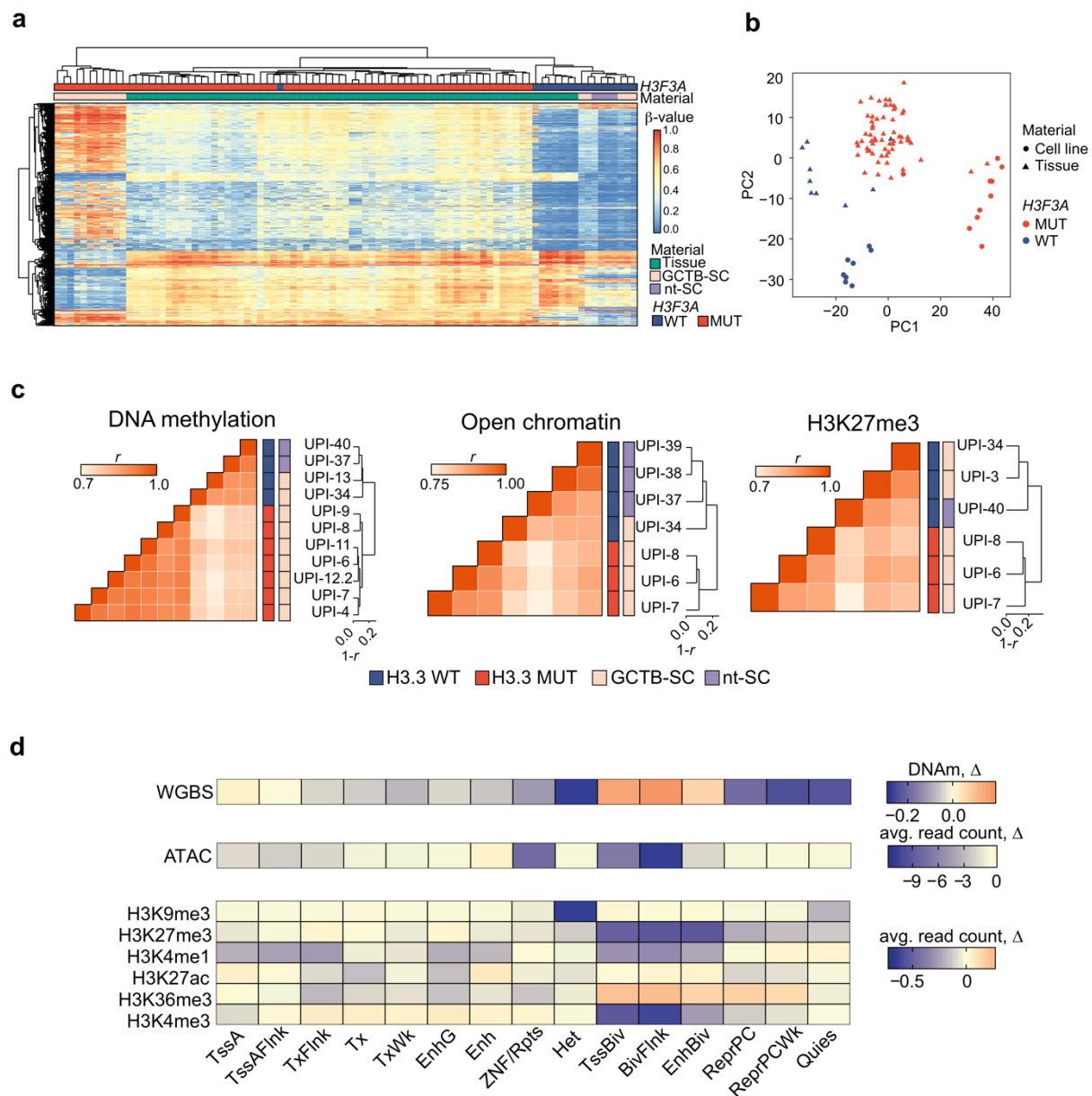
156 **Figure 1: Initial characterization of GCTB patient samples.**

157 **a.** Immunohistochemical staining of primary GCTB tumor resections with a H3.3-G34W-
 158 specific antibody (Active Motif) (red). UPI, unified patient identifier. **b.** Quantification of the
 159 mutation at position 103 in the *H3F3A* gene (c.103G>T) leading to the H3.3 G34W substitution
 160 in tumor resections and derived stromal cell lines using deep targeted resequencing. VAF,
 161 variant allele frequency. **c.** Overview of GCTB tissues and derived stromal cell lines analyzed
 162 within this paper. **d.** Circos plot of recurrent structural variants in osteosarcoma, H3.3-G34R-
 163 bearing pediatric glioblastoma (PedGBM) and GCTB cohorts based on whole genome
 164 sequencing data. Green lines represent translocations, blue lines deletions, red lines
 165 duplications, and black lines inversions. The variant recurrences are represented by bar plots.
 166 The outermost layer represents functional small variants (SNVs and small indels). The middle
 167 layer represents copy number variations. The innermost layer represents structural variations.
 168 All layers are normalized to the compared cohort size. Osteosarcoma cohort was sub-sampled
 169 at random to the size of the other two cohorts.

170 ***The epigenome of GCTB H3.3 MUT cells is globally altered***

171 We hypothesized that H3.3-G34W might exert its effects through other unidentified epigenetic
172 mechanisms and screened H3.3 MUT and H3.3 WT cells, as well as a large number of biopsies
173 of the four GCTB cohorts, using HumanMethylation450 arrays (89 samples in total). We
174 observed full-range DNA methylation changes of a large number of CpG sites (**Fig. 2a**).
175 Dimensionality reduction and clustering analysis of the methylation data showed the most
176 pronounced changes between primary patient-derived H3.3 MUT and H3.3 WT cell lines as
177 compared to GCTB biopsies. Most of the primary tumor samples had an intermediate
178 methylome identifying them as mixtures of normal and tumor cells (**Fig. 2a, b**), as also
179 confirmed by deep resequencing of the *H3F3A* gene locus of tissue-derived DNA showing
180 lower variant allele frequencies than theoretical 0.5 for a heterozygous mutation (**Fig. 1b**). This
181 observation prompted us to restrict all our subsequent analyses to H3.3 MUT and H3.3 WT
182 cells in order to obtain a clean view of H3.3-G34W-associated epigenetic alterations in pure
183 cell populations. To further a detailed analysis of the epigenome, we profiled DNA methylation
184 at single CpG resolution using whole genome bisulfite sequencing (WGBS), analyzed
185 chromatin accessibility with the assay for transposase-accessible chromatin using sequencing
186 (ATAC-seq) and analyzed global distribution of several histone marks (H3K4me1, H3K4me3,
187 H3K9me3, H3K27me3, H3K27ac, H3K36me3) to investigate their potential redistribution.
188 Global hierarchical clustering proved the H3.3-G34W substitution to be the major determinant
189 of variability in DNA methylation, chromatin accessibility and posttranslational histone
190 modifications (H3K27me3, H3K27ac, H3K4me1) with H3.3 MUT and H3.3 WT groups
191 forming separate clusters (**Fig. 2c, Fig. S2e**). In contrast, other modifications (H3K36me3,
192 H3K4me3) did not yield in H3.3 MUT- and H3.3 WT-separating clusters. Since nt-SCs
193 clustered with H3.3 WT cells, we combined all H3.3 wild type cells into a single control group
194 (H3.3 WT) for further analysis.

195 In order to identify loci with more pronounced epigenetic changes, we used a genome-wide
196 approach to stratify the observed differences from all epigenetic layers using the ENCODE
197 functional annotation of the MSC epigenome (21). We observed profound differences in several
198 groups of chromatin states including heterochromatic and Polycomb-repressed states as well as
199 bivalent domains (**Fig. 2d**). Heterochromatic regions were noticeably hypomethylated while
200 bivalent domains accumulated a whole range of alterations, including gain of DNA
201 methylation, loss of chromatin accessibility, as well as loss of several histone modifications
202 (H3K27me3, H3K4me3 etc). Collectively, we conclude that the epigenome of H3.3 MUT cells
203 shows significant and reproducible differences to that of H3.3 WT stromal cells.



204

205 **Figure 2: Genome-wide epigenetic distortion in H3.3 MUT cells.**

206 **a.** HumanMethylation450 profiles of GCTB tumor tissue samples, GCTB stromal cells and
 207 healthy nontumoral stromal cells (nt-SC). Heatmap displays 10,000 CpG sites with highest S.D.
 208 across all samples. Agglomerative, hierarchical clustering of rows (CpGs) and columns
 209 (samples) was performed with average-linkage based on Euclidean distance metric. **b.** Principal
 210 component analysis of HumanMethylation450 array-based DNA methylation profiles of GCTB
 211 tumor tissue, GCTB stromal cells and nontumoral stromal cells (nt-SC). **c.** Hierarchical
 212 clustering with correlation distance of DNA methylation (left), open chromatin (middle) and
 213 H3K27me3 profiles (right) in H3.3 WT (blue) and H3.3 MUT (red) cells. Heatmaps represent
 214 pairwise Pearson correlation coefficients (r) between respective modification profiles of two
 215 cell lines. Dendrograms were obtained with agglomerative hierarchical clustering with 1-r
 216 distance and average linkage. Heatmap color codes represent *H3F3A* mutational status (inner),
 217 and the cell type (outer), the common legend for all heatmaps is given at the bottom. nt-SC-
 218 nontumoral stromal cells. UPI, Unified patient identifier. **d.** Stratification of epigenetic
 219 differences using MSC-specific chromatin states as defined by ChromHMM. For DNA
 220 methylation (WGBS, “DNAm”) difference between average methylation levels of all CpGs

221 falling into corresponding states in H3.3 WT and H3.3 MUT cells is presented. For ATAC and
222 ChIP-seq of histone modifications, the difference between average normalized read counts over
223 all windows of a state is shown. Mnemonics for the ChromHMM states are defined by Roadmap
224 (TssA, active TSS; TssFlnk, active TSS flanking regions; Tx, transcribed regions; TxFlnk,
225 transcription flanking regions; TxWk, weakly transcribed regions; Enh, enhancers; EnhG,
226 genic enhancers, ZNF/Rpts; zinc finger genes and repeats; TssBiv, bivalent TSS; BivFlnk,
227 flanking bivalent regions; EnhBiv, bivalent enhancer; ReprPC, Polycomb repressed;
228 ReprPCWk, weak Polycomb repressed; Quies, quiescent).

229 *DNA hypomethylation in genomic megabase-scale domains indicates defects of*
230 *heterochromatin potentially contributing to stochastic genomic rearrangements*

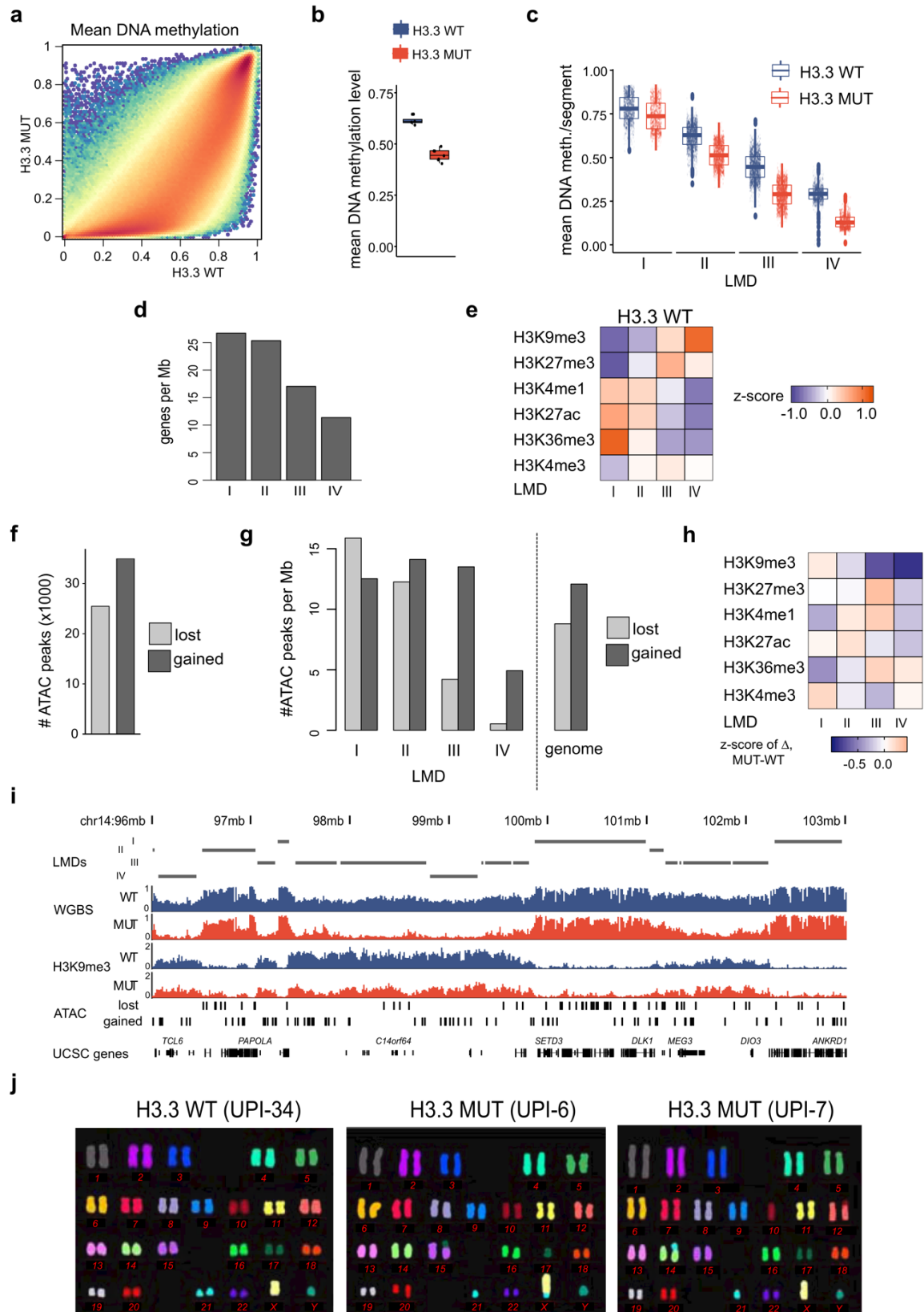
231 We first followed up changes associated with loss of DNA methylation. In contrast to the
232 homogeneity of DNA methylation profiles observed within each of the two groups (**Fig. S3a**),
233 DNA methylation was non-uniformly altered between H3.3 WT and H3.3 MUT cells across
234 the entire genome (**Fig. 3a**). The majority of significantly changed loci showed profound
235 reduction of DNA methylation in H3.3 MUT cells, while a minority was hyper-methylated
236 (**Fig.3a, Fig. S3b**). Overall, H3.3 MUT cells exhibited a 20% genome-wide reduction of DNA
237 methylation (**Fig.3b, Fig. S3c**). To systematically characterize the abundant DNA methylation
238 changes, we segmented the genome into large methylation domains (LMD I to IV) with sizes
239 >20kb (**Fig. S3d, Table S2**) using a combination of breakpoint analysis and clustering (see
240 Methods for details). The majority of regions (predominantly LMDs III and IV) matched the
241 criteria for partially methylated domains i.e. megabase-scale domains of predominantly
242 repressive chromatin with low gene density (22). Each LMD had a distinct level of DNA
243 methylation, a discrete pattern of histone marks and gene density, suggesting different
244 functional roles (**Fig. 3c-e**). For instance, LMD III showed enrichment of H3K27me3, while
245 LMD IV was associated with H3K9me3, suggesting an association with facultative and
246 constitutive heterochromatin (**Fig. 3e**). Furthermore, LMD III domains were often detected as
247 flanking to LMD IV. LMDs I to IV showed decreasing average methylation levels starting from
248 over 0.75 down to less than 0.25 (**Fig. 3c**). Lowly methylated LMDs III and IV which could be
249 characterized as heterochromatic (**Fig. 3e**) showed the most pronounced demethylation
250 comparing H3.3 MUT cells to H3.3 WT cells. We concluded that global epigenetic alterations
251 in GCTB are non-uniformly distributed along the genome, and that the most pronounced global
252 changes take place in large-scale domains associated with facultative and constitutive
253 heterochromatin, roughly corresponding to LMDs III and IV.

254 We next sought to associate hypomethylation with other epigenetic alterations. Consistent with
255 the global loss of DNA methylation, we observed increased chromatin accessibility in the H3.3

256 MUT cells with approximately 1.5-times more gained than lost differential ATAC-seq peaks
257 (**Fig. 3f**). ATAC peaks gained in H3.3 MUT cells were overrepresented in the LMDs III and
258 IV (**Fig. 3g**) indicating a more open state in respective genomic segments. These LMDs also
259 showed strong reduction in the heterochromatic histone mark H3K9me3 (**Fig. 3h**). Examples
260 of changes in different LMDs can be seen in **Fig. 3i**. The ATAC peaks gained in H3.3 MUT
261 showed a significant overlap with repetitive elements such as LINE, SINE and LTR elements
262 which are normally silenced by DNA methylation (**Fig. S3e, Table S3**). Looking at other
263 known repetitive regions, such as centromeres and telomers, we found them to be affected by
264 the genome-wide hypomethylation (**Fig. S3f**). m-FISH analysis found H3.3 WT cells to have a
265 normal karyotype, while H3.3 MUT cells in contrast displayed different, non-recurrent
266 centromeric fusions which could be a potential consequence of the heterochromatin defects
267 described above (**Fig. 3j**). We conclude that H3.3-G34W associates with heterochromatic
268 defects that potentially contribute to a genomic instability previously described as characteristic
269 for GCTB (23).

270 **Figure 3: Heterochromatin defects in H3.3 MUT stromal cells** (on the next page).

271 **a.** Binned scatterplot of DNA methylation profiles from WGBS of H3.3 WT cells versus H3.3
272 MUT cells. Hexagon color represents binned density gradient of 1 (blue), 1000 (yellow), and
273 10^6 (red) points. **b.** Genome-wide mean level of DNA methylation in H3.3 WT (blue) and H3.3
274 MUT (red) cells. Points represent individual patients, box-plot bar represents the mean, box the
275 inter-quartile range, whiskers extend from the smallest to the largest value within 1.5 IQR from
276 the lower and upper edges of the box, respectively. The difference is statistically significant
277 ($p=1.5 \cdot 10^{-5}$, two-sample t-test). **c.** Methylation level distribution at large methylation domains
278 (LMDs) in H3.3 WT (blue) and H3.3 MUT (red) cells. Each point represents a single LMD
279 segment. **d.** Mean gene density of LMDs. RefSeq genes were used and for each segment the
280 number of overlapping genes was divided by its length in Mb. **e.** Normalized levels of histone
281 modifications at LMDs in H3.3 WT cells. Heatmap represents z-scores of average signals
282 across all segments falling into the LMDs calculated using overall mean level and standard
283 deviation across all LMD segments. **f.** Absolute numbers of ATAC-seq peaks with differential
284 chromatin accessibility between H3.3 WT cells and H3.3 MUT cells. “lost” and “gained” refer
285 to peaks showing a decrease (disappearance) or increase (emergence) in H3.3 MUT cells,
286 respectively. **g.** Stratification of differential chromatin accessibility between H3.3 WT and H3.3
287 MUT cells by overlap with LMDs. Labels “lost” and “gained” differential ATAC peaks as in
288 (**f**); “genome”, distribution of ATAC peaks across the complete genome. **h.** Changes of histone
289 modifications between H3.3 WT and H3.3 MUT stromal cells at LMDs. Heatmap represents
290 difference of average read counts, Z-score-normalized with respect to zero. **i.** Genomic browser
291 view of a representative 7Mbp region on chromosome 14 (positions 96,000,000–103,000,000)
292 for whole-genome bisulfite sequencing, H3K9me3 ChIP-seq and ATAC-seq. Tracks are shown
293 for H3.3 WT (blue) and H3.3 MUT (red) cells. Bottom panel indicates gene locations based on
294 the UCSC default gene track. “lost” and “gained” differential ATAC-seq peaks as in (**f**). **j.**
295 Visualization of karyotypes using m-FISH in H3.3 WT and H3.3 MUT cells.



296

297 ***Isogenic cells expressing H3.3-G34W recapitulate DNA methylation changes found in H3.3***
298 ***MUT stromal cells***

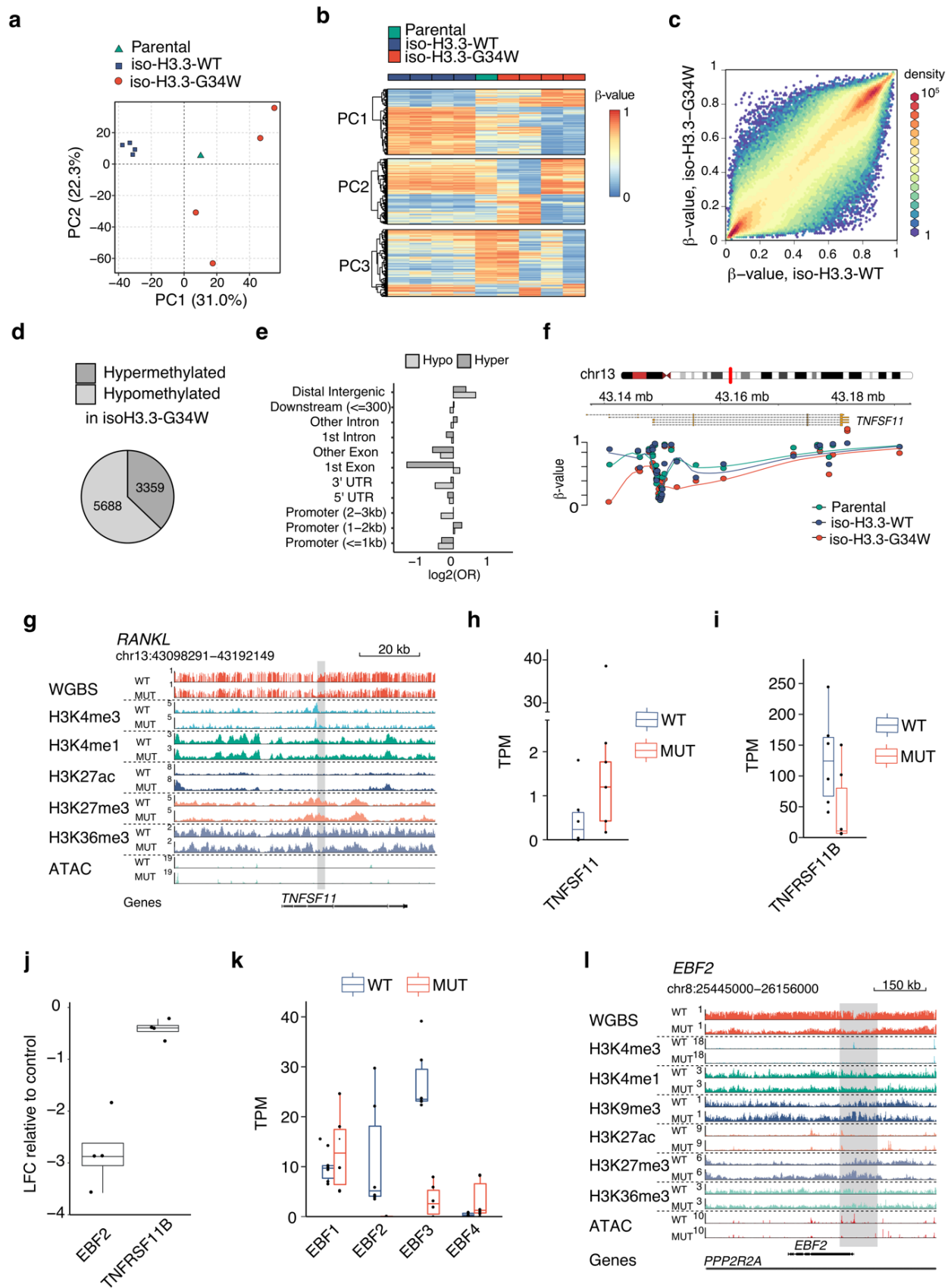
299 As H3.3 MUT stromal cells showed the strongest epigenetic differences to H3.3 WT stromal
300 cells on the DNA methylation level, we expect DNA methylation to play a major role in stromal
301 cell transformation leading to GCTB. To verify the observed epigenetic changes to be
302 dependent on H3.3-G34W expression, we aimed to recapitulate the findings in an unrelated cell
303 line system. To this end, we introduced the H3.3-G34W encoding mutation in HeLa cells by
304 targeting the endogenous *H3F3A* locus as earlier described (**Fig. S4a**) (24). Individual iso-H3.3-
305 WT and iso-H3.3-G34W clones were isolated of which four iso-H3.3-WT and iso-H3.3-G34W
306 monoclonal lines and one parental HeLa cell line were subject to HumanMethylationEPIC
307 DNA methylation analysis (**Fig. 4a**). The iso-H3.3-G34W samples showed similar alterations
308 as found when comparing H3.3 MUT and H3.3 WT stromal cells of GCTB (**Fig. 4a-e**).
309 Principal component analysis of the most variable methylation probes clustered iso-H3.3-WT
310 clones together while iso-H3.3-G34W clones dispersed off, indicating changes in DNA
311 methylation in iso-H3.3-G34W but not WT isogenic cell lines (**Fig. 4a**). Similar to primary
312 GCTB cells, the largest principal component (PC1, 31% of variance explained) captured
313 widespread hypomethylation in iso-H3.3-G34W accompanied by focal hypermethylation
314 events (**Fig. 4a**). Clustering analysis further enforced the distinct difference between iso-H3.3-
315 WT and G34W clones (**Fig. 4b**). A scatter plot of all probes indicated that iso-H3.3-G34W cells
316 showed predominantly hypomethylation like H3.3 MUT cells from GCTB (**Fig. 4c**). We found
317 9,047 differentially methylated probes ($\Delta \beta$ -value>0.2 and FDR<0.05) of which 5,688 (63%)
318 were hypomethylated in iso-H3.3-G34W cells (**Fig. 4d**). CpG sites that lost methylation were
319 more frequently localized to intergenic regions and promoter-proximal exons, while
320 hypermethylated ones were in addition moderately enriched at promoters and depleted at exons
321 (**Fig. 4e**). To a large extent, this recapitulated the situation in primary GCTB stromal cells and
322 confirmed that the changes in DNA methylation are associated with H3.3-G34W. As a specific
323 example, the *RANKL* locus showed hypomethylation at exon 3, a potential alternative promoter
324 element, (**Fig. 4f**) which is in line with the methylation differences observed between H3.3 WT
325 and H3.3 MUT stromal cells (**Fig. 4g**). RANKL signaling has been extensively studied in
326 GCTB (25). The expression of RANKL, a master regulator of osteoclast differentiation (26),
327 has been shown to be upregulated in GCTB stromal cells causing an osteolytic phenotype (25).
328 We verified the increased expression of *RANKL* in H3.3 MUT cells (**Fig. 4h, Fig. S4b**) and
329 additionally observed decreased expression and secretion of its decoy receptor Osteoprotegerin
330 (*OPG, TNFRSF11B*) (**Fig. 4i, Fig. S4 b, c**). Similar expression patterns were already described

331 by us earlier (24). *OPG* had decreased levels of the active histone marks H3K4me3 and
332 H3K27ac potentially indicating a missing activation by a transcription factor (**Fig. S4d**). One
333 known transcription factor of *OPG* is Early B-cell Factor 2 (*EBF2*) (27,28) which is a bivalent
334 gene (**Fig. S4l**) and belongs to key regulators of osteogenic differentiation in mice (29). *OPG*
335 became downregulated after siRNA mediated *EBF2* knockdown, confirming a role of *EBF2* in
336 *OPG* expression in stromal cells (**Fig. 4j, Fig. S4e**). The *EBF* family is a conserved group of
337 four transcription factors. Our RNA-seq analysis found *EBF2* and *EBF3* to be differentially
338 expressed between H3.3 WT and H3.3 MUT stromal cells (**Fig. 4k, Fig. S4f**). *EBF3*, previously
339 reported as a tumor suppressor in glioblastoma (30,31), showed reduced expression in H3.3
340 MUT cells whereas *EBF2* expression was almost completely lost. We found the *EBF2* locus to
341 be hypomethylated with a focal hypermethylation around the promoter region (**Fig. 4l**).
342 Increased H3K9me3 and H3K27me3 levels and decreased levels of H3K27ac supported a
343 repressed state of *EBF2*. Lost ATAC signals in H3.3 MUT cells indicated differentially closed
344 chromatin. These findings link the H3.3 G34W-associated epigenetic dysregulation of *EBF2*
345 expression to the osteolytic phenotype of GCTB.

346 **Figure 4: DNA methylation differences can be recapitulated in isogenic HeLa cells and link**
347 **epigenetic distortion to the osteolytic phenotype of GCTB** (on the next page).

348 **a.** Principal component analysis of HumanMethylationEPIC profiles of iso-H3.3-WT and iso-
349 H3.3-G34W HeLa cells. **b.** Methylation levels of top 2000 probes associated with each of the
350 first three principal components in iso-H3.3-WT and iso-H3.3-G34W cells. **c.** Scatter plot of
351 individual CpG probe methylation in iso-H3.3-WT (x-axis) vs. iso-H3.3-G34W (y-axis). Color
352 dots indicate incremental delta means of methylation, with orange representing >0.5 delta-
353 mean. **d.** Fractions of HumanMethylationEPIC probes, differentially methylated between iso-
354 H3.3-G34W and iso-H3.3-WT HeLa cells, by direction of DNA methylation change. **e.**
355 Enrichment (positive values) or depletion (negative values) of major gene model features
356 among differentially methylated probes. OR, odds ratio. **f.** Genome browser snapshot
357 displaying DNA methylation levels of CpGs in the vicinity of the *TNFSF11* locus in iso-H3.3-
358 WT and iso-H3.3-G4W cells. Points represent methylation values individual CpGs and the lines
359 depict LOESS curves with degree 1 and span 0.5. **g.** Genomic browser view of gene *TNFSF11*
360 encoding for RANKL. Each lane (dash-separated) represents normalized averaged signals of
361 DNA methylation, the level of H3K4me3, H3K4me, H3K9me3, H3K27ac, H3K27me3,
362 H3K36me3 and ATAC in several replicates of H3.3 WT (lane-wise top) or H3.3 MUT (lane-
363 wise bottom) cells. **h.** Expression levels of RANKL (*TNFSF11*) in H3.3 WT (blue) H3.3 MUT
364 (red) cells as analyzed by RNA-seq. TPM, transcripts per million. **i.** Expression levels of *OPG*
365 (*TNFRSF11B*) in H3.3 WT (blue) H3.3 MUT (red) cells as analyzed by RNA-seq. TPM,
366 transcripts per million. **j.** Expression analysis of *EBF2* and *OPG* (*TNFRSF11B*) by qPCR 48
367 hours after siRNA-mediated knockdown of *EBF2* in H3.3 WT cells (UPI-13). Log fold change
368 (LFC) relative to expression in cells transfected with a control siRNA. Each dot represents one
369 replicate. Boxplot bar represents mean, box gives the IQR, whiskers span additional 1.5*IQR
370 below and above the bars. Results showed significance in a one-sample t-test with p-values <
371 0.05. **k.** Expression levels of all members of the *EBF* family in H3.3 WT (blue) H3.3 MUT
372 (red) cells as analyzed by RNA-seq. TPM, transcripts per million. **l.** Genomic browser view of
373 the *EBF2* locus. Each lane (dash-separated) represents normalized averaged signals of DNA

374 methylation, the level of H3K4me3, H3K4me, H3K9me3, H3K27ac, H3K27me3, H3K36me3
 375 and ATAC in several replicates of H3.3 WT (lane-wise top) or H3.3 MUT (lane-wise bottom)
 376 cells.



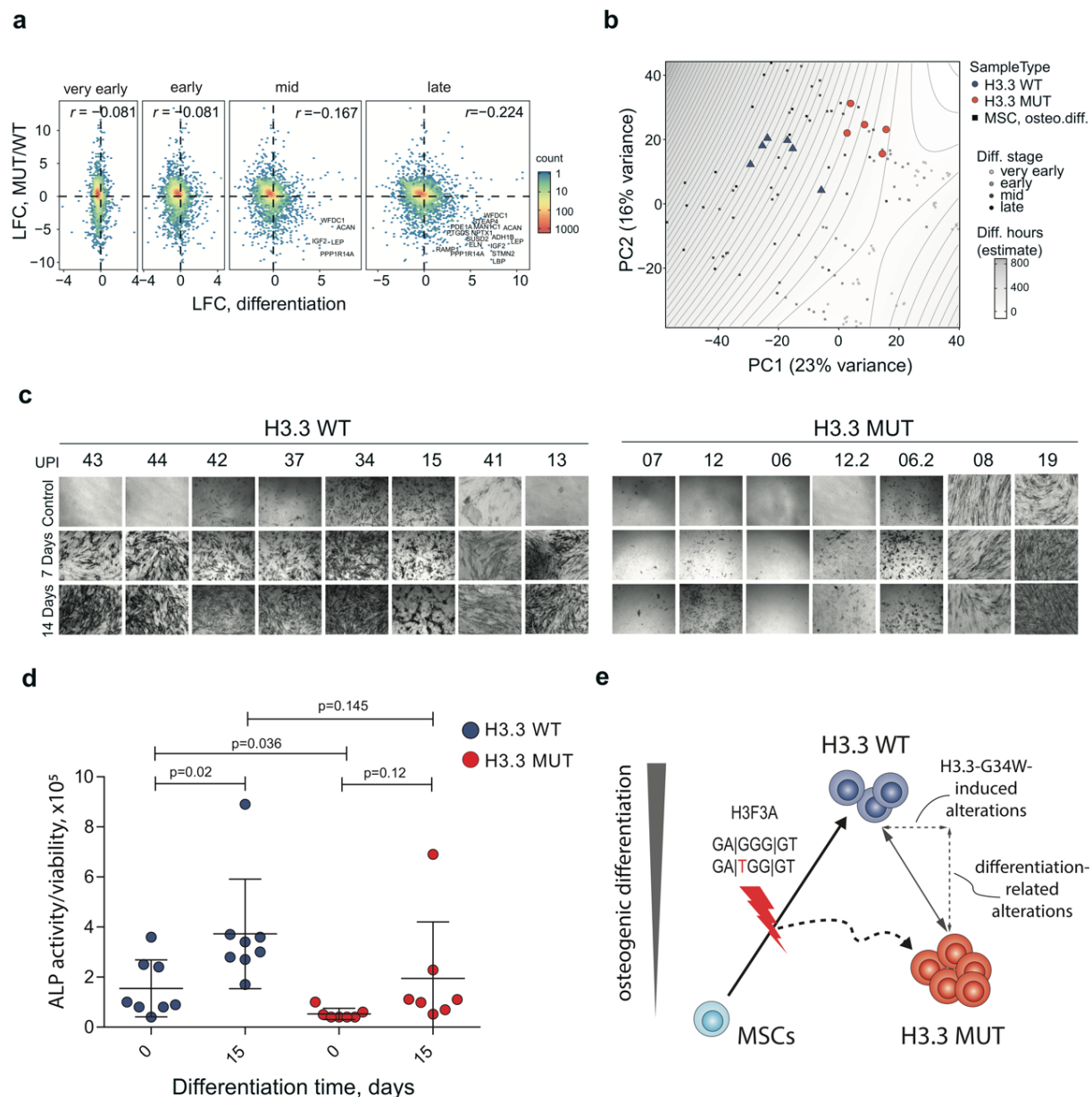
377

378 Taken together, we show that the stable introduction of H3.3-G34W into a GCTB unrelated
379 cell line recapitulates the DNA hypomethylation trend seen in GCTB stromal cells.
380 Furthermore, we could detect epigenetic alterations that directly and indirectly affected the
381 expression of key regulators of bone metabolism shedding novel light on the emergence of the
382 osteolytic phenotype in GCTB.

383 ***H3.3 MUT cells exhibits localized changes at bivalent promoters and have a distinct***
384 ***osteogenic differentiation state***

385 Besides dominant hypomethylation in heterochromatic regions, we additionally observed
386 profound epigenetic changes at bivalent domains (**Fig. 2d**), which are well known sites of H3.3
387 deposition (5,32). Changes included gain of DNA methylation which coincide with decreased
388 accessibility, decrease of H3K27me3 and increased levels of H3K36me3 (**Fig. 2d**). Most
389 promoter-associated ATAC peaks that were lost in H3.3 MUT cells overlapped with bivalent
390 promoters (**Fig. S5a**). As a consequence of epigenetic disturbance, bivalent genes comprised a
391 significant portion of differentially expressed genes both up- and downregulated in H3.3 MUT
392 cells (**Fig. S5b**). It was previously reported that comparable epigenetic and transcriptional
393 perturbations at bivalent genes occur in H3.3-deficient mouse ESCs (7). Genes with decreased
394 expression in H3.3 MUT were significant enriched in Polycomb-target genes, including several
395 developmental transcription factors (**Fig. S5c, Table S4**). This was also supported by genomic
396 overlap enrichment analysis of lost ATAC peaks showing strong enrichment at regions where
397 binding of PRC2 components EZH2 and SUZ12 was found in numerous cell lines (**Fig. S5d,**
398 **Table S3**). GO analysis of closed regions revealed many categories related to differentiation
399 (**Fig. S5e**). Impaired differentiation was already suggested for tumor entities harboring
400 mutations in histones including GCTB (15,33). To investigate whether H3.3 WT and H3.3
401 MUT cells differ in their osteogenic differentiation state, we performed a comparison across all
402 genes covered by RNA-seq with those obtained from an *in vitro* differentiation of MSCs. The
403 expression changes during osteogenic differentiation largely correlated inversely with the
404 expression changes observed between H3.3 WT and H3.3 MUT cells, with anticorrelation
405 increasing from very early to late differentiation stages (**Fig. 5a**). Accordingly, an overlap of
406 the differentially expressed genes from both experiments showed that the majority of matches
407 were upregulated during differentiation and downregulated in H3.3 MUT cells (**Fig. S5f**).
408 Examples include insulin growth factor 2 (*IGF2*) and leptin (*LEP*) previously implicated in
409 osteogenic differentiation (34,35) (**Fig. 5a**). A global principal component analysis confirmed
410 a less differentiated state of H3.3 MUT cells resembling MSC in a very early stage of

411 differentiation and H3.3 WT cells showing an expression profile more similar to MSC in early
412 to middle stage of osteogenic differentiation (**Fig. 5b**). To confirm differences in the osteogenic
413 differentiation state, we stained H3.3 WT and H3.3 MUT stromal cells for the activity of the
414 osteogenic marker alkaline phosphatase (ALP). Most of the H3.3 WT cells exhibited ALP
415 activity whereas only some H3.3 MUT cells showed ALP activity (**Fig. 5c**). Quantification of
416 ALP activity relative to viability as a surrogate for cell count confirmed reduced ALP activity
417 for H3.3 MUT stromal cells (**Fig. 5d**). Impaired differentiation of GCTB stromal cells was
418 suggested earlier based on the analysis of histological markers (36) and transcriptomic profiling
419 (37). We performed an osteogenic differentiation of H3.3 MUT and H3.3 WT cells from several
420 patients *in vitro* to analyze the potential of H3.3 MUT and H3.3 WT cells to differentiate. While
421 both H3.3 WT and H3.3 MUT stromal cells showed an increase of ALP activity indicating the
422 potential for osteogenic differentiation (**Fig. 5c and d**), we noticed that H3.3 MUT stromal cells
423 lagged behind and did not achieve the level of ALP activity reached by H3.3 WT stromal cells
424 (**Fig. 5c and d**). We conclude that H3.3-G34W associates with changes in bivalent regions and
425 that their deregulation potentially has an effect on osteogenic differentiation. H3.3 WT and
426 H3.3 MUT stromal cells therefore differ in their state of osteogenic differentiation in GCTB
427 patients and this difference contributes to the herein found epigenetic alterations (**Fig. 5e**).



428

429

Figure 5: Impairment of osteogenic differentiation in H3.3 MUT stromal cells

430

a. Correlation between differentiation- and H3.3 G34W-associated gene expression changes.

431

For the differentiation log-fold changes (LFC) were calculated between average of profiles

432

from each differentiation timepoint group (very early: 0.5 to 2 hours; early: 4 to 554 16 hours,

433

mid: 24 to 168 hours; and at: 336 to 504 hours) and the profiles of MSCs sampled prior to the

434

start of the experiment. Each hexagon reflects density of the underlying data points. Genes with

435

most pronounced changes in both data sets are marked. **b.** Principal component analysis of

436

RNA-Seq data from osteogenic differentiation of MSC and expression profiles of GCTB

437

samples added via a projection onto the subspace spanned by principal components 1 and 2.

438

The sampled differentiation timepoints were grouped as: very early (0.5 to 2 hours), early (4 to

439

16 hours), mid (24 to 168 hours) and late (336 to 504 hours). **c.** Alkaline phosphatase staining

440

of nt-SC, H3.3 WT and H3.3 MUT cells during osteogenic differentiation. **d.** Quantification of

441

alkaline phosphatase activity relative to the viability as a surrogate for cell count over the time

442

course of 14 days of *in vitro* osteogenic differentiation. **e.** Model for GCTB tumorigenesis: The

443

H3F3A mutation encoding H3.3-G34W occurs in osteoblastic precursor cells and leads to

444

alterations in osteogenic differentiation. Epigenetic differences described between H3.3 WT

445

and H3.3 MUT stromal cells are the result of measuring cells at different differentiation stages

446

and direct effects of the mutated histone.

447 Discussion

448 Mutant epigenetic regulator proteins, including histones, have been described in multiple cancer
449 types (2). Despite the fact that oncogenic mutations in histones influence the epigenetic
450 landscape, deep insight into the mechanistic ramifications to cancer initiation is still incomplete.
451 Several studies investigated the mechanism of how the G34W substitution in histone variant
452 H3.3 affects the epigenetic machinery responsible for its post-translational modifications. To
453 this end, reduced H3K36 methylation and increased H3K27me3 *in cis* was described in model
454 systems and verified by us in this work (16,38). However, whether these effects can be found
455 in patients and whether epigenetic changes are involved in tumorigenesis of GCTB was not
456 known so far. In this paper, we therefore analyzed GCTB tissue as well as patient-derived
457 primary stromal cells from four different centers to investigate H3.3-G34W-associated
458 epigenetic changes and their contribution to the GCTB neoplasia. The high incidence of H3.3-
459 G34W in GCTB in a largely unaltered genomic context leaves GCTB as a suitable system to
460 study histone mutation-driven tumorigenesis.

461 We found that, despite that H3.3-G34W is incorporated into chromatin, its presence did not
462 lead to changes in the global amount of H3K36me2/me3 as shown for other substitutions
463 affecting H3.3 lysine 27 and lysine 36 (11,13,39). This ruled out a possible *in trans* effects of
464 H3.3-G34W on histone posttranslational modifications analyzed in this study. Using ChIP
465 sequencing, we identified high confidence H3.3-G34W enrichment sites in patient cell lines
466 and analyzed common histone marks at these regions. No pronounced differences were found
467 that could recapitulate the *in cis* effects on H3.3 lysine methylation observed in HEK293 cells.
468 The absence of clear-cut *in cis* effects upon chromatin modifications could be explained by an
469 inability to distinguish wild-type and mutant H3.3 in GCTB samples by western blot analysis.
470 In ChIP analysis, nucleosomes of mutant cells could include both, wild-type as well as mutated
471 H3.3 obliterating possible *in cis* effects. Moreover, antibodies against histone modifications
472 cannot distinguish between H3 variants and the overall small fraction of the mutant H3.3 (25%
473 of overall H3.3) in the total H3 pool makes analysis of H3.3 modifications challenging.

474 We show that H3.3-G34W, the sole recurrent alteration of GCTB, associates with large-scale
475 differences in multiple epigenetic marks. H3.3 MUT stromal cells show heterochromatic
476 defects, specifically reduced genome-wide levels of DNA methylation and gained accessibility
477 at heterochromatic regions which is in line with previous reports on pediatric glioblastomas
478 with the G34R substitution in H3.3 (39). These changes could potentially contribute to the

479 genomic instability described for GCTB (23). We furthermore observed localized changes at
480 bivalent regions, some of the same sites previously described as targets of H3.3 deposition (5,7).
481 DNA methylation changes could be recapitulated using an isogenic system verifying effects on
482 the methylome to be H3.3-G34W-specific. Moreover, we show H3.3-G34W-associated
483 methylome changes to directly or indirectly affect the expression of the main players of bone
484 metabolism, RANKL and OPG. These results connect the H3.3-G34W-associated epigenetic
485 dysregulation to the osteolytic phenotype, a hallmark of GCTB.

486 While the epigenomic differences between H3.3 WT and H3.3 MUT stromal cells can be
487 partially ascribed to direct effects of H3.3-G34W, other alterations observed can be explained
488 by the fact that H3.3 MUT and H3.3 WT stromal cells represent distinct differentiation stages.
489 Assuming epigenetic reprogramming during differentiation as demonstrated in recent studies
490 (40,41), this could also contribute to the epigenetic differences seen between H3.3 WT and
491 H3.3 MUT cells. While expression profiles from H3.3 MUT cells resemble precursor states of
492 osteoblasts, the H3.3 WT cells resemble more mature osteoblasts, suggesting impaired or
493 delayed differentiation of H3.3 MUT cells as already suggested by (42). A weak differentiation
494 signature is in line with previous reports suggesting that neoplastic GCTB stromal cells are
495 expressing markers of an early osteoblastic differentiation (43). Genes such as insulin growth
496 factor 2 (*IGF2*) (34), leptin (*LEP*) and stathmin-like 2 (*STMN2*), found to be dysregulated in
497 H3.3 MUT stromal cells, were described to play a role in mesenchymal stem cell differentiation
498 and osteoclastogenesis in mice, in promoting bone differentiation (35) and as a marker for
499 osteogenesis (44), respectively. The impairment of differentiation of H3.3 MUT stromal cells
500 could be mediated by H3.3-G34W- associated downregulation or prevention of induction of
501 genes upregulated during osteogenic differentiation through a yet unknown mechanism.
502 Epigenetic repression of bivalent genes and dysregulation of PRC2 targets described in this
503 study could potentially contribute to the observed phenotype. H3.3 MUT stromal cells are still
504 able to undergo osteogenic differentiation, as it was also found in studies not separating H3.3-
505 mutant neoplastic from H3.3 -wild-type stromal cells of GCTB tissue (45,46).

506 Taken together, our results suggest that the molecular mechanism behind H3.3-G34W-induced
507 epigenomic alterations and neoplastic transformation in GCTB is different to the action of other
508 mutant histone variants such as H3.3-K27M and H3.3-K36M, where the mutated amino acid is
509 a direct target for histone modification (11,13,15). This suggests that the H3.3-G34W
510 substitution alters the N-terminal interactome as previously described (24), recruiting novel
511 binders or disrupting interactions with binders. The described DNA methylation changes can
512 be indicative of H3.3-G34W affecting either binding or activity of enzymes writing (DNMTs)

513 or erasing (TETs) this modification. On the structural side it is known that DNMT3a/b
514 recognize H3K36me3 with their PWWP-domains (47,48). The G34W substitution might
515 interfere with binding of DNMT3a/b PWWP preventing proper establishment of DNA
516 methylation patterns. The thereby induced remodeling of the epigenome associated with H3.3-
517 G34W, in particular changes at heterochromatic and bivalent regions, could contribute to an
518 impaired osteogenic differentiation and the phenotypes of GCTB: stochastic chromosomal
519 rearrangements and increased RANKL signaling. Precise sequences of molecular events
520 leading to these phenotypes will be a matter of further studies.

521 **Methods**

522 *Patient samples*

523 Biopsies and derived primary cell lines were obtained from patients in the Orthopedic
524 University Hospital, Heidelberg (UHOK), the Korean Cancer Center (KCC), the University
525 Clinic of Leipzig (UKL) and the University Medical Center Hamburg-Eppendorf (UKHE). The
526 complete list of samples used in the study is available in **Table S1**. The use of patient samples
527 and the experiments performed in this study was approved by and in accordance with guidelines
528 and regulations by the Ethics Committees of the University of Heidelberg, University Clinic of
529 Leipzig, University Medical Center Hamburg-Eppendorf, and the National Cancer Center of
530 Korea (IRB NCC2015-0070). All H3.3 WT and H3.3 MUT cells analyzed in the manuscript
531 were obtained from different patients or indicated as relapse samples by addition of .2 to the
532 UPI.

533 *Isolation of GCTB stromal cells from patient tissue*

534 Tumor tissue from surgical resections was mechanically cut into small pieces and digested with
535 1.5 mg/ml collagenase B (Roche Diagnostics, Mannheim, Germany) at 37°C for 3 h in
536 Dulbecco's Modified Eagle Medium (DMEM) (Lonza GmbH, Wuppertal, Germany) high
537 glucose supplemented with 10% fetal calf serum (FCS) (Biochrom, Berlin, Germany), and
538 100 U/ml penicillin/streptomycin (Lonza GmbH). Cells were collected by centrifugation,
539 washed twice in PBS and cultured in DMEM as described above. Twenty-four hours after
540 plating, cells were carefully treated with Trypsin/EDTA (Lonza GmbH) leaving the giant cells
541 attached in the culture flask. Detached cells were cultured for further 2 passages to eliminate
542 histiocytes and remaining giant cells.

543 ***Isolation of nt-SC cell lines from the bone marrow***

544 Nt-SCs were isolated from fresh bone marrow samples derived from the iliac crest under the
545 approval of the Ethics Committee of the University of Heidelberg. Bone marrow cells were
546 purified on a Ficoll-Paque™ Plus density gradient (GE Healthcare, München, Germany),
547 washed in PBS and treated with erythrocyte lysis buffer (0.154 M NH₄Cl, 10 mM KHCO₃,
548 0.1 mM EDTA) to remove erythrocytes. The nt-SC-enriched fraction was seeded and cultured
549 in DMEM high-glucose supplemented with 12.5% FCS (Biochrom), 2 mM L-glutamine,
550 100 units/ml penicillin, 100 µg/ml streptomycin (Lonza), 4 ng/ml basic fibroblast growth factor
551 (Merck Chemicals GmbH, Darmstadt, Germany, 50 µM 2-mercaptoethanol and 1% non-
552 essential amino acids (Invitrogen, Karlsruhe, Germany). After 48 h, cultures were washed with
553 PBS to remove non-adherent material. During expansion, medium was replaced twice a week.

554 ***Cell line maintenance***

555 The GCTB stromal and nt-SC cell lines were cultivated in DMEM/F12 no phenol red (Gibco,
556 Invitrogen, Life Technologies, Paislay, UK) supplemented with 10% FCS (Biochrom, Berlin,
557 Germany) and 2 ng/µl basic FGF (Biolegend, San Diego, USA) at 37°C and 5% CO₂. To avoid
558 contact inhibition, cells were split after reaching maximum 90% of confluency.

559 ***Isogenic cell lines***

560 The establishment of H3.3-GFP isogenic cell lines in HeLa cells utilized the zinc-finger (ZF)
561 targeting methodology as previously described (24). In brief, after transfection with ZF and
562 targeting constructs with electroporation, cells were cultured in media under neomycin
563 selection using G418 for four weeks. Surviving cells were FACS-sorted and individual clones
564 were allowed to propagate. HEK293T were lentivirally-transduced with human H3.3-HA-
565 3XFLAG constructs in pLVX expression vectors and selected with puromycin. A pLVX-
566 AcGFP construct was used as control. Cells were maintained in DMEM (Invitrogen™ Life
567 Technologies, Carlsbad, USA) supplemented with 10% FCS (Biochrom, Berlin, Germany) and
568 penicillin-streptomycin.

569 ***Immunohistochemistry***

570 For immunohistochemical detection of the H3.3-G34W mutation formalin fixed paraffin
571 embedded GCT tissue sections were deparaffinized in Roti-Histol (Carl Roth GmbH,
572 Karlsruhe, Germany) and rehydrated in isopropanol. Antigen retrieval was performed using

573 Dako target retrieval solution pH 6 (Dako, Hamburg, Germany) for 5 min at 121°C in a
574 pressure cooker. Sections were blocked for 30 min at room temperature using PBS
575 supplemented with 5 % bovine serum albumin (BSA). The primary rabbit anti-H3G34W
576 antibody (Active Motiv, Carlsbad, USA) was diluted 1:1000 in PBS/1% BSA and incubated
577 over night at 4°C. The signal was amplified using the BrightVision +Poly-AP kit (VWR,
578 Darmstadt, Germany) according to the manufacturer's instructions. Samples were
579 counterstained with hematoxylin (Carl Roth GmbH) and mounted using Neo-Mount (Merck,
580 Darmstadt, Germany). All used antibodies are listed in **Table S5**.

581 *Lineage verification by flow-cytometric analysis*

582 Stromal cells were harvested by trypsination and transferred to 1,5 ml Eppendorf in PBS. Cells
583 were washed with PBS 2% FCS and centrifuged for 5 min at 250 g at 4°C. The antibody staining
584 cocktail was prepared in PBS 2% FCS and cells were stained 20 min at 4°C (all antibodies are
585 specified in the **Table S5**). Afterwards cells were washed with PBS 2% FCS to remove
586 unbound antibodies and centrifuged at 250 g for 5 min. To exclude dead cells during flow
587 cytometric analysis, propidium iodide was added prior to flow analysis. Data were acquired on
588 a BD FACSAria Fusion (Beckton, Dickinson & Company (BD)) and data were analyzed using
589 FlowJo (BD). Prism (GraphPad Software) was used to generate bar graphs.
590 Hematopoietic/erythroid cells were defined as CD45+ CD235+, CD45-CD235- CD105+ and
591 CD45-CD235- CD90+ cells were defined as mesenchymal cells. Freshly frozen iliac crest bone
592 marrow aspirates from healthy donors were used as healthy controls.

593 *OPG expression quantification with enzyme-linked immunosorbent assay (ELISA)*

594 5000 cells were seeded into a 96-well plate and cultivated in 100µl media for two days. ELISA
595 was performed with the Abcam human Osteoprotegerin ELISA Kit following the
596 manufacturer's instructions.

597 *Cell titer blue assay*

598 Cells were cultivated in a 96-well plate. 100µl fresh media (DMEM/F12 no phenol red (Gibco,
599 Invitrogen, Life Technologies, Paislay, UK) supplemented with 10% FCS (Biochrom, Berlin,
600 Germany)) were added and supplemented with 20µl CTB reagent from Promega. After 2.5h of
601 cultivation at 37°C Fluorescence was recorded at 560Ex/590Em.

602 *siRNA knockdown*

603 120.00 cells were seeded into a 6-well plate. After 24 hours, medium was changes and
604 transfection was performed using 1µl Dharmafect1 (Horizon Discovery, Waterbeach, UK)
605 reagent in combination with 2 µl of a 25 µM siRNA SMARTpool of Dharmacon (Horizon
606 Discovery, Waterbeach, UK) targeting EBF2. Cells were harvested 48 hours after transfection.

607 ***Detection of H3F3A-G34W mutation by mutation-specific PCR***

608 Genomic DNA was isolated using the Quick DNA Miniprep kit (Zymo research, Freiburg,
609 Germany) according to the manufacturer's protocol. PCR amplification was performed using
610 H3F3A wild-type and H3F3A-G34W specific primer, respectively. The reaction consisted of
611 2 U Platinum Taq polymerase (Thermo Fisher Scientific, Dreieich, Germany), 0.6 µl MgCl₂
612 (50 mM), 0.4 µl dNTPs (10 mM each), 0.5 µl of each primer (10 µM) and 100 ng genomic
613 DNA as template in a total volume of 20 µl. Samples were incubated at 94°C for 3 min followed
614 by 34 cycles of denaturation at 94°C for 15 s, annealing at 66°C for 20 s and extension at 72°C
615 for 30 s and a final extension step at 72°C for 7 min. PCR products were separated on a 1.6%
616 agarose gel, visualized by Midori Green (Biozym, Hessisch Oldendorf, Germany) and imaged.
617 All primers are listed in **Table S6**.

618 ***Detection of H3F3A-G34W mutation by Sanger sequencing***

619 DNA was extracted using the Quiamp Mini Kit (Qiagen, Hilden, Germany), according to the
620 manufacturer's instructions. The PCR reaction to amplify the mutation spanning *H3F3A*
621 region consisted of 1U Hot Star Taq polymerase (Thermo Fisher Scientific, Dreieich,
622 Germany), 0.8 µl dNTPs (10 mM each, Fermentas, St. Leon-Rot, Germany), 2 µl of each
623 primer (10 µM, Sigma-Aldrich, Taufkirchen, Germany) and 100 ng genomic DNA as template
624 in a total volume of 40 µl. Samples were incubated at 95°C for 15 min followed by 10 cycles
625 of denaturation at 94°C for 45 s, annealing at 61-56°C for 30 s (touchdown -0.5°C/cycle) and
626 extension at 72°C for 60 s followed by 25 cycles of denaturation at 94°C for 45s, annealing at
627 56°C for 30 s and extension at 72°C for 60 s and a final extension step at 72°C for 10 min.
628 Sequencing was performed by GATC Biotech AG, Konstanz, Germany. All primers are listed
629 in **Table S6**.

630 ***Deep targeted resequencing using MiSeq***

631 Deep resequencing of *H3F3A* amplicons was performed as earlier described (49). In brief, DNA
632 was extracted using the Quiamp Mini Kit (Qiagen, Hilden, Germany), according to the

633 manufacturer's instructions. The PCR reaction to amplify the mutation spanning region of the
634 *H3F3A* gene consisted of 0.35U HotStart Q5 polymerase (NEB, Ipswich, USA), 0.6 µl dNTPs
635 (10 mM each, Fermentas, St. Leon-Rot, Germany), 0.3 µl of each primer (10 µM, Sigma-
636 Aldrich, Taufkirchen, Germany) and 25 ng genomic DNA as template in a total volume of
637 25 µl. Samples were incubated at 98°C for 1 min followed by 33 cycles of denaturation at 98°C
638 for 10 s, annealing at 61°C for 30 s and extension at 72°C for 20 s, and a final extension step at
639 72°C for 2 min. Primer sequences are listed in **Table S6**. Primers included a sequence
640 complementary to the primers used for library preparation. Samples were separated on a 1.2%
641 agarose gel and DNA was visualized by Ethidium Bromide. Gel extraction was performed using
642 the Gel extraction Kit (Qiagen, Hilden, Germany). Libraries were prepared using 12.5µl NEB
643 Next HF 2x PCR mix (New England Biolabs, USA) in combination with 0.75 µl of 10 µM IDT
644 primers with Nextera handles and TruSeq Unique-Dual Indices (IDT, USA), 0.3 µl 100x Sybr
645 green and 11 µl of DNA (2ng). Amplification was performed for 6 cycles with the following
646 progression: 98°C 30sec, 98°C 10 sec, 62°C 30 sec and 72°C 15 sec. Libraries were pooled and
647 sequenced on a single flow-cell lane in a or 300 bp paired-end MiSeq run (Illumina, San Diego,
648 CA, USA). The reads were demultiplexed using custom scripts, aligned to the GRCh37
649 assembly using *bwa mem* (v. 0.7.8) (50) with default settings. Base frequencies were read into
650 R and analyzed using package *deepSNV* (v. 1.24.0) (51).

651 ***Detection of genomic structural variations using multiplex fluorescence in situ hybridization,*** 652 ***M-FISH***

653 M-FISH was performed as described in (52). Briefly, seven pools of flow-sorted whole
654 chromosome painting probes were amplified and directly labelled using DEAC-, FITC-, Cy3,
655 TexasRed, and Cy5-conjugated nucleotides or biotin-dUTP and digoxigenin-dUTP,
656 respectively, by degenerative oligonucleotide primed (DOP)-PCR. Prior hybridization,
657 metaphase preparations of the H3.3 WT and H3.3 MUT cells were digested with pepsin (0.5
658 mg/ml; Sigma) in 0.2N HCL (Roth) for 10 min at 37°C, washed in PBS, post-fixed in 1%
659 formaldehyde, dehydrated with a degraded ethanol series and air dried. Slides were denatured
660 in 70% formamide/1 x SSC/15% dextran sulfate for 2 min at 72°C. Hybridization mixture
661 consisting of 50% formamide, 2 x SSC, Cot1-DNA, and labeled DNA probes was denatured
662 for 7 min at 75°C, preannealed for 20 min at 37°C, and hybridized to the denatured metaphase
663 preparations. After 48 hours incubation at 37°C slides were washed at room temperature in 2 x
664 SSC, 3x 5 min, followed by 2x 5 min in 0.2% SSC/0.2% Tween-20 at 56°C. For indirect labeled
665 probes, a two-step immunofluorescence detection was carried out using biotinylated goat anti

666 avidin followed by streptavidin Laser Pro IR790, and rabbit anti digoxin followed by goat anti
667 rabbit Cy5.5. Slides were washed in 4 x SSC/0.2% Tween-20, counterstained with 4.6-
668 diamidino-2-phenylindole (DAPI) and covered with antifade solution. Images of 20 metaphase
669 spreads of the H3.3 WT and H3.3 G34W cells were captured for each fluorochrome using
670 highly specific filter sets (Chroma technology, Brattleboro, VT), and processed using the Leica
671 MCK software (Leica Microsystems Imaging Solutions, Cambridge, UK), respectively.

672 *Chromatin fractionation*

673 Cell pellets were resuspended in lysis buffer (10 mM HEPES pH 7.6, 10 mM KCl, 0.05%
674 NP40) with protease inhibitors and incubated on ice for 30 minutes. Samples were centrifuged
675 at 13200 rpm at 4°C for 10 min and the supernatant was taken as cytosolic fraction. Leftover
676 pellet was further lysed with low salt buffer (10 mM Tris HCl pH 7.5, 3 mM MgCl₂)
677 supplemented with 1% Triton X-100 for 15 min on ice and centrifuged at 13200 rpm at 4°C for
678 10 min. Supernatant was taken as nuclear proteins. Leftover pellet was resuspended in 0.2M
679 HCl, incubated on ice for 20 min and centrifuged at 13200 rpm at 4°C for 10 min. The
680 supernatant was neutralized with 1M Tris HCl buffer (pH8) and used as the chromatin fraction.

681 *Western blot*

682 Proteins were separated using SDS polyacrylamide gel electrophoresis (SDS-PAGE) and
683 analyzed by Western blotting using the antibodies listed in the **Table S5**. Chemiluminescence
684 signals were imaged using Amersham Imager 680 (GE, Boston, USA). For the separation of
685 endogenous histone H3 proteins and ectopically-expressed H3.3-HA-3XFLAG in HEK293T,
686 we used 8-16% Mini-PROTEAN® TGX gels (Bio-Rad Laboratories).

687 *Differentiation and alkaline phosphatase staining*

688 50.000 cells were seeded in 24-well plates. When confluent medium was changed to osteogenic
689 differentiation medium: DMEM high Glucose mit L-Glutamin supplemented with 10%FCS,
690 0.1 μM Dexamethason, 0.17mM Ascorbinsäure-2-phosphat and 10mM β-Glycerophosphat.
691 Alkaline phosphatase activity was analyzed using BCIP/NBT Alkaline Phosphatase Substrate
692 Kit (Vector Laboratories, Burlingame, CA, USA) after 1min of 4% PFA fixation.

693 *Quantification of ALP activity*

694 50,000 cells were seeded in 24-well plates. When confluent media was changed to
695 differentiation media (compare Differentiation and alkaline phosphatase staining) and the first
696 measurement was performed. After CTB as described above, wells were PBS washed and
697 fixated for 60sec with 4% PFA in PBS. After PBS washing 500 μ l Alkaline Phosphatase Yellow
698 (pNpp) Liquid Substrate (System for ELISA, Sigma-Aldrich, USA) was added and incubated
699 at 37°C for 8 minutes. Absorbance was measured at 405nm.

700 ***Whole-genome sequencing and analysis***

701 Read pairs were mapped to the human reference genome (build 37, version hs37d5), using *bwa*
702 *mem* (v. 0.7.8) (50) with minimum base quality threshold set to zero [-T 0] and remaining
703 settings left at default values, followed by coordinate-sorting with *bamsort* (with compression
704 option set to fast REF1) and marking duplicate read pairs with *bammarkduplicates* (with
705 compression option set to best (53)); both are part of *biobambam* package (v.0.0.148) (54).
706 Somatic SNVs were identified with the DKFZ SNV-calling workflow (55). Somatic SNVs and
707 indels in matched tumor normal pairs were identified using the DKFZ core variant calling
708 workflows of the ICGC Pan-cancer Analysis of Whole Genomes (PCAWG) project
709 (<https://dockstore.org/containers/quay.io/pancancer/pcawg-dkfz-workflow>). Tumor and
710 matched control samples were analyzed by *Platypus* (v. 1.0) (56) to identify indel events. SNVs
711 and indels from all samples were annotated using *ANNOVAR* (v. 2017Jul16) (53) according to
712 GENCODE gene annotation (v. 19) and overlapped with variants from dbSNP10 (build 141)
713 and the 1000 Genomes Project database. Genomic structural rearrangements were detected
714 using *SOPHIA* (v.34.0) (<https://bitbucket.org/utoprak/sophia/src>) as described in (57). Briefly,
715 *SOPHIA* uses supplementary alignments as produced by *bwa mem* as indicators of a possible
716 underlying SV. SV candidates are filtered by comparing them to a background control set of
717 sequencing data obtained using normal blood samples from a background population database
718 of 3261 patients from published TCGA and ICGC studies and both published and unpublished
719 DKFZ studies, sequenced using Illumina HiSeq 2000 (100 bp), 2500 (100 bp) and HiSeq X
720 (151 bp) platforms and aligned uniformly using the same workflow as in this study. An SV
721 candidate is discarded if (i) it has more than 85% of read support from low quality reads; (ii)
722 the second breakpoint of the SV was unmappable in the sample and the first breakpoint was
723 detected in 10 or more background control samples; (iii) an SV with two identified breakpoints
724 had one breakpoint present in at least 98 control samples (3% of the control samples); or (iv)
725 both breakpoints have less than 5% read support. Statistics over SVs for 9 samples with
726 matched control and integrated variant analysis over all samples were based on *SOPHIA* calls.

727 Allele-specific copy-number aberrations were detected using *ACEseq* (v. 1.0) (58). SVs called
728 by *SOPHIA* were incorporated to improve genome segmentation.

729 *ATAC-sequencing and analysis*

730 Libraries for ATAC-sequencing were prepared as previously published with modifications
731 (59). Briefly, cells were lysed by 1% NP40 and tagmented at 55°C for 8 minutes in a reaction
732 mix with 2.5µl of TDE1 (Nextera Illumina DNAKit), 25µl Tagmentation buffer (Nextera
733 Illumina DNA Kit) and 25µl of lysed cells. Reaction was stopped by adding 10 µl Guanidium
734 (5 M) and samples were purified using Ampure Beads. Libraries were generated using
735 NEBNext High Fidelity PCR Mix and sequenced on the Illumina HiSeq 2000 platform.
736 Sequencing reads were adaptor-trimmed using *cutadapt* (v. 1.10) (60). Genomic alignments
737 were performed against the human reference genome (hg19, NCBI build 37.1) using *Bowtie2*
738 (v. 2.3.0) (61). The non-default parameters “-q 20 -s” were used. PCR duplicates were removed
739 by *Picard MarkDuplicates* (v. 1.125). Signal tracks were generated using *deepTools* (v. 2.3.3)
740 (62). Peaks were called using *Macs2* (v. 2.1.1.) (63) with the parameters “--nomodel --shift -50
741 --extsize 100 --qvalue 0.01”. All peaks were merged to create a common bed file with read
742 counts before differential analysis using *edgeR* (v. 0.3.16) (64). Gene annotations were made
743 using *ChIPpeakAnno* (v. 3.18.0) (65). Transcription binding motif analysis was performed
744 using *HOMER* (v. 4.9) (66). Motifs with a *P*-value < 0.01 and a ratio of motif to background
745 above 1.1 were defined as significantly enriched.

746 *ChIP-sequencing and analysis*

747 We used the ChIP-mentation protocol (67) to map the genomic distribution of WT and G34W
748 H3.3, total H3, as well as 6 histone modifications (H3K4me1, H3K4me3, H3K9me3, H3K27ac,
749 H3K27me3 and H3K36me3) in a subset of samples (see details in **Table S1**). To validate the
750 specificity of the H3.3 and H3.3 G34W antibody used for ChIP-Seq analysis, we performed a
751 validation experiment with a Histone code peptide array (JPT, Berlin, Germany) containing
752 short peptides that densely cover most of the known histones and their modifications. We did
753 not observe any significant binding of the G34W-specific antibody to H3.3 peptides and a
754 highly specific binding of the H3.3 antibody. Sequence reads were preprocessed using *cutadapt*
755 (v. 1.10) (60) and aligned with *Bowtie2* (v. 2.3.0) (61) with the default command line options.
756 We used *deepTools* (v. 2.3.3) (62) with non-default options “--binSize 10 --extendReads 400 -
757 --normalizeTo1x 2451960000 --ignoreForNormalization chrX” to quantify genomic coverage
758 in fixed-size window intervals for meta-plots and heatmaps.

759 ***Calling of H3.3-G34W enriched regions***

760 We used the Poisson test-based binarization module of the ChromHMM software (v. 1.18) (68)
761 to generate 200bp windows with statistically significant enrichment of the wild-type H3.3 or
762 H3.3-G34W signal over a simulated background. Adjacent windows were merged to generate
763 primary enrichment regions. The regions were filtered against a union of the ENCODE ChIP-
764 seq blacklists.

765 ***Whole-genome bisulfite sequencing and analysis***

766 DNA sequencing libraries were prepared using the TruSeq NanoDNALibrary PrepKit
767 (Illumina, San Diego, CA, USA) following the manufacturer's instructions. Paired-end
768 sequencing (2×150 bp) was performed using one lane of a HiSeqX (Illumina) for every sample.
769 Basic statistics about the sequencing results is given in **Table S1**. Raw reads were processed
770 using *Trimmomatic* (v. 0.36) (69) and aligned against reference sequence of the Genome
771 Research Consortium (v. 37) using *bwa mem* (v. 0.7.8) (50) with default parameters, except for
772 invoking “-T 0”. After alignment duplicates were marked by applying *Picard MarkDuplicates*
773 (v. 1.125). Methylation calling was performed with *MethylDackel* (v. 0.3.0). Coverage filtering
774 still missing). *BSmooth* was used (v. 1.4.0) with default parameters to smooth the methylation
775 profiles in all samples (70). We then used *DSS* (v. 2.27.0) (71) to call DMRs for pairwise
776 comparison between H3.3 WT and H3.3 MUT cells. Regions with at least 3 CpGs, a minimum
777 length of 50bp and a Benjamin-Hochberg corrected *P* value < 0.05 were selected. All DMRs
778 were filtered requiring a minimal mean methylation-value difference of 0.1.

779 ***HumanMethylation450 and HumanMethylationEPIC analysis***

780 Methylation analysis using HumanMethylation450 arrays was performed by the Genomics and
781 Proteomics Core Facility according to the manufacturer's instructions. Profiling of isogenic
782 HeLa cell lines with HumanMethylationEPIC arrays was conducted at Korean Cancer Center
783 according to the manufacturer's instructions. Unnormalized signals (IDAT files) were loaded
784 into R using RnBeads software (v. 2.2.0) (72) and subjected to preprocessing with default
785 option settings. 10,000 sites most variable across all samples were used for both, Principal
786 Component Analysis and clustering analysis, visualized as a heatmap.

787 ***RNA-sequencing and analysis***

788 Poly-A RNA sequencing of GCTB samples was performed according to the standard
789 protocol published elsewhere (73). In brief, total RNA was prepared for each cell line by
790 using the RNAeasy Mini kit (Qiagen, Hilden, Germany) and library preparation was done
791 using TruSeq Stranded mRNA Kit (Illumina), according the manufacturer's instruction.
792 Paired-end 125bp sequencing runs were performed on Illumina Hiseq2000 v4 machines. Raw
793 sequence reads were preprocessed using *cutadapt* (v. 1.10) (60) to remove sequencing
794 primers and adapters. Reads were aligned to the GRCh37 human reference genome with
795 *HISAT2* (v. 2.0.4) (74) with additional non-default parameters "--max-intronlen 20000 --no-
796 unal --dta". Transcripts were assembled and quantified with *StringTie* (v. 1.3.3) (75) with
797 the GRCh37 transcript database. Differential expression analysis was performed using
798 *DeSeq2* (v. 1.18.1) (76). Genes were called differentially expressed at FDR 0.05 and the
799 absolute log-fold difference of greater or equal to one. To compare expression with the data
800 from all TCGA cohorts, the raw data was alternatively quantified with *kallisto*
801 (v. 0.43.1) (77), and log₂-transformed counts were combined with identically processed data
802 downloaded from UCSC Xena portal.

803 **RNA-Seq of the differentiation samples**

804 RNA-seq of the MSC differentiation samples was performed according to the following
805 protocol. 50,000 cells were seeded in 6-well plates in maintenance medium: MesenPRO-RS™
806 (Thermo Fisher Scientific, Massachusetts, USA). After 3 days, medium was replaced with
807 osteogenic differentiation medium: DMEM high glucose supplemented with 10% FBS, 1x non-
808 essential amino acids (NEAA), 2mM L-glutamine, 0.28mM ascorbic acid, 10 mM β
809 glycerophosphate, and 10 nM dexamethasone (Sigma-Aldrich, USA). At each time-point: 0,
810 0.5, 1, 2, 4, 6, 8, 12, 16, 24, 48, 72, 125, 168, 336, 504 (hrs) during osteogenesis, total RNA
811 was isolated using TRIzol (Invitrogen™ Life Technologies, Carlsbad, USA) with the Direct-
812 zol RNA kit (ZymoResearch, USA) according to the manufacturer's instruction. RNA-seq
813 library preparation was carried out using the NEBNext Poly(A) mRNA Magnetic Isolation
814 Module and NEBNext Ultra Directional RNA Library Prep Kit for Illumina (New England
815 Biolabs, USA) according to the manufacturer's instruction. The quantity and quality of the
816 cDNA library were assessed using the Agilent 2200 tapestation (Agilent Technologies, Santa
817 Clara, USA). Paired-end sequencing of the pooled library was carried out using the Illumina
818 NextSeq 500 v2 kit (Illumina, San Diego, USA) according to the manufacturer's instruction.

819 ***qRT-PCR expression analysis***

820 Total RNA was extracted using the RNeasy Mini kit (Qiagen, Hilden, Germany). cDNA was
821 synthesized using random hexamers (Qiagen, Hilden, Germany), and Superscript III Reverse
822 Transcriptase (Invitrogen, Life Technologies, Paisley, UK) according to the manufacturer's
823 instructions. qPCR mixture consisted of 3.5µl Light Cyclers 480 Probe master (Roche
824 Diagnostics, Mannheim, Germany) 1 µl 10 µM Primer mix (Sigma-Aldrich, Taufkirchen,
825 Germany) and 0.05µl UPL probe (Roche Diagnostics, Mannheim, Germany). 2.5 µl of a 1:10
826 dilution of cDNA served as template. Expression analysis was performed on the LightCycler
827 480-2 (Roche) system with the following progression: 10 min 95°C and 45 cycles of 10 s 95°C,
828 20 s 55°C, 1 s 72°C. Alternative to the UPL system, Sybr green qPCR was performed in a total
829 volume of 10 µl including 5µl Prima Quant Mix (Steinbrenner, Wiesenbach, Germany), 0.6 µl
830 Primermix (10 µM each, Sigma-Aldrich) and 2 µl of 1:10 dilution of the template cDNA.
831 Samples were incubated at 95°C for 15 min and 15 s at 95°C, 30 s at 55°C and 10 s at 72°C for
832 45 cycles. Target gene expression was normalized to the housekeeping gene GAPDH using the
833 Δ CT method (relative expression is equal to $2^{-\Delta$ CT). All used primers are listed in **Table S6**.

834 ***Targeted DNA methylation analysis using MassARRAY***

835 Bisulfite treatment was performed with the EZ DNA Methylation Kit from Zymo Research
836 following the manufacturer's protocol. PCR was performed with primers listed in **Table S6**
837 using the Qiagen HotStar Taq (Qiagen, Hilden, Germany). The Shrimp alkaline phosphatase
838 step and in vitro transcription were performed with the EpiTYPER Reagent Set from Agena
839 Bioscience following the manufacturer's protocol. Analyses were performed on the Sequenom
840 Platform (Agena Bioscience, San Diego, USA).

841 ***Gene and genomic feature annotations***

842 Unless specified otherwise, Ensembl transcript and gene annotations were used for the GRCh37
843 assembly (build 87). List of bivalent genes in H3.3 WT was compiled by overlapping the
844 consensus H3.3 WT peaks of H3K27me3 and H3K4me3 within 2kb from a RefSeq TSS. A
845 consensus list of bivalent genes in human ESCs was obtained from (78). A list of PRC2 target
846 genes was found in (79). A comprehensive list of imprinted genes was obtained from (80).
847 Replication timing domains for the annotation of LMDs originates from Repli-Seq data of (81),
848 processed and publicly deposited by *RepliScan* package (82). MSC-specific chromatin states
849 were taken from the 15-state *ChromHMM* model (68) for bone-marrow derived MSCs
850 generated by the Roadmap Epigenomics consortium (21) (sample E026).

851 ***Identification of large-scale methylation domains***

852 DNA methylation data was summarized in 20 kb tiling windows to eliminate the small-scale
853 variability (e.g. related to CpG islands). The changepoints were then called using R package
854 *changepoint* (v. 2.2.2). DNA methylation data was summarized in the obtained segments and
855 the latter were clustered using standard hierarchical clustering resulting in 6 stable clusters.
856 After clustering adjacent segments that belonged to the same cluster were merged. Two smallest
857 clusters were removed since one contained less than 10 segments and the other one exclusively
858 Y-chromosome segments.

859 ***Genomic overlap enrichment analysis***

860 We tested the significance of overlap of differential ATAC-seq peaks, H3.3 G34W
861 incorporation regions and other regions of interest with publicly available genomic annotations
862 using *LOLA* (v. 1.8.0) (83). In the case of ATAC peaks a union of all called peaks in H3.3 WT
863 and H3.3 MUT was used as background to test enrichments at H3.3 G34W gained and lost
864 peaks relative to each other. For H3.3 WT and G34W enrichment regions the background set
865 consisted of regions called for wild-type H3.3 in H3.3 WT and H3.3 MUT groups, as well as
866 the H3.3-G34W regions. We used the *LOLA* core database for most of the analyses. Enrichment
867 of repeat elements was based on a custom *LOLA* database created using the UCSC Repeat
868 Masker track (<http://www.repeatmasker.org>).

869 ***Gene set overrepresentation and enrichment analysis***

870 We used gene sets from the Molecular Signatures Database v. 6.2 (MSigDB) (84) to test for the
871 overrepresentation of DEGs or genes associated with differential ATAC peaks, and the gene
872 set enrichment analysis (GSEA) of DEGs. Overrepresentation analysis was performed with the
873 help of R package *GeneOverlap* (v. 1.14.0). GSEA was performed with *fgsea* package (v. 1.4.1)
874 (85) using 10,000 permutations.

875 ***Availability of data and code***

876 All raw sequencing data from WGS, WGBS, ATAC-seq, ChIP-seq, RNA-seq and deep targeted
877 resequencing were obtained from patient samples (therefore restricted) and are being uploaded
878 to European Genome-Phenome Archive (EGA): EGA: EGAS00001003730. Processed
879 sequencing data and microarray data are being uploaded to ArrayExpress database (E-MTAB-

880 7184). All presented results were obtained with the use of published and publicly available
881 software tools introduced in the Methods section.

882

883 References

884

- 885 1. Mohammad F, Helin K. Oncohistones: drivers of pediatric cancers. *Genes Dev* **2017**;31(23-24):2313-24 doi 10.1101/gad.309013.117.
- 886 2. Nacev BA, Feng L, Bagert JD, Lemiesz AE, Gao J, Soshnev AA, *et al.* The expanding
887 landscape of 'oncohistone' mutations in human cancers. *Nature* **2019** doi
888 10.1038/s41586-019-1038-1.
- 889 3. Bennett RL, Bele A, Small EC, Will CM, Nabet B, Oyer JA, *et al.* A Mutation in Histone
890 H2B Represents a New Class of Oncogenic Driver. *Cancer Discov* **2019**;9(10):1438-51
891 doi 10.1158/2159-8290.CD-19-0393.
- 892 4. Szenker E, Ray-Gallet D, Almouzni G. The double face of the histone variant H3.3.
893 *Cell Res* **2011**;21(3):421-34 doi 10.1038/cr.2011.14.
- 894 5. Goldberg AD, Banaszynski LA, Noh KM, Lewis PW, Elsaesser SJ, Stadler S, *et al.*
895 Distinct factors control histone variant H3.3 localization at specific genomic regions.
896 *Cell* **2010**;140(5):678-91 doi 10.1016/j.cell.2010.01.003.
- 897 6. Voon HP, Wong LH. New players in heterochromatin silencing: histone variant H3.3
898 and the ATRX/DAXX chaperone. *Nucleic acids research* **2016**;44(4):1496-501 doi
899 10.1093/nar/gkw012.
- 900 7. Banaszynski LA, Wen D, Dewell S, Whitcomb SJ, Lin M, Diaz N, *et al.* Hira-dependent
901 histone H3.3 deposition facilitates PRC2 recruitment at developmental loci in ES cells.
902 *Cell* **2013**;155(1):107-20 doi 10.1016/j.cell.2013.08.061.
- 903 8. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, *et al.* A bivalent
904 chromatin structure marks key developmental genes in embryonic stem cells. *Cell*
905 **2006**;125(2):315-26 doi 10.1016/j.cell.2006.02.041.
- 906 9. Schwartzenuber J, Korshunov A, Liu XY, Jones DT, Pfaff E, Jacob K, *et al.* Driver
907 mutations in histone H3.3 and chromatin remodelling genes in paediatric glioblastoma.
908 *Nature* **2012**;482(7384):226-31 doi 10.1038/nature10833.
- 909 10. Behjati S, Tarpey PS, Presneau N, Scheipl S, Pillay N, Van Loo P, *et al.* Distinct H3F3A
910 and H3F3B driver mutations define chondroblastoma and giant cell tumor of bone.
911 *Nature genetics* **2013**;45(12):1479-82 doi 10.1038/ng.2814.
- 912 11. Fang D, Gan H, Lee JH, Han J, Wang Z, Riester SM, *et al.* The histone H3.3K36M
913 mutation reprograms the epigenome of chondroblastomas. *Science*
914 **2016**;352(6291):1344-8 doi 10.1126/science.aae0065.
- 915 12. Chan KM, Fang D, Gan H, Hashizume R, Yu C, Schroeder M, *et al.* The histone
916 H3.3K27M mutation in pediatric glioma reprograms H3K27 methylation and gene
917 expression. *Genes Dev* **2013**;27(9):985-90 doi 10.1101/gad.217778.113.
- 918 13. Lewis PW, Muller MM, Koletsky MS, Cordero F, Lin S, Banaszynski LA, *et al.*
919 Inhibition of PRC2 Activity by a Gain-of-Function H3 Mutation Found in Pediatric
920 Glioblastoma. *Science* **2013** doi 10.1126/science.1232245.
- 921 14. Bender S, Tang Y, Lindroth AM, Hovestadt V, Jones DT, Kool M, *et al.* Reduced
922 H3K27me3 and DNA hypomethylation are major drivers of gene expression in K27M
923 mutant pediatric high-grade gliomas. *Cancer cell* **2013**;24(5):660-72 doi
924 10.1016/j.ccr.2013.10.006.
- 925 15. Lu C, Jain SU, Hoelper D, Bechet D, Molden RC, Ran L, *et al.* Histone H3K36
926 mutations promote sarcomagenesis through altered histone methylation landscape.
927 *Science* **2016**;352(6287):844-9 doi 10.1126/science.aac7272.
- 928

- 929 16. Shi L, Shi J, Shi X, Li W, Wen H. Histone H3.3 G34 Mutations Alter Histone H3K36
930 and H3K27 Methylation In Cis. *J Mol Biol* **2018** doi 10.1016/j.jmb.2018.04.014.
- 931 17. Amanatullah DF, Clark TR, Lopez MJ, Borys D, Tamurian RM. Giant cell tumor of
932 bone. *Orthopedics* **2014**;37(2):112-20 doi 10.3928/01477447-20140124-08.
- 933 18. Raskin KA, Schwab JH, Mankin HJ, Springfield DS, Hornicek FJ. Giant cell tumor of
934 bone. *J Am Acad Orthop Surg* **2013**;21(2):118-26 doi 10.5435/JAAOS-21-02-118.
- 935 19. Goldring SR, Roelke MS, Petrisson KK, Bhan AK. Human giant cell tumors of bone
936 identification and characterization of cell types. *The Journal of clinical investigation*
937 **1987**;79(2):483-91 doi 10.1172/JCI112838.
- 938 20. International Cancer Genome Consortium PedBrain Tumor P. Recurrent MET fusion
939 genes represent a drug target in pediatric glioblastoma. *Nat Med* **2016**;22(11):1314-20
940 doi 10.1038/nm.4204.
- 941 21. Roadmap Epigenomics C, Kundaje A, Meuleman W, Ernst J, Bilenky M, Yen A, *et al.*
942 Integrative analysis of 111 reference human epigenomes. *Nature* **2015**;518(7539):317-
943 30 doi 10.1038/nature14248.
- 944 22. Zhou W, Dinh HQ, Ramjan Z, Weisenberger DJ, Nicolet CM, Shen H, *et al.* DNA
945 methylation loss in late-replicating domains is linked to mitotic cell division. *Nature*
946 *genetics* **2018**;50(4):591-602 doi 10.1038/s41588-018-0073-4.
- 947 23. Moskovszky L, Szuhai K, Krenacs T, Hogendoorn PC, Szendroi M, Benassi MS, *et al.*
948 Genomic instability in giant cell tumor of bone. A study of 52 cases using DNA ploidy,
949 relocalization FISH, and array-CGH analysis. *Genes Chromosomes Cancer*
950 **2009**;48(6):468-79 doi 10.1002/gcc.20656.
- 951 24. Lim J, Park JH, Baude A, Yoo Y, Lee YK, Schmidt CR, *et al.* The histone variant H3.3
952 G34W substitution in giant cell tumor of the bone link chromatin and RNA processing.
953 *Sci Rep* **2017**;7(1):13459 doi 10.1038/s41598-017-13887-y.
- 954 25. Balke M. Denosumab treatment of giant cell tumour of bone. *The Lancet Oncology*
955 **2013**;14(9):801-2 doi 10.1016/S1470-2045(13)70291-2.
- 956 26. Boyle WJ, Simonet WS, Lacey DL. Osteoclast differentiation and activation. *Nature*
957 **2003**;423(6937):337-42 doi 10.1038/nature01658.
- 958 27. Kieslinger M, Folberth S, Dobrev G, Dorn T, Croci L, Erben R, *et al.* EBF2 regulates
959 osteoblast-dependent differentiation of osteoclasts. *Dev Cell* **2005**;9(6):757-67 doi
960 10.1016/j.devcel.2005.10.009.
- 961 28. Boyce BF, Xing L, Chen D. Osteoprotegerin, the bone protector, is a surprising target
962 for beta-catenin signaling. *Cell Metab* **2005**;2(6):344-5 doi
963 10.1016/j.cmet.2005.11.011.
- 964 29. Wolock SL, Krishnan I, Tenen DE, Matkins V, Camacho V, Patel S, *et al.* Mapping
965 Distinct Bone Marrow Niche Populations and Their Differentiation Paths. *Cell Rep*
966 **2019**;28(2):302-11 e5 doi 10.1016/j.celrep.2019.06.031.
- 967 30. Zardo G, Tiirikainen MI, Hong C, Misra A, Feuerstein BG, Volik S, *et al.* Integrated
968 genomic and epigenomic analyses pinpoint biallelic gene inactivation in tumors. *Nature*
969 *genetics* **2002**;32(3):453-8 doi 10.1038/ng1007.
- 970 31. Shapira SN, Lim HW, Rajakumari S, Sakers AP, Ishibashi J, Harms MJ, *et al.* EBF2
971 transcriptionally regulates brown adipogenesis via the histone reader DPF3 and the BAF
972 chromatin remodeling complex. *Genes Dev* **2017**;31(7):660-73 doi
973 10.1101/gad.294405.116.
- 974 32. Voon HP, Hughes JR, Rode C, De La Rosa-Velazquez IA, Jenuwein T, Feil R, *et al.*
975 ATRX Plays a Key Role in Maintaining Silencing at Interstitial Heterochromatic Loci
976 and Imprinted Genes. *Cell Rep* **2015**;11(3):405-18 doi 10.1016/j.celrep.2015.03.036.
- 977 33. Harutyunyan AS, Krug B, Chen H, Papillon-Cavanagh S, Zeinieh M, De Jay N, *et al.*
978 H3K27M induces defective chromatin spread of PRC2-mediated repressive

- 979 H3K27me2/me3 and is essential for glioma tumorigenesis. *Nat Commun*
980 **2019**;10(1):1262 doi 10.1038/s41467-019-09140-x.
- 981 34. Hardouin SN, Guo R, Romeo PH, Nagy A, Aubin JE. Impaired mesenchymal stem cell
982 differentiation and osteoclastogenesis in mice deficient for Igf2-P2 transcripts.
983 *Development* **2011**;138(2):203-13 doi 10.1242/dev.054916.
- 984 35. Xu JC, Wu GH, Zhou LL, Yang XJ, Liu JT. Leptin improves osteoblast differentiation
985 of human bone marrow stroma stem cells. *Eur Rev Med Pharmacol Sci*
986 **2016**;20(16):3507-13.
- 987 36. Huang L, Teng XY, Cheng YY, Lee KM, Kumta SM. Expression of preosteoblast
988 markers and Cbfa-1 and Osterix gene transcripts in stromal tumour cells of giant cell
989 tumour of bone. *Bone* **2004**;34(3):393-401 doi 10.1016/j.bone.2003.10.013.
- 990 37. Lau CP, Kwok JS, Tsui JC, Huang L, Yang KY, Tsui SK, *et al.* Genome-Wide
991 Transcriptome Profiling of the Neoplastic Giant Cell Tumor of Bone Stromal Cells by
992 RNA Sequencing. *J Cell Biochem* **2017**;118(6):1349-60 doi 10.1002/jcb.25792.
- 993 38. Zhang Y, Shan C-M, Wang J, Bao K, Tong L, Jia S. Molecular basis for the role of
994 oncogenic histone mutations in modulating H3K36 methylation. *Scientific Reports*
995 **2017**;7:43906 doi 10.1038/srep43906.
- 996 39. Sturm D, Witt H, Hovestadt V, Khuong-Quang DA, Jones DT, Konermann C, *et al.*
997 Hotspot Mutations in H3F3A and IDH1 Define Distinct Epigenetic and Biological
998 Subgroups of Glioblastoma. *Cancer cell* **2012**;22(4):425-37 doi
999 10.1016/j.ccr.2012.08.024.
- 1000 40. de la Rica L, Rodriguez-Ubreva J, Garcia M, Islam AB, Urquiza JM, Hernando H, *et*
1001 *al.* PU.1 target genes undergo Tet2-coupled demethylation and DNMT3b-mediated
1002 methylation in monocyte-to-osteoclast differentiation. *Genome Biol* **2013**;14(9):R99
1003 doi 10.1186/gb-2013-14-9-r99.
- 1004 41. Rauch A, Haakonsson AK, Madsen JGS, Larsen M, Forss I, Madsen MR, *et al.*
1005 Osteogenesis depends on commissioning of a network of stem cell transcription factors
1006 that act as repressors of adipogenesis. *Nature genetics* **2019** doi 10.1038/s41588-019-
1007 0359-1.
- 1008 42. Lau YS, Sabokbar A, Gibbons CL, Giele H, Athanasou N. Phenotypic and molecular
1009 studies of giant-cell tumors of bone and soft tissue. *Hum Pathol* **2005**;36(9):945-54 doi
1010 10.1016/j.humpath.2005.07.005.
- 1011 43. Murata A, Fujita T, Kawahara N, Tsuchiya H, Tomita K. Osteoblast lineage properties
1012 in giant cell tumors of bone. *J Orthop Sci* **2005**;10(6):581-8 doi 10.1007/s00776-005-
1013 0946-0.
- 1014 44. Chiellini C, Grenningloh G, Cochet O, Scheideler M, Trajanoski Z, Ailhaud G, *et al.*
1015 Stathmin-like 2, a developmentally-associated neuronal marker, is expressed and
1016 modulated during osteogenesis of human mesenchymal stem cells. *Biochem Biophys*
1017 *Res Commun* **2008**;374(1):64-8 doi 10.1016/j.bbrc.2008.06.121.
- 1018 45. Liu L, Aleksandrowicz E, Fan P, Schonsiegel F, Zhang Y, Sahr H, *et al.* Enrichment of
1019 c-Met⁺ tumorigenic stromal cells of giant cell tumor of bone and targeting by
1020 cabozantinib. *Cell Death Dis* **2014**;5:e1471 doi 10.1038/cddis.2014.440.
- 1021 46. Wulling M, Delling G, Kaiser E. The origin of the neoplastic stromal cell in giant cell
1022 tumor of bone. *Hum Pathol* **2003**;34(10):983-93 doi 10.1053/s0046-8177(03)00413-1.
- 1023 47. Baubec T, Colombo DF, Wirbelauer C, Schmidt J, Burger L, Krebs AR, *et al.* Genomic
1024 profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation.
1025 *Nature* **2015**;520(7546):243-7 doi 10.1038/nature14176.
- 1026 48. Dhayalan A, Rajavelu A, Rathert P, Tamas R, Jurkowska RZ, Ragozin S, *et al.* The
1027 Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA
1028 methylation. *J Biol Chem* **2010**;285(34):26114-20 doi 10.1074/jbc.M109.089433.

- 1029 49. Souren NY, Gerdes LA, Kumpfel T, Lutsik P, Klopstock T, Hohlfeld R, *et al.*
1030 Mitochondrial DNA Variation and Heteroplasmy in Monozygotic Twins Clinically
1031 Discordant for Multiple Sclerosis. *Hum Mutat* **2016**;37(8):765-75 doi
1032 10.1002/humu.23003.
- 1033 50. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler
1034 transform. *Bioinformatics* **2009**;25(14):1754-60 doi 10.1093/bioinformatics/btp324.
- 1035 51. Gerstung M, Beisel C, Rechsteiner M, Wild P, Schraml P, Moch H, *et al.* Reliable
1036 detection of subclonal single-nucleotide variants in tumour cell populations. *Nat*
1037 *Commun* **2012**;3:811 doi 10.1038/ncomms1814.
- 1038 52. Geigl JB, Uhrig S, Speicher MR. Multiplex-fluorescence in situ hybridization for
1039 chromosome karyotyping. *Nat Protoc* **2006**;1(3):1172-84 doi 10.1038/nprot.2006.160.
- 1040 53. Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants
1041 from high-throughput sequencing data. *Nucleic acids research* **2010**;38(16):e164 doi
1042 10.1093/nar/gkq603.
- 1043 54. Tischler G, Leonard S. biobambam: tools for read pair collation based algorithms on
1044 BAM files. *Source Code for Biology and Medicine* **2014**;9(1):13 doi 10.1186/1751-
1045 0473-9-13.
- 1046 55. Jones DT, Hutter B, Jager N, Korshunov A, Kool M, Warnatz HJ, *et al.* Recurrent
1047 somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. *Nature genetics*
1048 **2013**;45(8):927-32 doi 10.1038/ng.2682.
- 1049 56. Rimmer A, Phan H, Mathieson I, Iqbal Z, Twigg SRF, Consortium WGS, *et al.*
1050 Integrating mapping-, assembly- and haplotype-based approaches for calling variants in
1051 clinical sequencing applications. *Nature genetics* **2014**;46(8):912-8 doi
1052 10.1038/ng.3036.
- 1053 57. Sahm F, Toprak UH, Hubschmann D, Kleinheinz K, Buchhalter I, Sill M, *et al.*
1054 Meningiomas induced by low-dose radiation carry structural variants of NF2 and a
1055 distinct mutational signature. *Acta Neuropathol* **2017**;134(1):155-8 doi
1056 10.1007/s00401-017-1715-9.
- 1057 58. Kleinheinz K, Bludau I, Huebschmann D, Heinold M, Kensche P, Gu Z, *et al.* ACEseq
1058 - allele specific copy number estimation from whole genome sequencing. *bioRxiv* **2017**
1059 doi 10.1101/210807.
- 1060 59. Buenrostro JD, Giresi PG, Zaba LC, Chang HY, Greenleaf WJ. Transposition of native
1061 chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding
1062 proteins and nucleosome position. *Nat Methods* **2013**;10(12):1213-8 doi
1063 10.1038/nmeth.2688.
- 1064 60. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing
1065 reads. *2011* **2011**;17(1):3 doi 10.14806/ej.17.1.200.
- 1066 61. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods*
1067 **2012**;9(4):357-9 doi 10.1038/nmeth.1923.
- 1068 62. Ramirez F, Dundar F, Diehl S, Gruning BA, Manke T. deepTools: a flexible platform
1069 for exploring deep-sequencing data. *Nucleic acids research* **2014**;42(Web Server
1070 issue):W187-91 doi 10.1093/nar/gku365.
- 1071 63. Zhang Y, Liu T, Meyer CA, Eeckhoutte J, Johnson DS, Bernstein BE, *et al.* Model-based
1072 analysis of ChIP-Seq (MACS). *Genome Biol* **2008**;9(9):R137 doi 10.1186/gb-2008-9-
1073 9-r137.
- 1074 64. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor RNA-
1075 Seq experiments with respect to biological variation. *Nucleic acids research*
1076 **2012**;40(10):4288-97 doi 10.1093/nar/gks042.
- 1077 65. Zhu LJ. Integrative analysis of ChIP-chip and ChIP-seq dataset. *Methods Mol Biol*
1078 **2013**;1067:105-24 doi 10.1007/978-1-62703-607-8_8.

- 1079 66. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, *et al.* Simple combinations
1080 of lineage-determining transcription factors prime cis-regulatory elements required for
1081 macrophage and B cell identities. *Mol Cell* **2010**;38(4):576-89 doi
1082 10.1016/j.molcel.2010.05.004.
- 1083 67. Schmidl C, Rendeiro AF, Sheffield NC, Bock C. ChIPmentation: fast, robust, low-input
1084 ChIP-seq for histones and transcription factors. *Nat Methods* **2015**;12(10):963-5 doi
1085 10.1038/nmeth.3542.
- 1086 68. Ernst J, Kellis M. Chromatin-state discovery and genome annotation with ChromHMM.
1087 *Nat Protoc* **2017**;12(12):2478-92 doi 10.1038/nprot.2017.124.
- 1088 69. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina
1089 sequence data. *Bioinformatics* **2014**;30(15):2114-20 doi
1090 10.1093/bioinformatics/btu170.
- 1091 70. Hansen KD, Langmead B, Irizarry RA. BSmooth: from whole genome bisulfite
1092 sequencing reads to differentially methylated regions. *Genome Biol* **2012**;13(10):R83
1093 doi 10.1186/gb-2012-13-10-r83.
- 1094 71. Park Y, Wu H. Differential methylation analysis for BS-seq data under general
1095 experimental design. *Bioinformatics* **2016**;32(10):1446-53 doi
1096 10.1093/bioinformatics/btw026.
- 1097 72. Assenov Y, Muller F, Lutsik P, Walter J, Lengauer T, Bock C. Comprehensive analysis
1098 of DNA methylation data with RnBeads. *Nat Methods* **2014**;11(11):1138-40 doi
1099 10.1038/nmeth.3115.
- 1100 73. Weischenfeldt J, Simon R, Feuerbach L, Schlangen K, Weichenhan D, Minner S, *et al.*
1101 Integrative genomic analyses reveal an androgen-driven somatic alteration landscape in
1102 early-onset prostate cancer. *Cancer cell* **2013**;23(2):159-70 doi
1103 10.1016/j.ccr.2013.01.002.
- 1104 74. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory
1105 requirements. *Nat Methods* **2015**;12(4):357-60 doi 10.1038/nmeth.3317.
- 1106 75. Pertea M, Pertea GM, Antonescu CM, Chang TC, Mendell JT, Salzberg SL. StringTie
1107 enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat*
1108 *Biotechnol* **2015**;33(3):290-5 doi 10.1038/nbt.3122.
- 1109 76. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for
1110 RNA-seq data with DESeq2. *Genome Biol* **2014**;15(12):550 doi 10.1186/s13059-014-
1111 0550-8.
- 1112 77. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq
1113 quantification. *Nat Biotechnol* **2016**;34(5):525-7 doi 10.1038/nbt.3519.
- 1114 78. Court F, Arnaud P. An annotated list of bivalent chromatin regions in human ES cells:
1115 a new tool for cancer epigenetic research. *Oncotarget* **2017**;8(3):4110-24 doi
1116 10.18632/oncotarget.13746.
- 1117 79. Bracken AP, Dietrich N, Pasini D, Hansen KH, Helin K. Genome-wide mapping of
1118 Polycomb target genes unravels their roles in cell fate transitions. *Genes Dev*
1119 **2006**;20(9):1123-36 doi 10.1101/gad.381706.
- 1120 80. Skaar DA, Li Y, Bernal AJ, Hoyo C, Murphy SK, Jirtle RL. The human imprintome:
1121 regulatory mechanisms, methods of ascertainment, and roles in disease susceptibility.
1122 *ILAR J* **2012**;53(3-4):341-58 doi 10.1093/ilar.53.3-4.341.
- 1123 81. Hansen RS, Thomas S, Sandstrom R, Canfield TK, Thurman RE, Weaver M, *et al.*
1124 Sequencing newly replicated DNA reveals widespread plasticity in human replication
1125 timing. *Proceedings of the National Academy of Sciences of the United States of*
1126 *America* **2010**;107(1):139-44 doi 10.1073/pnas.0912402107.
- 1127 82. Zynda GJ, Song J, Concia L, Wear EE, Hanley-Bowdoin L, Thompson WF, *et al.*
1128 Repliscan: a tool for classifying replication timing regions. *BMC Bioinformatics*
1129 **2017**;18(1):362 doi 10.1186/s12859-017-1774-x.

- 1130 83. Sheffield NC, Bock C. LOLA: enrichment analysis for genomic region sets and
1131 regulatory elements in R and Bioconductor. *Bioinformatics* **2016**;32(4):587-9 doi
1132 10.1093/bioinformatics/btv612.
- 1133 84. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, *et al.*
1134 Gene set enrichment analysis: a knowledge-based approach for interpreting genome-
1135 wide expression profiles. *Proceedings of the National Academy of Sciences of the*
1136 *United States of America* **2005**;102(43):15545-50 doi 10.1073/pnas.0506580102.
- 1137 85. Sergushichev A. An algorithm for fast preranked gene set enrichment analysis using
1138 cumulative statistic calculation. *bioRxiv* **2016** doi 10.1101/060012.

1139 **Acknowledgements**

1140 We acknowledge the Genomics and Proteomics Core Facility of the German Cancer Research
1141 Center (head Stephan Wiemann) for their assistance with the high-throughput sequencing, and
1142 the Omics and Data Management Core Facility (heads Jürgen Eils, Christian Lawerenz and
1143 Ivo Buchhalter) for the help with processing and management of the sequencing data. We are
1144 thankful to Christa Winkelmann (Institute of Pathology, Erlangen) for excellent performance
1145 of immunostainings, Cansu Cirzi for the help with ChIP-mentation libraries and Oliver
1146 Mücke for his help with MassARRAY. We thank ActiveMotif for providing a novel specific
1147 antibody against H3.3-G34W, as well as JPT (Pavlo Holenya and Ulf Reimer) for kindly
1148 providing a histone peptide array for antibody validation. We thank Stephan Pfister and David
1149 Jones for the permission to use PedGBM cohort, and Ana Banito for a thoughtful feedback
1150 about the manuscript.

1151 **Author Contributions**

1152 C.P. and A.L. supervised the study. F.H., J.Z. and J.F. provided GCTB samples and performed
1153 their initial characterization. A.B., S.Ö., D.M., J.F., A. Jau., A.K., M.H., J.L., M. Schl.,
1154 M.Schuh., D.V., F.G., S.H., V.H.N., S.L., J.H.P., Y.J.P., D.K., C.J., F.R. and D.W. performed
1155 experimental work. P.L., S.Ö., A.M., U.T., R.T., J.H., K.B., C.J., U.O. performed
1156 bioinformatical data analysis and interpretation. P.L., A.B., S.Ö., D.M., A.K., C.P. and A.L.
1157 wrote the manuscript.

1158 **Competing interests' statement**

1159 The authors declare no potential conflicts of interest.

1160 Requests for materials and other correspondence should be directed to corresponding authors.