

## ARTICLE

Received 2 Apr 2014 | Accepted 29 Oct 2014 | Published 4 Dec 2014

DOI: 10.1038/ncomms6690

# Extrachromosomal driver mutations in glioblastoma and low-grade glioma

Sergey Nikolaev<sup>1,\*</sup>, Federico Santoni<sup>1,2,\*</sup>, Marco Garieri<sup>1</sup>, Periklis Makrythanasis<sup>1,2</sup>, Emilie Falconnet<sup>1</sup>, Michel Guipponi<sup>2</sup>, Anne Vannier<sup>2</sup>, Ivan Radovanovic<sup>3,4</sup>, Frederique Bena<sup>2</sup>, Françoise Forestier<sup>2</sup>, Karl Schaller<sup>3,4</sup>, Valerie Dutoit<sup>5</sup>, Virginie Clement-Schatlo<sup>3,4</sup>, Pierre-Yves Dietrich<sup>5</sup> & Stylianos E. Antonarakis<sup>1,2,6</sup>

Alteration of the number of copies of double minutes (DMs) with oncogenic *EGFR* mutations in response to tyrosine kinase inhibitors is a novel adaptive mechanism of glioblastoma. Here we provide evidence that such mutations in DMs, called here amplification-linked extrachromosomal mutations (ALEMs), originate extrachromosomally and could therefore be completely eliminated from the cancer cells. By exome sequencing of seven glioblastoma patients we reveal ALEMs in *EGFR*, *PDGFRA* and other genes. These mutations together with DMs are lost by cancer cells in culture. We confirm the extrachromosomal origin of such mutations by showing that wild-type and mutated DMs may coexist in the same tumour. Analysis of 4,198 tumours suggests the presence of ALEMs across different tumour types with the highest prevalence in glioblastomas and low-grade gliomas. The extrachromosomal nature of ALEMs explains the observed drastic changes in the amounts of mutated oncogenes (like *EGFR* or *PDGFRA*) in glioblastoma in response to environmental changes.

<sup>1</sup>Department of Genetic Medicine and Development, University of Geneva Medical School, 1 rue Michel Servet, 1211 Geneva 4, Switzerland. <sup>2</sup>Geneva University Hospitals—HUG, Service of Genetic Medicine, 4 Rue Gabrielle-Perret-Gentil, 1211 Geneva 4, Switzerland. <sup>3</sup>Department of Clinical Neuroscience, University of Geneva Medical School, 1 rue Michel Servet, 1211 Geneva 4, Switzerland. <sup>4</sup>Department of Neurosurgery, Geneva University Hospitals—HUG, 4 Rue Gabrielle-Perret-Gentil, 1211 Geneva 4, Switzerland. <sup>5</sup>Center of Oncology, Geneva University Hospitals—HUG, 4 Rue Gabrielle-Perret-Gentil, 1211 Geneva 4, Switzerland. <sup>6</sup>IGE3 institute of Genetics and Genomics of Geneva, 1 rue Michel Servet, 1211 Geneva 4, Switzerland. \* These authors contributed equally to this work. Correspondence and requests for materials should be addressed to S.N. (email: sergey.nikolaev@unige.ch) or to F.S. (email: federico.santoni@unige.ch) or to S.E.A. (email: stylianos.antonarakis@unige.ch).

Human cancers are characterised by the presence of genomic instability<sup>1,2</sup>. One form of genomic instability which makes tumours susceptible to acquiring point mutations is often referred to as the mutator phenotype<sup>3–5</sup>. It is expected that the probability of acquisition of gain-of-function mutations in oncogenes is considerably lower than the probability of acquisition of loss-of-function mutations in tumour suppressor genes, because the first occurs in a few critical sites, whereas the latter occurs anywhere in the coding sequence of the gene. Despite that, oncogenes harbour ~80% of the driver mutations<sup>6</sup>. This could be partially explained by the frequent genomic focal amplification (FA) of some oncogenes (that is, RTKs like *EGFR*, *PDGFRA*)<sup>7,8</sup> which may increase the probability of acquisition of gain-of-function mutations<sup>9–14</sup>.

Several sources of evidence suggest that regions of genomic rearrangements including FAs in cancer may be associated with high mutation loads. In the germline, the DNA mutation loads depend on the number of replication cycles<sup>15,16</sup>, and genomic rearrangements frequently coexist with the concomitant mutations<sup>17</sup>. Notably, the break-induced replication repair pathway<sup>18,19</sup> was recently suggested to be responsible for the frequent genomic duplications in human cancers<sup>20</sup>.

Double minutes (DMs) and homogeneously staining regions (HSR) are the cytogenetic hallmarks of genomic FAs in cancer<sup>21</sup>. DMs are extrachromosomal circular DNA molecules without centromere and are found in the nucleus or cytoplasm enveloped by a nuclear-like membrane (micronuclei) allowing the transcription and DNA replication<sup>22</sup>. The absence of a centromere in DMs results in a random segregation between daughter cells through ‘hitchhiking’<sup>23</sup>. DMs were found in many tumour types including glioblastomas (GBM)<sup>13,24</sup>, low-grade gliomas (LGG), ovary<sup>25</sup>, breast<sup>26</sup>, lung<sup>27</sup>, colon<sup>28,29</sup> and neuroblastoma<sup>25,30</sup>. The probable mechanism of DM formation involves non-homologous end joining<sup>31–33</sup> which is active in different tumours, especially in those with defective homologous recombination<sup>34</sup>. Therefore the mutation load in DMs is expected to be higher than that in chromosomal DNA, because the repair of DNA damage by non-homologous end joining results in acquisition of point mutations and small indels and the DNA damage repair mechanism is less efficient in the micronuclei compared with the nucleus<sup>29,35</sup>. It is also expected that the mutational load in the regions amplified as DMs may be considerably higher than that in the chromosomal non-amplified DNA as this kind of amplification may reach hundreds of copies per cell or more<sup>36</sup>.

In this work we describe a novel class of mutations in cancer, amplification-linked extrachromosomal mutations (ALEMs) which occur in DMs. ALEMs are detected in GBMs because they disappear from tumour cells during cell culture. While ALEMs are most prevalent in GBMs and LGG they also exist in other tumour types. Based on these findings, we propose a novel

mechanism of acquisition of gain-of-function extrachromosomal mutations mediated by FAs which may underlie the acquisition of resistance to therapies.

Results

**Amplification-linked mutations.** We investigated the genetic heterogeneity of GBM by exome sequencing of primary tumour fragments and derived gliomaspheres from seven patients. GBMs were selected for the study, because these tumours are characterised by frequent FAs in their genomes (> 50% of the cases) predominantly in the form of DMs<sup>37</sup>. We took advantage of the fact that the cultured GBM spheres in certain conditions can lose DMs<sup>38,39</sup>, in order to monitor the fate of point mutations within FAs.

We observed eight mutations present within FAs in the primary tumours, and remarkably, all of them were lost in the gliomaspheres after several passages (Table 1). Neither LOH nor chromosomal abnormalities have been detected in the corresponding regions in gliomaspheres. Notably, one individual had four mutations associated with FAs in the primary tumour which were lost in the spheres (GBM IV-34) (Fig. 1). One of these mutations, *PDGFRA* N659K (COSM22415), was within a 5.1 Mb amplification (chr4:52.86–57.98 Mb), while the other three: *MARS* p.G888E, *DDIT3* p.P11S and *DDIT3* p.S31L were within an amplification of 1.5 Mb on chromosome 12 (chr12:57.86–59.31 Mb). The fraction of reads supporting these mutations was close to 100% in the primary tumour sample (86, 97, 99 and 100%, respectively). In addition we performed fluorescence *in situ* hybridization (FISH) analysis of the interphase nuclei of the GBM IV-34 primary tumour cells and gliomaspheres. The number and localisations of the *PDGFRA* signals in the primary GBM cells as well as their loss in the cell culture strongly suggests that this FA is present in a form of DMs (Supplementary Fig. 1a,b).

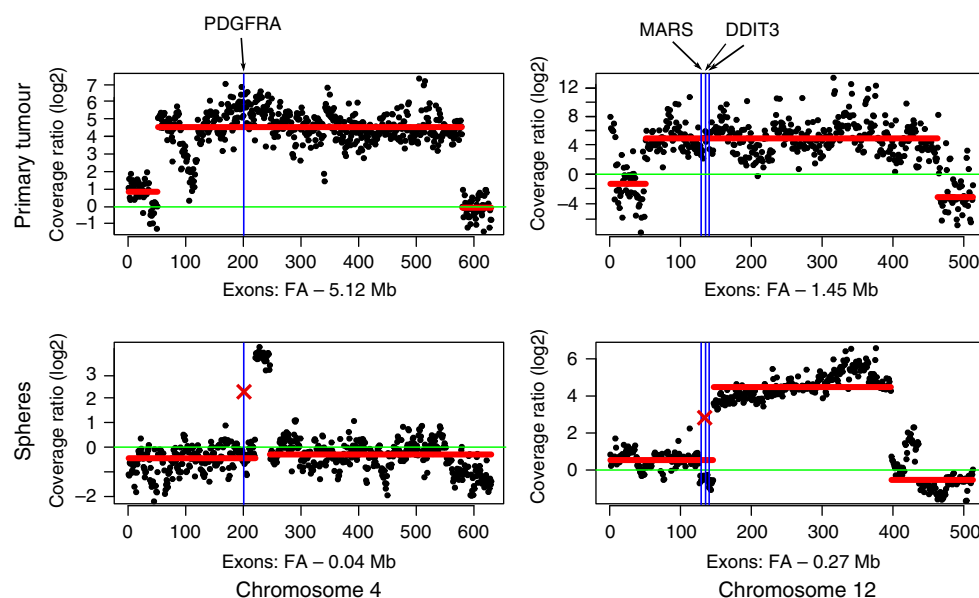
Similar observations were made comparing variants in primary tumours versus spheres from patients GBM IV-19 and GBM IV-39. In both tumours, the *EGFR* (p.A289V and p.S227Y) mutations and FAs were present in the primary tissues and gliomaspheres at passage 0; however, they were completely lost at later passages (Table 1). The DMs amplifications containing the *EGFR* locus in the primary tumours and their loss in gliomaspheres in both GBM IV-19 and GBM IV-39 were confirmed by FISH analyses (Fig. 2, Supplementary Fig. 1c,d).

In addition, we performed metaphase FISH analysis on a GBM cell line (GBM6 kindly provided by Prof. Paul S. Mischel) characterised by strong amplification of the *EGFR* gene. This cell line is of particular interest as amplified *EGFR* copies harbour the in-frame deletion of exons 2–7 coding for the extracellular ligand (EGFRvIII)<sup>9</sup>.

FISH was performed with probes targeting *EGFR* and the centromere of chromosome 7. *EGFR* was present in multiple

Table 1   Mutations lost in the spheres and focal amplifications.						
ID	Gene (mut)	Spheres (% reads)	Tumour (% reads)	Overlap to amplification in tumour	Amplification in spheres	Passages of spheres
GBM IV-34	<i>DDIT3</i> (p.P11S)	0	99	yes	No	6
GBM IV-34	<i>MARS</i> (p.G888E)	0	97	yes	No	6
GBM IV-34	<i>DDIT3</i> (p.S31L)	0	100	yes	No	6
GBM IV-34	<i>PDGFRA</i> (p.N659K)	0	86	yes	No	6
SK01600*	<i>EGFR</i> (p.C326S)	0	36	yes	No	0
GBM IV-19	<i>EGFR</i> (p.A289V)	57 (0†)	77	yes	Yes (No†)	0 (3†)
GBM IV-39	<i>EGFR</i> (p.S227Y)	88 (0†)	98	yes	Yes (No†)	0 (4†)

\*The sample was reanalysed from Yost et al.<sup>40</sup>.  
†Later passages have lost the *EGFR* mutations and amplifications as revealed through Sanger sequencing and FISH.



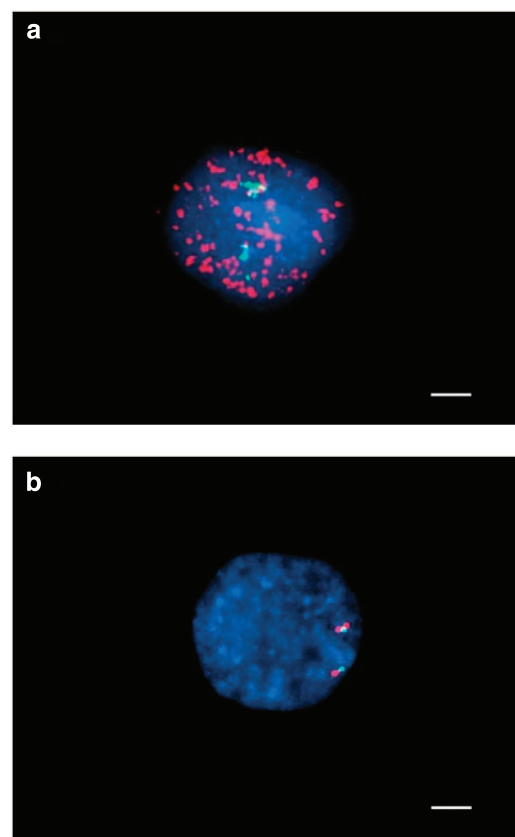
**Figure 1 | Two examples of Focal Amplifications in primary GBM IV-34 in the tumour tissues which are lost in the gliomaspheres.** Y axis- normalised log2 ratios of the sequence coverages between the tumour and the normal samples. X axis—equidistantly plotted exons. Green line—diploid state in the tumour. Blue vertical lines depict positions of the mutations. Crosses represent the loss of mutations in gliomaspheres. Red horizontal lines represent hidden Markov models prediction of the regions of amplifications. Focal Amplifications are estimated taking into account the fraction of tumour cells in the tumour samples.

extrachromosomal copies >100. After culturing with erlotinib in the media, we have repeated the FISH with the same conditions and, in agreement with our previous observations, all extrachromosomal copies have disappeared. Only the chromosomal *EGFR* was detectable.

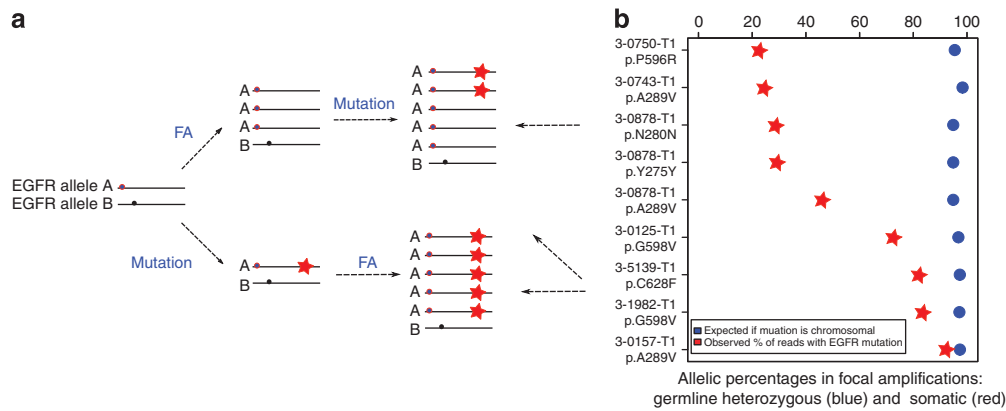
We have also analysed the data of one patient with GBM reported in the literature<sup>40</sup>, where both the primary tumour tissue and spheres were sequenced. The *EGFR* p.C326S (COSM1600351) mutation was within the focally amplified region and was identified in 36% of reads of the primary tumour but was lost in the spheres.

**Mutations of extrachromosomal origin.** These results raise the question of the origin of the mutations associated with FAs that are present in the primary GBMs and disappear in neurospheres. If mutations associated to FAs are of chromosomal origin and therefore cannot be lost without causing an LOH then their absence in the spheres could be explained by the expansion of a different clone without such mutations (Fig. 3a lower panel). On the other hand if mutations occur after the formation of DMs and therefore are of extrachromosomal origin, their absence in the spheres could be explained by the loss of DMs from the tumour cells (Fig. 3a upper panel). In this case, in one cell wild-type DM copies must for some time coexist with DM mutated copies.

To validate this hypothesis we reanalysed the DNA sequencing data from GBM primary tumours from The Cancer Genome Atlas (TCGA) study<sup>10</sup> with FAs and point mutations in *EGFR*, focusing on samples where *EGFR* mutations were present in <90% of reads (Fig. 3b). The allelic ratio of the germline heterozygous variants in the tumours was used to estimate the extent of FAs. Cases with suspicion of more than one FA of the *EGFR* locus were excluded (Supplementary Fig. 2). In the seven remaining TCGA GBM tumours, we detected nine *EGFR* mutations showing a level of amplification of more than 36 copies per cell. The allelic ratio in these FAs was close to 1 indicating that almost all sequence reads (= >95%) originate



**Figure 2 | FISH analysis for the detection of *EGFR* amplification in GBM IV-19 cells.** (a) FISH in primary tumour cells demonstrates euploid chromosome 7 (green signal) and multiple copies of *EGFR* scattered all over the nucleus (red signal). (b) Cultured tumour cells shows euploid chromosome 7 in green and not amplified *EGFR* (red signal). Scale bar = 5 µm.



**Figure 3 | Somatic mutations in EGFR occur after focal amplifications in GBM.** (a) Models of extrachromosomal mutations (in double minutes) (upper panel) and chromosomal mutation followed by amplification (lower panel). (b) Allelic percentages of heterozygous germline variants and somatic mutations in EGFR focal amplifications. Heterozygous germline variants of allele A and B (blue and black circles); Somatic mutations (red stars). GBM tumours were reanalysed from TCGA consortium.

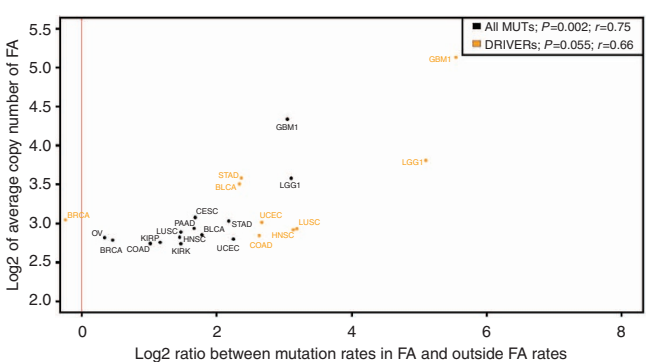
from the amplified copies. In addition, four out of nine somatic point mutations in *EGFR* were present in less than 50% of reads, indicating that a fraction of DM molecules did not contain the mutation (Fig. 3a upper panel). Therefore, these data confirm the existence of GBM tumours in which the mutated and wild-type DM copies coexist and therefore support the scenario, where the DMs are first formed and the mutation subsequently occurred in one of the DM copies. Thereafter both populations of DM molecules coexist until the DMs carrying the mutation are lost or fixed. We named this new class of mutations as ALEMs.

**Prevalence of ALEMs in different tumour types.** In order to investigate the prevalence of co-localisation of FAs and mutations in various tumour types, we investigated 4,198 tumours from 17 tumour types from the TCGA collection, for which both exome sequencing and CNV analyses were performed (Supplementary Table 1).

The FAs matching the characteristics of DMs and HSR<sup>31</sup> (more than four copies and length <6 Mb) were included in the study as described in the methods. In total we have identified 1,129 somatic mutations across all tumour types which map within regions of FAs and comprise 0.58% of all studied mutations. We found a positive correlation between mutation rates and extent of FAs across all tumour types ( $P = 0.002$ ,  $R = 0.75$ ). In each tumour type the mutation rates within FAs were higher than outside, with an average increase of  $3.67\text{-fold} \pm 2.68$ . The most important increase of mutation rates within FAs were observed in brain tumours: LGG (9-fold) and GBM (8-fold) (Fig. 4). These tumour types are characterised by frequent FAs in the form of DMs. Since DMs are isolated circular DNA molecules, we hypothesised that a competition between DM copies bearing different somatic mutations may result in positive selection for copies with the strongest oncogenic driver mutation.

To test this hypothesis, we generated a data set of mutations enriched in oncogenic driver mutations in 54 documented oncogenes<sup>6</sup>.

Remarkably 27% of all ‘oncogenic’ driver mutations were located within FAs. The probability of fixation of ‘oncogenic’ driver mutations in FAs as compared with the total number of mutations increased in almost all tumour types with the most pronounced effect in LGG (4-fold) and GBM (6-fold). A similar result was obtained in a different data set enriched in putative driver mutations, where only mutations with at least three occurrences in the COSMIC v67 database (Supplementary Fig. 3) were included. Interestingly, when a similar analysis was



**Figure 4 | Correlation of increase of mutation rates in FAs with the FA copy number.** Statistical significance was assessed with analysis of variance. X axis—log2 of the ratio between mutations rates inside FAs and outside. Y axis—log2 of the average copy number in FAs. Each data point represents the tumour type. Black—all mutations (N=14). Orange—mutations in oncogenes (N=9). Red line represents equal mutations rates inside and outside of FAs.

performed with only passenger mutations, many of these were also localised with FAs in all tumour types, however with a smaller enrichment compared with what was observed with putative driver mutations (Supplementary Fig. 3).

Next we investigated which genes exhibit non-random co-localisation of mutations with FAs. When all tumours were analysed, 212 genes revealed a significant enrichment of mutations in focally amplified regions ( $P < 0.05$ ). This list of genes included RTK such as *EGFR*, *PDGFRA*, *ERBB2* and *KIT*; other receptors associated to cancer such as *NOTCH3*, *EPHA6* and other oncogenes including *CCNE1*, *BCL11A*, *WHSC1L1* and *CDK8* (Supplementary Data 1).

We reasoned that genes mapping near known drivers found in FAs, may also display increased mutation rate. We observed this effect in *MED1* which is located 0.25 Mb from *ERBB2*; and in *SHANK2* and *PPFIA1* genes near *CCND1* (0.84 and 0.65 Mb, respectively) which is known to be amplified in a form of DMs<sup>27,39,41–45</sup>. These closely located genes on the chromosome are likely to be co-amplified in the same FA (Fig. 5).

To reveal tumour-specific oncogenes that exhibit the pattern of co-localisation of mutations within FAs, we repeated the gene-by-gene analyses independently per each tumour type. The strongest enrichment of mutations in FAs was observed in *EGFR* in GBM, low-grade glioma, head and neck squamous-cell carcinoma, lung

and uterine cancer. *PDGFRA* mutations were enriched in FAs in GBM, low-grade glioma and lung cancer; and *KIT* mutations in lung cancer. Similar co-localisations were observed for *NOTCH3* mutations in ovarian and breast cancers; *CCNE1* mutations in uterine cancer; *BCL11A* mutations in lung cancer and *WHSC1L1* mutations in head and neck squamous-cell carcinoma and lung cancer (Supplementary Fig. 4).

The fraction of putative driver mutations that occurred in FAs was increased for all tested oncogenes. The most remarkable effect was observed in *EGFR* where the percentage of mutations in FA in all tumours increased from 39% (all mutations) to 65% (putative driver mutations) followed by *PDGFRA* (from 17 to

50% in LGG) and *ERBB2* (from 7 to 11% in all tumours). These results taken together suggest a positive selection for DM clones carrying oncogenic driver mutations.

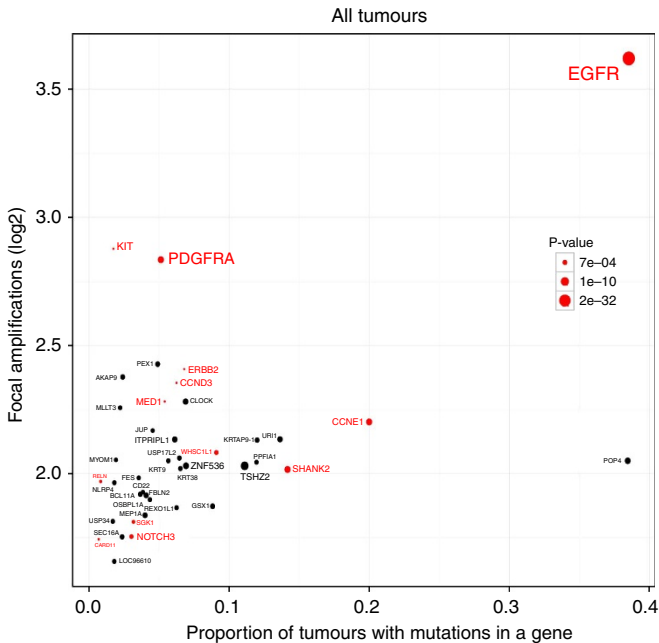
Discussion

In this work we demonstrate the existence of a non-random association of FAs and likely driver mutations in tumours. In addition, we propose a mechanism for acquisition of gain-of-function driver mutations in oncogenes mediated by the higher mutation load observed in DMs.

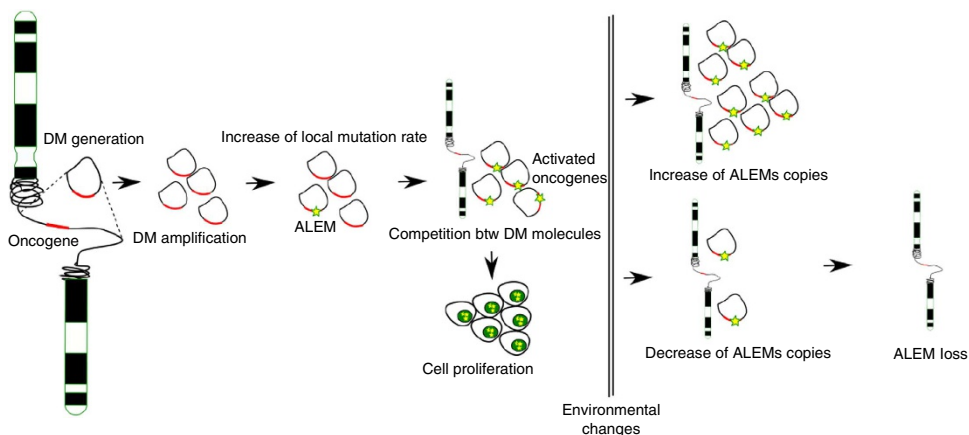
In several independent cases from this study and from Yost *et al.*<sup>40</sup>, both amplifications and mutations present in the primary GBM tumours were lost during cell culturing suggesting their extrachromosomal origin. DM origin of such mutations was confirmed by revealing the GBM tumours where wild type and mutated DM copies coexist (Fig. 3a upper panel).

We hypothesise that this phenomenon is taking place in several steps. The process begins with the generation of the DM molecules, which may happen in an almost random fashion across the genome. When the increased number of copies of a gene provides a proliferative advantage to the tumour cell, this event has a probability of being expanded in the tumour. The amplified DNA region is prone to acquisition of an increased number of variants because of a higher number of DNA copies (similar mutation rates with corresponding locus of genomic DNA) or a higher rates of acquisition of variants (higher mutation rates than in the corresponding locus of genomic DNA) or a combination of both. Subsequently, DMs with the oncogenic variants may be subjected to selection based on the random distribution of DMs among daughter cells. After cell division the cell with the highest number of DMs harbouring the driver mutation will have a proliferative advantage. The end point of this process is the presence of a high number of DMs per cell, where almost all copies have the driver ALEMs (Fig. 6).

An important consequence of this model is that, in the case of changes of environmental conditions, the number of DMs could be modulated and even reduced to zero resulting in the complete loss of ALEMs. The same mechanism would not be possible if these amplifications were in a form of HSR. Indeed the selection and competition between copies of amplified DNA with different genetic background is only possible between spatially isolated molecules, such as in the case of DMs. Moreover strong gain-of-function mutations in HSR amplifications would be detectable only if they have occurred at the early steps of amplification and were found in a high proportion of copies. Another consequence



**Figure 5 | Co-localisations of mutations and amplifications on gene-by-gene basis across 17 tumour types and *N*=4,198 tumour samples.** X axis—proportion of mutations in Focal Amplifications, Y axis—log2 of the average copy number in FAs. The area in circle is inversely proportional to the log2 of the log2 of the *P*-value (Fisher test). All oncogenes are selected in red. All oncogenes are presented if they have at least one mutation in FAs and a *P*-value less than 0.15, the other genes are presented if they have at least two mutations in FAs and *P*-value less than 0.01.



**Figure 6 | Model of generation and function of ALEMs.** After random generation of the DM molecules, the amplified DNA region is prone to acquisition of an ALEM due to a higher number of DNA copies. The cell with the highest number of DMs harbouring the ALEM will have a proliferative advantage. In response to environmental stress the cells may accordingly change the amount of DMs (see text for details).



of this model observed in this study is the enrichment of passenger mutations in FAs which can be explained by the 'hitchhiking' effect.

By studying a large number of tumours from publicly available data (TCGA consortium), we have detected co-localisation of mutations with amplifications in tumours known to harbour DMs such as GBM, LGG and LUSC<sup>13,24,27</sup>. Interestingly, genes highly affected by ALEMs were members of RTK family, such as *EGFR*, *PDGFRA* and *ERBB2*. We also noted that driver ALEMs were 26-fold more frequent in GBM and 13-fold in LGG than the passenger ALEMs. These two facts confirm the expected positive selection for the driver mutations in DMs.

Remarkably, our model explains the observations made by Nathanson *et al.*<sup>9</sup>, where the extrachromosomal EGFRvIII mutation disappeared in response to tyrosine kinase inhibitors.

ALEMs may make tumour cells fast-adaptable to the environmental changes including those induced by anticancer treatments. For example, we speculate that this mechanism may be utilised to acquire resistance to vemurafenib treatment in the BRAF V600E positive tumours<sup>46</sup>. According to our model amplification of EGFR which is not a strong oncogenic event *per se*<sup>47–50</sup> may increase the mutation load and enhance the probability of acquisition of the driver mutations in *EGFR*.

In conclusion, we provide evidence to support a novel type of cancer variants, the ALEMs. They result from a mutagenic process which is based on the increased mutational load of DMs that include proliferation-promoting genes such as tyrosine-kinases receptors leading to an increased adaptive potential of the tumour cells.

## Methods

**Processing of tumours and gliomasphere cultures.** In patient samples, tumour resections were obtained after surgery at the University Hospital of Geneva. After approval of the ethics committee of the Geneva University Hospitals informed written consent was obtained for all subjects. Primary tumour samples were cut in pieces and fresh frozen until analysis. Human biopsies were chopped mechanistically and digested with papain and DNase to generate a cell suspension. Gliomaspheres were thereof generated as previously described<sup>51</sup>. Briefly, media (DMEM-F12, B27 2%) and growth factors (EGF and bFGF at 10 ng ml<sup>-1</sup>) were renewed once every 5 days. Peripheral blood was obtained at the time of surgery. Peripheral blood mononuclear cells were isolated over a Ficoll gradient and frozen in liquid nitrogen in 10% DMSO until analysis.

**DNA extraction and exome sequencing.** The overall methodology was as previously described<sup>52–54</sup>. Briefly, DNA was extracted from the two distant fragments of frozen tissues, neurosphere cultures (spheres) and peripheral blood lymphocytes using the QIAamp DNA Mini Kit (Qiagen) for seven patients with GBM. When little material was available (<0.5 µg), Whole genome amplification was performed using REPLI-g Mini Kit (Qiagen). Exome capture was conducted using the SureSelect Human Exon v3 50 Mb (Agilent Technologies) reagents and sequencing was performed on Illumina HiSeq2000 instrument (Illumina) with paired-end 105 nt reads. Burrows–Wheeler Aligner (BWA) software<sup>55</sup> was used to align the sequence reads to the human reference genome (NCBI build GRCh37/hg19). SAMtools<sup>56</sup> was used to remove polymerase chain reaction (PCR) duplicates and to call single-nucleotide variants (SNV). Detection of small insertions and deletions (smINDEL) was conducted with Pindel 0.2.2 software<sup>57</sup>. The average sequencing coverage was 155 × per DNA sample (Supplementary Table 2). The search for somatic mutations was restricted to the regions covered at least 20-fold in both the normal and tumour samples.

**Calling of SNVs.** The initial list of SNVs was filtered against the common (>1%) germline polymorphisms present in the dbSNP137 and 1,000 genomes databases. SNVs present in the normal tissue sample from the same patient at a frequency of >1% were also filtered out. In contrast to the SNVs, smINDELs were called with lower accuracy and, therefore, we report only those smINDELs that were validated by Sanger sequencing. For both SNVs and smINDELs, we focused on the mutations that map to the protein coding sequences and to splice sites, as the untranslated exonic regions were less well covered by the commercially available exome capture reagents used in this study (Supplementary Data 2).

**Calling of LOHs and focal amplifications.** Two sources of information from exome sequencing were used to estimate the somatic copy number alterations:

(i) the fractions of reads with the heterozygous germline variants; and (ii) the ratios of the coverage of the tumour sample and the corresponding normal DNA. An in-house hidden Markov models based algorithm was used to predict the regions of FAs taking into account both sources of data. FAs were confirmed with quantitative PCR (Supplementary Fig. 5). Primers for quantitative PCR are reported in the Supplementary Table 3.

**FISH analysis.** Genes amplifications were investigated by FISH analysis using the LSI EGFR Spectrum Orange/CEP7 Spectrum Green probe (Vysis, Abbott Laboratories, IL, USA) and the BAC probe RP11-231c18 directed against *PDGFRA* (chr4:55,127,335–55,259,498; hg19, spectrum green) and the control probe dj963K6 (4qter, spectrum red).

The FISH signals for each locus-specific FISH probe were assessed under an Zeiss AxioScop microscope (Zeiss) equipped with specific filters (DAPI/Green/Red). DAPI II (4,6-diamino-2-phenylindole-2-hydrochloride) was used for chromatin counterstaining.

**TCGA data analysis.** The tumours for which exome-sequencing and CNV data were publicly available in TCGA<sup>10</sup> were analysed in this study (Supplementary Table 1). When several kinds of SNV analyses were available for the same tumours, we selected those which covered the largest number of samples. We selected amplifications of more than 4-fold and length less than 6 Mb<sup>31</sup>.

The Fisher exact test was applied for statistical assessment of non-random co-localisation of FAs and point mutations.

## References

- Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: the next generation. *Cell* **144**, 646–674 (2011).
- Luo, J., Solimini, N. L. & Elledge, S. J. Principles of cancer therapy: oncogene and non-oncogene addiction. *Cell* **136**, 823–837 (2009).
- Loeb, L. A., Springgate, C. F. & Battula, N. Errors in DNA replication as a basis of malignant changes. *Cancer Res.* **34**, 2311–2321 (1974).
- Loeb, L. A., Bielas, J. H. & Beckman, R. A. Cancers exhibit a mutator phenotype: clinical implications. *Cancer Res.* **68**, 3551–3557, discussion 3557 (2008).
- Loeb, L. A. Human cancers express mutator phenotypes: origin, consequences and targeting. *Nat. Rev. Cancer* **11**, 450–457 (2011).
- Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).
- Cancer Genome Atlas, N. Comprehensive molecular characterization of human colon and rectal cancer. *Nature* **487**, 330–337 (2012).
- Cancer Genome Atlas, N. Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
- Nathanson, D. A. *et al.* Targeted therapy resistance mediated by dynamic regulation of extrachromosomal mutant EGFR DNA. *Science* **343**, 72–76 (2014).
- Brennan, C. W. *et al.* The somatic genomic landscape of glioblastoma. *Cell* **155**, 462–477 (2013).
- Sanborn, J. Z. *et al.* Double minute chromosomes in glioblastoma multiforme are revealed by precise reconstruction of oncogenic amplicons. *Cancer Res.* **73**, 6036–6045 (2013).
- Snuderl, M. *et al.* Mosaic amplification of multiple receptor tyrosine kinase genes in glioblastoma. *Cancer Cell* **20**, 810–817 (2011).
- Cancer Genome Atlas, N. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. *Nature* **455**, 1061–1068 (2008).
- Ercan, D. *et al.* Amplification of EGFR T790M causes resistance to an irreversible EGFR inhibitor. *Oncogene* **29**, 2346–2356 (2010).
- Kong, A. *et al.* Rate of *de novo* mutations and the importance of father's age to disease risk. *Nature* **488**, 471–475 (2012).
- Nikolaev, S. I. *et al.* Life-history traits drive the evolutionary rates of mammalian coding and noncoding genomic elements. *Proc. Natl Acad. Sci. USA* **104**, 20443–20448 (2007).
- Carvalho, C. M. *et al.* Replicative mechanisms for CNV formation are error prone. *Nat. Genet.* **45**, 1319–1326 (2013).
- Deem, A. *et al.* Break-induced replication is highly inaccurate. *PLoS Biol.* **9**, e1000594 (2011).
- Malkova, A. & Haber, J. E. Mutations arising during repair of chromosome breaks. *Annu. Rev. Genet.* **46**, 455–473 (2012).
- Costantino, L. *et al.* Break-induced replication repair of damaged forks induces genomic duplications in human cells. *Science* **343**, 88–91 (2013).
- Storlazzi, C. T. *et al.* Gene amplification as double minutes or homogeneously staining regions in solid tumors: origin and structure. *Genome Res.* **20**, 1198–1206 (2010).
- Shimizu, N. Molecular mechanisms of the origin of micronuclei from extrachromosomal elements. *Mutagenesis* **26**, 119–123 (2011).
- Botchan, M. Hitchhiking without covalent integration. *Cell* **117**, 280–281 (2004).

24. Canute, G. W. *et al.* The hydroxyurea-induced loss of double-minute chromosomes containing amplified epidermal growth factor receptor genes reduces the tumorigenicity and growth of human glioblastoma multiforme. *Neurosurgery* **42**, 609–616 (1998).
25. Lundberg, G. *et al.* Binomial mitotic segregation of MYCN-carrying double minutes in neuroblastoma illustrates the role of randomness in oncogene amplification. *PLoS One* **3**, e3099 (2008).
26. Lalic, H., Volavsek, C. & Radošević-Stasić, B. Chromosomal instability and double minute chromosomes in a breast cancer patient. *Acta Med. Okayama* **58**, 51–58 (2004).
27. Nielsen, J. L. *et al.* Evidence of gene amplification in the form of double minute chromosomes is frequently observed in lung cancer. *Cancer Genet. Cytogenet.* **65**, 120–124 (1993).
28. Von Hoff, D. D. *et al.* Elimination of extrachromosomally amplified MYC genes from human tumor cells reduces their tumorigenicity. *Proc. Natl Acad. Sci. USA* **89**, 8165–8169 (1992).
29. Shimizu, N., Misaka, N. & Utani, K. Nonselective DNA damage induced by a replication inhibitor results in the selective elimination of extrachromosomal double minutes from human cancer cells. *Genes Chromosomes Cancer* **46**, 865–874 (2007).
30. Ambros, I. M. *et al.* Neuroblastoma cells can actively eliminate supernumerary MYCN gene copies by micronucleus formation—sign of tumour cell reversion? *Eur. J. Cancer* **33**, 2043–2049 (1997).
31. Vogt, N. *et al.* Molecular structure of double-minute chromosomes bearing amplified copies of the epidermal growth factor receptor gene in gliomas. *Proc. Natl Acad. Sci. USA* **101**, 11368–11373 (2004).
32. Lengauer, C., Kinzler, K. W. & Vogelstein, B. Genetic instabilities in human cancers. *Nature* **396**, 643–649 (1998).
33. Collis, S. J., DeWeese, T. L., Jeggo, P. A. & Parker, A. R. The life and death of DNA-PK. *Oncogene* **24**, 949–961 (2005).
34. Nik-Zainal, S. *et al.* Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979–993 (2012).
35. Yu, L. *et al.* Gemcitabine eliminates double minute chromosomes from human ovarian cancer cells. *PLoS One* **8**, e71988 (2013).
36. Quinn, L. A., Moore, G. E., Morgan, R. T. & Woods, L. K. Cell lines from human colon carcinoma with unusual cell products, double minutes, and homogeneously staining regions. *Cancer Res.* **39**, 4914–4924 (1979).
37. Bigner, S. H. & Vogelstein, B. Cytogenetics and molecular genetics of malignant gliomas and medulloblastoma. *Brain Pathol.* **1**, 12–18 (1990).
38. Schulte, A. *et al.* Glioblastoma stem-like cell lines with either maintenance or loss of high-level EGFR amplification, generated via modulation of ligand concentration. *Clin Cancer Res.* **18**, 1901–1913 (2012).
39. Szerlip, N. J. *et al.* Intratumoral heterogeneity of receptor tyrosine kinases EGFR and PDGFRA amplification in glioblastoma defines subpopulations with distinct growth factor response. *Proc. Natl Acad. Sci. USA* **109**, 3041–3046 (2012).
40. Yost, S. E. *et al.* High-resolution mutational profiling suggests the genetic validity of glioblastoma patient-derived pre-clinical models. *PLoS One* **8**, e56185 (2013).
41. Bigner, S. H., Mark, J. & Bigner, D. D. Cytogenetics of human brain tumors. *Cancer Genet. Cytogenet.* **47**, 141–154 (1990).
42. Fischer, U. *et al.* A different view on DNA amplifications indicates frequent, highly complex, and stable amplicons on 12q13–21 in glioma. *Mol. Cancer Res.* **6**, 576–584 (2008).
43. Phillips, J. J. *et al.* PDGFRA amplification is common in pediatric and adult high-grade astrocytomas and identifies a poor prognostic group in IDH1 mutant glioblastoma. *Brain Pathol.* **23**, 565–573 (2013).
44. Del Rey, J. *et al.* Centrosome clustering and cyclin D1 gene amplification in double minutes are common events in chromosomal unstable bladder tumors. *BMC Cancer* **10**, 280 (2010).
45. Gibaud, A. *et al.* Extrachromosomal amplification mechanisms in a glioma with amplified sequences from multiple chromosome loci. *Hum. Mol. Genet.* **19**, 1276–1285 (2010).
46. Su, F. *et al.* Resistance to selective BRAF inhibition can be mediated by modest upstream pathway activation. *Cancer Res.* **72**, 969–978 (2012).
47. Holland, E. C., Hively, W. P., DePinho, R. A. & Varmus, H. E. A constitutively active epidermal growth factor receptor cooperates with disruption of G1 cell-cycle arrest pathways to induce glioma-like lesions in mice. *Genes Dev.* **12**, 3675–3685 (1998).
48. Weiss, W. A. *et al.* Genetic determinants of malignancy in a mouse model for oligodendroglioma. *Cancer Res.* **63**, 1589–1595 (2003).
49. Wei, Q. *et al.* High-grade glioma formation results from postnatal pten loss or mutant epidermal growth factor receptor expression in a transgenic mouse glioma model. *Cancer Res.* **66**, 7429–7437 (2006).
50. Corcoran, R. B. *et al.* EGFR-mediated re-activation of MAPK signaling contributes to insensitivity of BRAF mutant colorectal cancers to RAF inhibition with vemurafenib. *Cancer Discov.* **2**, 227–235 (2012).
51. Galli, R. *et al.* Isolation and characterization of tumorigenic, stem-like neural precursors from human glioblastoma. *Cancer Res.* **64**, 7011–7021 (2004).
52. Nikolaev, S. I. *et al.* A single-nucleotide substitution mutator phenotype revealed by exome sequencing of human colon adenomas. *Cancer Res.* **72**, 6279–6289 (2012).
53. Nikolaev, S. I. *et al.* Exome sequencing identifies putative drivers of progression of transient myeloproliferative disorder to AMKL in infants with Down syndrome. *Blood* **122**, 554–561 (2013).
54. Nikolaev, S. I. *et al.* Frequent cases of RAS-mutated Down syndrome acute lymphoblastic leukaemia lack JAK2 mutations. *Nat. Commun.* **5**, 4654 (2014).
55. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
56. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
57. Ye, K., Schulz, M. H., Long, Q., Apweiler, R. & Ning, Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. *Bioinformatics* **25**, 2865–2871 (2009).

## Acknowledgements

This study was supported by grants from the Swiss Cancer League (LSCC 2939-02-2012) and SER-Swiss Russia to S.I.N.; SLCC2939-02-2012, Novartis14B065, 'Dinu Lipatti' to S.I.N.; ERA-NET to SEA and S.I.N.; SNF 144082, ERC 249968, Foundation 'ChildCare' to SEA and Foundation 'Damm-Etienne' to I.R.; 'Fondation Artères, Genève' and 'Ligue genevoise contre le cancer' to P.Y.D. The authors thank Prof. Paul Mischel for kindly providing the GBM6 cell line and Prof. Thanos Halazonetis for constructive discussion and comments on the manuscript.

## Author contributions

S.I.N., F.A.S., V.C.S., I.R., K.S. and S.E.A. designed the study; S.I.N., F.A.S., M.G. and P.M. performed statistical analysis; K.S., P.Y.D., V.C.S., I.R. and V.D. collected the samples and collated clinical data; E.F., A.V., F.B., F.F., V.D. and M.Gu. performed the experiments; and S.I.N., F.A.S., P.M., S.E.A. analysed the results and wrote the manuscript.

## Additional information

**Accession codes:** Sequence data for exome-sequencing of primary tumour fragments, matched blood samples and derived gliomaspheres have been deposited in GenBank/EMBL/DBJ under the accession code PRJNA263837.

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npng.nature.com/reprintsandpermissions/>

**How to cite this article:** Nikolaev, S. *et al.* Extrachromosomal driver mutations in glioblastoma and low grade glioma. *Nat. Commun.* 5:5690 doi: 10.1038/ncomms6690 (2014).