Intrachromosomal tandem duplication and repeat expansion during attempts to inactivate the subtelomeric essential gene *GSH1* in *Leishmania*

Angana Mukherjee[1], Lance D. Langston[2] and Marc Ouellette[1,*]

[1]Centre de Recherche en Infectiologie and Département de Microbiologie, Immunologie and Infectiologie, Université Laval, Québec, Canada, G1V 4G2 and [2]Laboratory of DNA Replication, Howard Hughes Medical Institute, The Rockefeller University, New York, NY, 10065, USA

ABSTRACT

Gamma-glutamylcysteine synthetase encoded by *GSH1* is the rate-limiting enzyme in the biosynthesis of glutathione and trypanothione in *Leishmania*. Attempts to generate *GSH1* null mutants by gene disruption failed in *Leishmania infantum*. Removal of even a single allele invariably led to the generation of an extra copy of *GSH1*, maintaining two intact wild-type alleles. In the second and even third round of inactivation, the markers integrated at the homologous locus but always preserved two intact copies of *GSH1*. We probed into the mechanism of *GSH1* duplication. *GSH1* is subtelomeric on chromosome 18 and Southern blot analysis indicated that a 10-kb fragment flanked by 466-bp direct repeated sequences was duplicated in tandem on the same chromosomal allele each time *GSH1* was targeted. Polymerase chain reaction analysis and sequencing confirmed the generation of novel junctions created at the level of the 466-bp repeats consequent to locus duplication. In loss of heterozygosity attempts, the same repeated sequences were utilized for generating extra-chromosomal circular amplicons. Our results are consistent with break-induced replication as a mechanism for the generation of this regional poly-ploidy to compensate for the inactivation of an essential gene. This chromosomal repeat expansion through repeated sequences could be implicated in locus duplication in *Leishmania*.

INTRODUCTION

The protozoan parasite *Leishmania* is the causative agent for leishmaniasis, and it belongs to the Kinetoplastida, one of the oldest eukaryotic lineages. *Leishmania* has a plastic genome and this was first illustrated by diverse karyotypes when different species of *Leishmania* were compared (1). *Leishmania* is considered as a diploid organism but several studies have shown that a portion of its genome can become aneuploid (2–4). Recently, fluorescence *in situ* hybridization studies have shown that aneuploidy appears to be common with variable chromosomal ploidy among individual cells in *Leishmania* (5). Gene amplification as a part of linear minichromosomes or extrachromosomal circles is also frequent in *Leishmania* and is a manifestation of genome plasticity (6–8). This kind of gene amplification can be found in unselected stocks or after drug selection (6,9). These gene rearrangements usually occur at the level of direct or inverted repeats (3,4,10,11). Another manifestation of genome plasticity is in the attempts to generate null mutant of reputed essential genes. Indeed, upon inactivation, gene rearrangement takes place maintaining at least one intact allele of the gene to be targeted. This has been reported abundantly (2,12,13), although the mechanism to generate this ploidy has not been studied into details, one study highlighted a change in chromosomal ploidy (2) and in another there was a translocation of a chromosomal segment to another chromosome (12).

Recently, we have attempted to generate a null mutant of the gene *GSH1* encoding gamma-glutamylcysteine synthetase (γ-GCS), the rate-limiting enzyme in the biosynthesis of glutathione in *Leishmania*. We reported that *GSH1* is an essential gene in *Leishmania* (14). Indeed, all

our attempts to generate chromosomal null mutants failed except if cells were first provided with a rescue *GSH1* plasmid encoding γ-GCS. While generation of polyploidy is frequently observed when attempting disruption of essential genes in *Leishmania* (2,12), in the case of *GSH1*, the deletion of one single allele also led to gene rearrangements. In this current study, we provide evidence for a mechanism of intrachromosomal gene duplication where the *GSH1* locus becomes polyploid preserving two intact copies of the essential gene upon sequential rounds of gene inactivation.

MATERIALS AND METHODS

Strains and culture conditions

Promastigotes of *Leishmania infantum* (MHOM/MA/67/ ITMAP-263) and its transfectants were grown in SDM-79 medium supplemented with 10% fetal bovine serum, 5 μg/ml of hemin at pH 7.0 and 25°C.

DNA constructs and transfection

GSH1 inactivation cassettes with hygromycin phosphotransferase B, neomycin phosphotransferase, blasticidin deaminase for *L. infantum* were constructed using a polymerase chain reaction (PCR) fusion-based strategy and transfected as described previously (14). The *LinJ_V3.1670* inactivation cassettes were also generated by a similar PCR fusion-based strategy with hygromycin and neomycin phosphotransferases as the selectable markers. Two micrograms of linear fragments for transfection were obtained by PCR amplification, gel purified and transfected into promastigotes by electroporation (15). Recombinants were preselected initially in the presence of 200 μg/ml hygromycin, 20 μg/ml G418 (Geneticin, Gibco-BRL) or 50 μg/ml blasticidin S (Invitrogen). After 24 h, the transfected cells were grown in the presence of higher drug concentrations and cells growing in highest drug selection (600 μg/ml hygromycin, 80 μg/ml G418; 100 μg/ml blasticidin) were cloned.

Pulsed-field gel electrophoresis

Intact chromosomes were prepared from *Leishmania* promastigotes harvested from late log phase, washed and lysed *in situ* in 1% low melting agarose plugs. Briefly, cells were resuspended in HEPES-NaCl buffer at a density of 5×10^8 cells/ml and mixed with equal volume of low melting-point agarose. Cells were lysed in the presence of 0.5 M ethylenediaminetetraacetic acid (EDTA; pH 9.5), 1% sodium dodecyl sulfate (SDS) and 350 μg/ml proteinase K overnight at 50°C. For HpaI digestion of chromosomes embedded in agarose blocks, the blocks were washed twice in 25 vol of TE to remove proteinase K. The blocks were then equilibrated in restriction buffer at room temperature for 30 min after which fresh buffer was added and 100 U of HpaI in a reaction volume of 200 μl were incubated at 37°C overnight. *Leishmania* chromosomes were separated by Pulsed-field gel electrophoresis (PFGE) using a Bio-Rad CHEF-DR III apparatus at 5 V/cm, 120° separation angle as described previously (12). The range of chromosome separation was between 100 and 900 kb and between 500 and 1000 kb depending on the conditions used. Digested chromosomes were separated under conditions optimal for separating between 25 and 100 kb for 27 h.

Southern blot analysis

Genomic DNA of the clones was isolated using DNAzol (Invitrogen) and circular DNA was isolated by Promega Wizard miniprep kit following the manufacturer's instructions. Digested genomic DNA or circular DNA or separated chromosomes were subjected to Southern blot hybridization with [α-^{32}P]dCTP-labeled DNA according to standard protocols (16). All probes were obtained by PCR from *Leishmania* genomic DNA. Densitometric analyses of Southern blots were performed using Image J and Agfa Arcus 2 scanner.

RESULTS

GSH1 polyploidy following gene disruption attempts

The *GSH1* gene is single copy (17) and essential in *Leishmania* (14). Removal of even a single wild-type (WT) allelic copy of *GSH1* invariably led to the generation of an extra copy of *GSH1*, maintaining two intact WT alleles (14). As demonstrated previously (14) successive integration of the selectable markers hygromycin phosphotransferase (*HYG*), neomycin phosphotransferase (*NEO*) and blasticidin deaminase (*BLA*) in the *GSH1* gene of *L. infantum* (Figure 1A) (14) led not only to the integration of all the three markers in the *GSH1* locus but also to preservation of intact *GSH1* alleles (14). This was further confirmed with additional restriction digestions and Southern blots (Figure 1). A 3.0-kb band was observed in WT cells when hybridized to a 550-bp probe within the 5′ flank of *GSH1* (Figure 1B). Upon sequential integration of the *HYG, NEO* or *BLA* resistance cassettes, digestion with HindIII and probing with the 5′ flank *GSH1* probe of 550 bp yielded 5.2, 5.0 and 4.7 kb hybridizing bands, respectively, in addition to the WT band of 3.0 kb (Figure 1B). Densitometric analyses were done with Image J software and values were represented as Net Integrated Optical density after background subtraction which revealed a 2:1 ratio for WT alleles compared to the *HYG, NEO* or *BLA* alleles (Figure 1B).

In an attempt to further characterize the events leading to polyploidy at the *GSH1* locus, we separated the chromosomes of the WT and recombinants by PFGE. Ethidium bromide staining of the gel did not reveal the presence of linear amplicons or gross karyotypic changes in the recombinants due to *GSH1* rearrangement (Figure 2A) and hybridizing to a *GSH1* probe indicated that the gene was present on the 720-kb chromosome for all recombinants with no evidence for circular amplification or gross translocation elsewhere on the genome (Figure 2A). Integration of the *HYG*, *NEO* and *BLA* markers at the *GSH1* locus was further confirmed by hybridization with no evidence for gross gene rearrangements (Figure 2B–D). Chromosome 18 is disomic in the *L. infantum* strain studied. Indeed, we could easily obtain

A

B
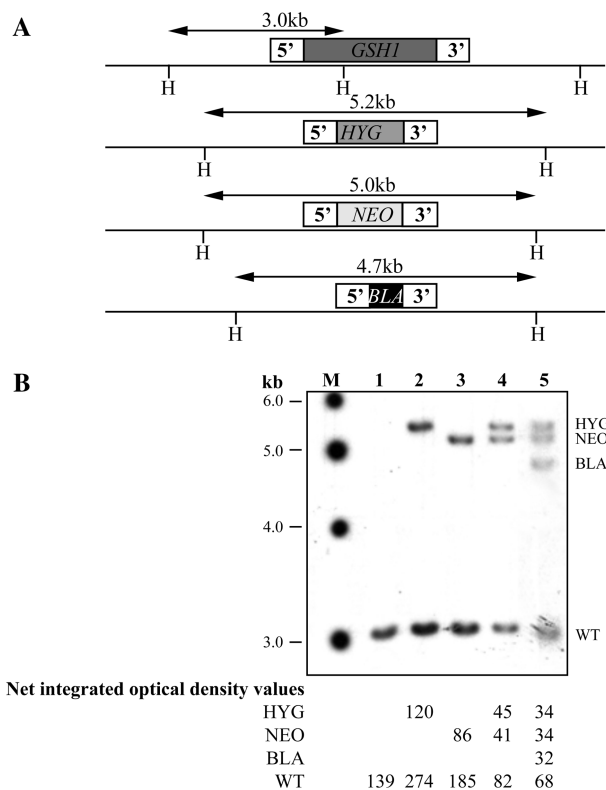
Figure 1. *GSH1* gene inactivation in *L. infantum*. (**A**) Schematic drawing of the *GSH1* locus in *L. infantum* before and after integration of the inactivation cassettes (hygromyin phosphotransferase B, *HYG*; neomycin phosphotransferase, *NEO*; blasticidin deaminase, *BLA*) and the relevant HindIII restriction sites (H). (**B**) Southern blot analysis with genomic DNA digested with HindIII from the WT and recombinant clones and hybridization with a probe covering the 5′ flank *GSH1* region. Molecular weights (M) are indicated on the left, the various alleles are pinpointed on the right and underneath is the raw data (Net integrated optical density values) of densitometric analysis of each allele as determined using the Image J software. 1, *L. infantum* WT, *GSH1/GSH1*; 2–5, *L. infantum* with the following genotypes at the *GSH1* locus: 2, *GSH1/GSH1/HYG*; 3, *GSH1/GSH1/NEO*; 4, *GSH1/GSH1/NEO/HYG*; 5, *GSH1/GSH1/NEO/HYG/BLA*.

a *GSH1* chromosomal null mutant in two rounds of transfection provided that a rescue plasmid was present (14). Additionally, we obtained a chromosomal null mutant of *LinJ18_V3.1670*, the gene downstream of *GSH1* (Figure 3), by two successive integration of *NEO* and *HYG* (Supplementary Figure S1).

Gene rearrangements at the level of direct repeated sequences

The generation of *GSH1* polyploidy is neither due to translocation of a segment elsewhere on the genome nor by chromosomal aneuploidy. *L. infantum* promastigotes with the *GSH1/GSH1/NEO/HYG* genotype were selected for high concentration of G418 and hygromycin B in order to try to induce loss of heterozygosity (18) but this failed and instead extrachromosomal circular elements with the *NEO* and *HYG* markers were generated (14). Extrachromosomal circles are often formed by homologous recombination (HR) between direct repeated

sequences (7,10,11) and this prompted us to search for such repeated sequences in the vicinity of the *GSH1* gene. We discovered direct repeats of 466 bp with 98.9% identity flanking a region of 10.136 kb and encompassing a region with four genes (Figure 3A). Interestingly, the last 100 bp of the 3′ end of the *GSH1* inactivation cassette construct were part of the 466-bp repeated sequences. The extrachromosomal circular elements derived from these cells were isolated (Figure 3B, extreme left), digested with NcoI and hybridized to the *HYG* and *NEO* probes, and after longer migration to the 5′ flank *GSH1* probe. Both the *HYG* and *NEO* alleles were amplified as circles giving rise to the expected restriction fragments after NcoI digestion (Figure 3B, middle panels). To ascertain that the extrachromosomal circular amplicon was generated by HR between the 466-bp direct repeated sequences, a pair of primers (1a and 1b) was designed to amplify the new junction of the circle generated by recombination from the DNA isolated by alkaline lysis (Figure 3A and C). PCR amplification with this set of primers indeed amplified the expected 850-bp product from the DNA isolated from *GSH1* recombinant cells subjected to high concentration of the drugs and not from WT under the conditions tested (Figure 3D). Sequencing of this amplified fragment confirmed that the circles were formed by HR between the 466-bp repeats (data not shown).

We next investigated whether the direct 466-bp repeated sequences also played a role in locus duplication following targeting of *GSH1*. The copy number of neighboring genes of *GSH1* on chromosome 18 was investigated by digesting genomic DNA from the WT and recombinants with NcoI and by Southern blot hybridization using nearby genes as probes (Figure 4A). Restriction fragments hybridizing to the probes used are shown in Figure 4A, the superscripts indicating the panel number in Figure 4B and C. We used probes recognizing genes upstream, within and downstream of the 466-bp repeated sequences. *LinJ23_V3.0310* (*PTR1*) was used as a control for normalization (Figure 4B; panel, control). The copy number of genes *LinJ18_V3.1600* to *LinJ18_V3.1620* (upstream the direct repeated sequence) remained diploid (Figure 4B, panels a–c) and similarly the copy number of the genes *LinJ18_V3.1670* and *LinJ18_V3.1680*, downstream the repeated sequences were unchanged (Figure 4B, panels d and e). However, when the blots were probed with *LinJ18_V3.1630* and *LinJ18_V3.1640*, we observed an increase in copy number of the gene by one copy in the mutants every time *GSH1* was targeted in the parental strain. For example in *GSH1/GSH1/HYG* (Figure 4C, panels i and ii lanes 2) and *GSH1/GSH1/NEO* (Figure 4C, panels i and ii lanes 3), densitometric analyses revealed that the intensity of the bands were approximately higher by 1.5 times than WT. Interestingly, the band intensities of *LinJ18_V3.1630* and *LinJ18_V3.1640* were ~2 and 2.5 times higher than WT in *GSH1/GSH1/NEO/HYG* and *GSH1/GSH1/NEO/HYG/BLA* mutants respectively (Figure 4C, panels i and ii lanes 4 and 5). These blots (Figure 4C, panels i and ii) were normalized with *LinJ18_V3.1620* (Figure 4C, panel control), a gene on chromosome 18 which did not change
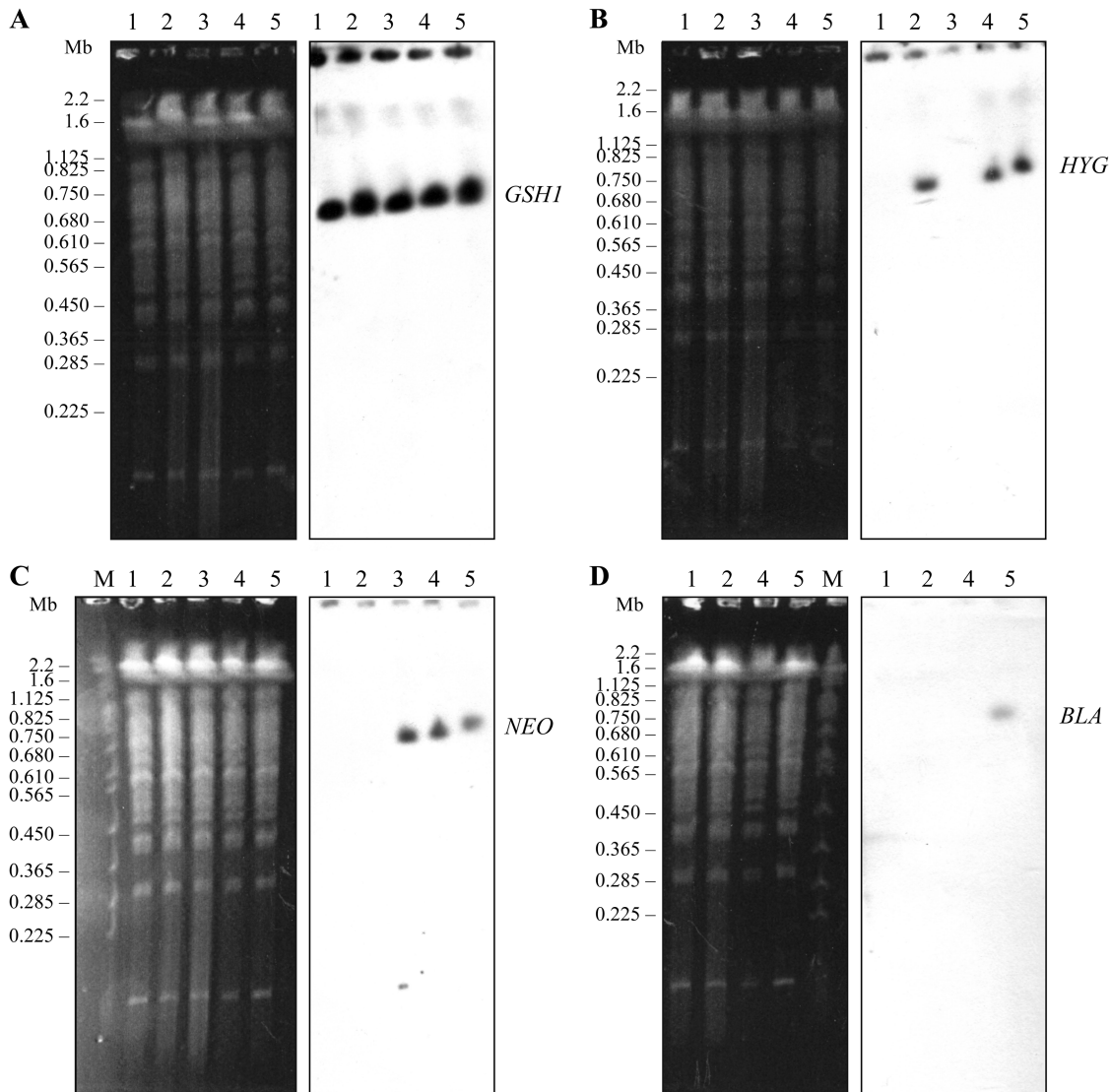
Figure 2. *GSH1* polyploidy is not due to gross translocation, circular/linear amplification or chromosomal aneuploidy. Pulsed-field gel electrophoresis was used for separating chromosomes between 100 kb and 900 kb and the gels were visualized by ethidium bromide staining and transferred and hybridized to a *GSH1* ORF (**A**), *HYG* (**B**), *NEO* (**C**), *BLA* (**D**) probes. Molecular weight markers (0.225–2.2 Mb *S. cerevisiae* chromosomes) are indicated on the left. 1, WT *L. infantum, GSH1/GSH1*; 2, *GSH1/ GSH1/ HYG*; 3, *GSH1/GSH1/NEO*; 4, *GSH1/GSH1/NEO/HYG*; 5, *GSH1/GSH1/NEO/HYG/BLA*.

its copy number when *GSH1* was targeted. Probing with *LinJ18_V3.1650* yielded multiple bands because of the nearby integration of the resistance markers *NEO, HYG* and *BLA* (Figure 4C, panel iii). Densitometric analyses revealed that all mutants retained two WT alleles in addition to the integration of a copy of the markers (Figure 4C, panel iii and data not shown). When the blot was probed with *LinJ18_V3.1660* (*GSH1*) we observed an identical copy number of the gene in WT and all recombinants (Figure 4C, panel iv). Digesting genomic DNA of these recombinants with ClaI and probing with the same genes between or outside the repeats essentially yielded similar results (data not shown). The results suggested that the region between the 466-bp repeated sequences was duplicated after each attempt of inactivating *GSH1*.

Chromosomes of WT parasites and recombinants were separated by PFGE and Southern blot analysis was performed with probes covering *LinJ18_V3.1620* to *LinJ18_V3.1670* (Figure 5A) and probes specific for each resistance marker (Figure 5B). The targeting of either *HYG* or *NEO* marker in WT indicated that it integrated into chromosome 18 (Figure 5B, lanes 2 and 3) and that chromosome 18 remained intact but with a slight increase in intensity as determined by hybridization with four probes within the repeats (Figure 5A, lanes 2 and 3). Targeting of the *HYG* marker in the *GSH1* locus in the *GSH1/GSH1/NEO* parasites led also to the integration of the *HYG* cassette in chromosome 18 but we observed a further increase in the intensity in the hybridization signal with the four probes *LinJ18_V3.1630* to
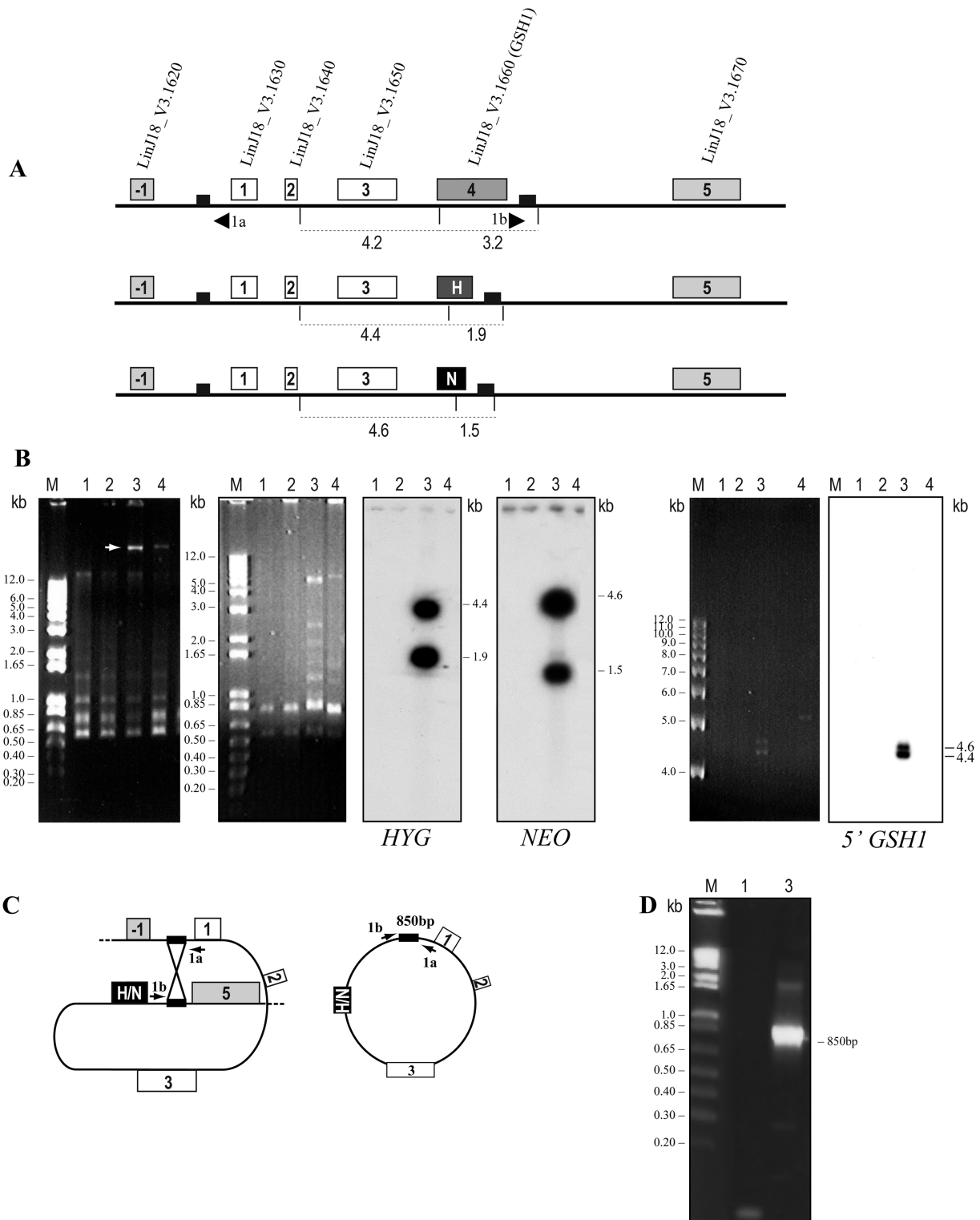
Figure 3. Formation of extrachromosomal circles by homologous recombination between direct repeated sequences. (**A**) Schematic drawing of the *GSH1* locus in *L. infantum,* before and after integration of the *HYG* and *NEO* cassettes with the direct repeated sequences of 466 bp as small black boxes. Genes within the repeated sequences are represented as 1, 2, 3 and 4 where 4 is *LinJ18_V3.1660* (*GSH1*). The gene upstream the repeat is represented as −1 (*LinJ18_V3.1620*) and the gene downstream the repeat is represented as 5 (*LinJ18_V3.1670*). The sizes of the NcoI digests are shown as dotted lines. H, hygromycin phosphotransferase B; N, neomycin phosphotransferase. (**B**) Isolation of circular DNA amplicons using the Promega Wizard Plus kit. The white arrow represents the presence of extrachromosomal circular DNA (extreme left, lanes 3 and 4). The plasmid was digested with NcoI (second gel from the left), followed by Southern blot hybridization with *HYG* and *NEO* probes. The NcoI digested DNA was also run longer and hybridized with the 5′ flank *GSH1* (right panel) (**C**) Model for the generation of the extrachromosomal circular amplicon, formed by HR between the 466-bp direct repeated sequences. (**D**) PCR amplification with primers 1a and 1b as shown in (A) and (C) were used to amplify the 850-bp fragment from circular DNA isolated by alkaline lysis. 1, WT (*GSH1/GSH1*); 2, *GSH1/GSH1/NEO/HYG*; 3, *GSH1/GSH1/NEO/HYG* grown in the presence of 10 X EC$_{50}$ of the selective drugs; 4, WT transfected with an episomal vector (PspαZeoα-GSH1) as positive control for plasmid isolation.
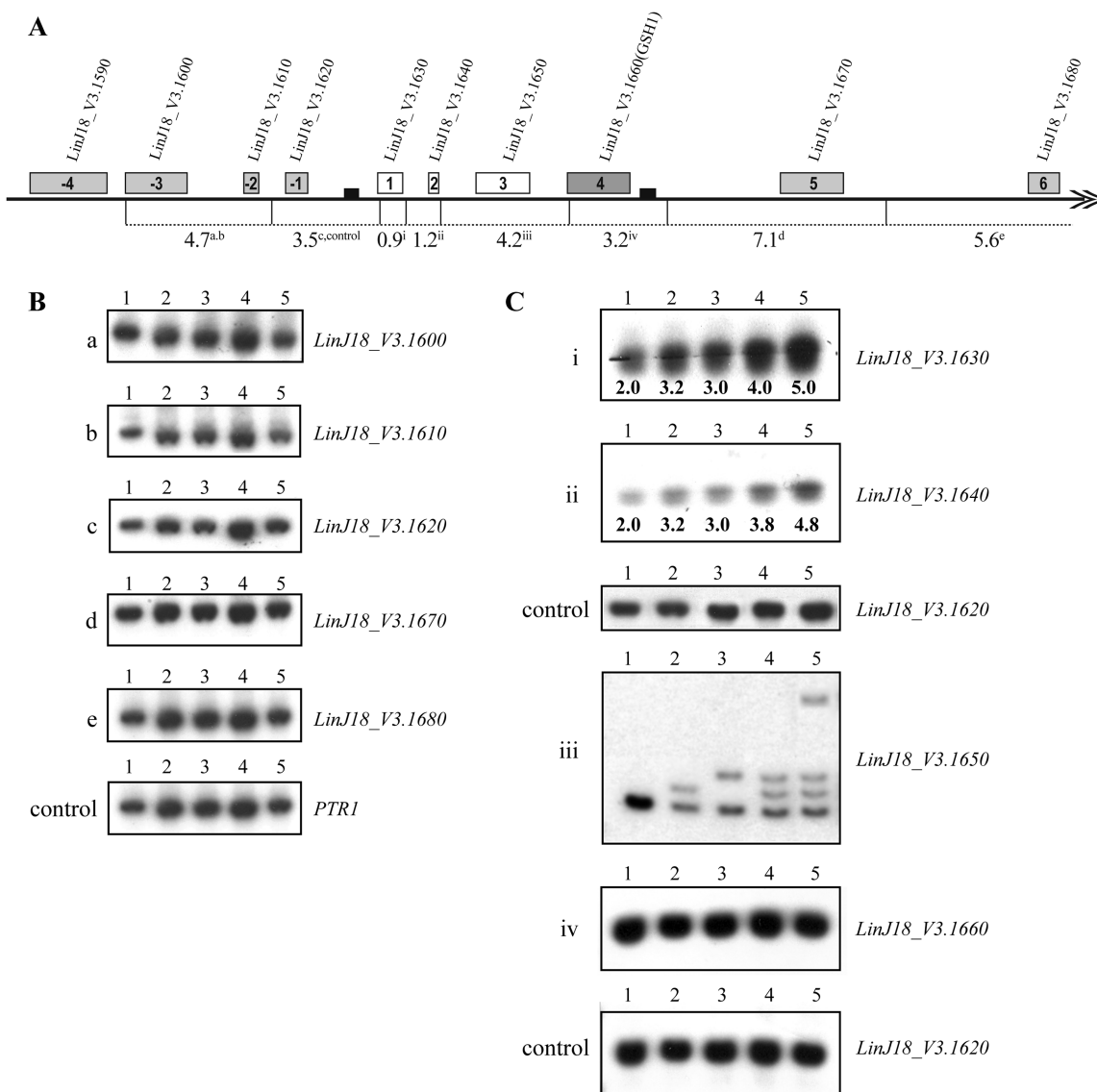
Figure 4. Chromosomal rearrangement following inactivation of the *GSH1* gene. (**A**) Schematic drawing of the *GSH1* locus on chromosome 18 in *L. infantum*, showing the NcoI restriction pattern. Direct repeated sequences of 466 bp are indicated as small black boxes. Genes within the repeated sequences are represented as 1, 2, 3 and 4 where 4 is *LinJ18_V3.1660* (*GSH1*). Genes upstream to the repeats are represented with a negative sign and genes downstream the repeats are 5 and 6. The telomere is located after 6 (*LinJ18_V3.1680*) and shown as a double arrowhead. NcoI restriction fragments' sizes are shown by dotted lines and the superscripts on the restriction fragments denote the panel numbers in (**B**) and (**C**). (B) Southern blot hybridizations of digested genomic DNA with NcoI using specific probes from the genes located upstream (a–c) and downstream (d and e) of the direct repeats. *LinJ23_V3.0310* (*PTR1*) was used as a control for DNA. (C) Southern blot hybridizations of digested genomic DNA with NcoI using specific probes from the genes between the repeated sequences (i, ii, iii and iv). Relative copy numbers of the genes in different recombinants compared to WT are indicated below the blots after densitometric analysis (panels i and ii). *LinJ18_V3.1620* was used as a control for DNA loading for panels i, ii and iv. 1, WT *L. infantum, GSH1/GSH1*; 2, *GSH1/GSH1/HYG*; 3, *GSH1/GSH1/NEO*; 4, *GSH1/GSH1/NEO/HYG*; 5, *GSH1/GSH1/NEO/HYG/BLA*.

3′ *LinJ18_V3.1660* (550 bp) that are within the direct repeated sequences but not with the genes *LinJ18_V3.1620* and *LinJ18_V3.1670* just outside the repeats (Figures 4A and 5A, lanes 4). However, with all these probes we saw a small size upper shift in the hybridizing band due to an increase in size of the chromosome. This shift was also observed with the *HYG* and *NEO* hybridization (Figure 5B, lanes 4). This was validated further with the integration of the *BLA* marker in the *GSH1/GSH1/NEO/HYG* parasites. In this case, two

homologs of chromosome 18 were clearly observed as discrete bands, the band with the higher molecular weight hybridizing with the *HYG*, *NEO* and *BLA* probes (Figure 5B, lanes 5) and the four genes within the repeats clearly giving rise to a stronger hybridization signal on the upper band compared to the two genes outside the repeats (*LinJ18_V3.1620* and *LinJ18_V3.1670*) (Figure 5A, lanes 5). The data are compatible with intrachromosomal duplication of the *GSH1* locus between the 466-bp repeats, after each inactivation attempt (Figure 6). To prove this,
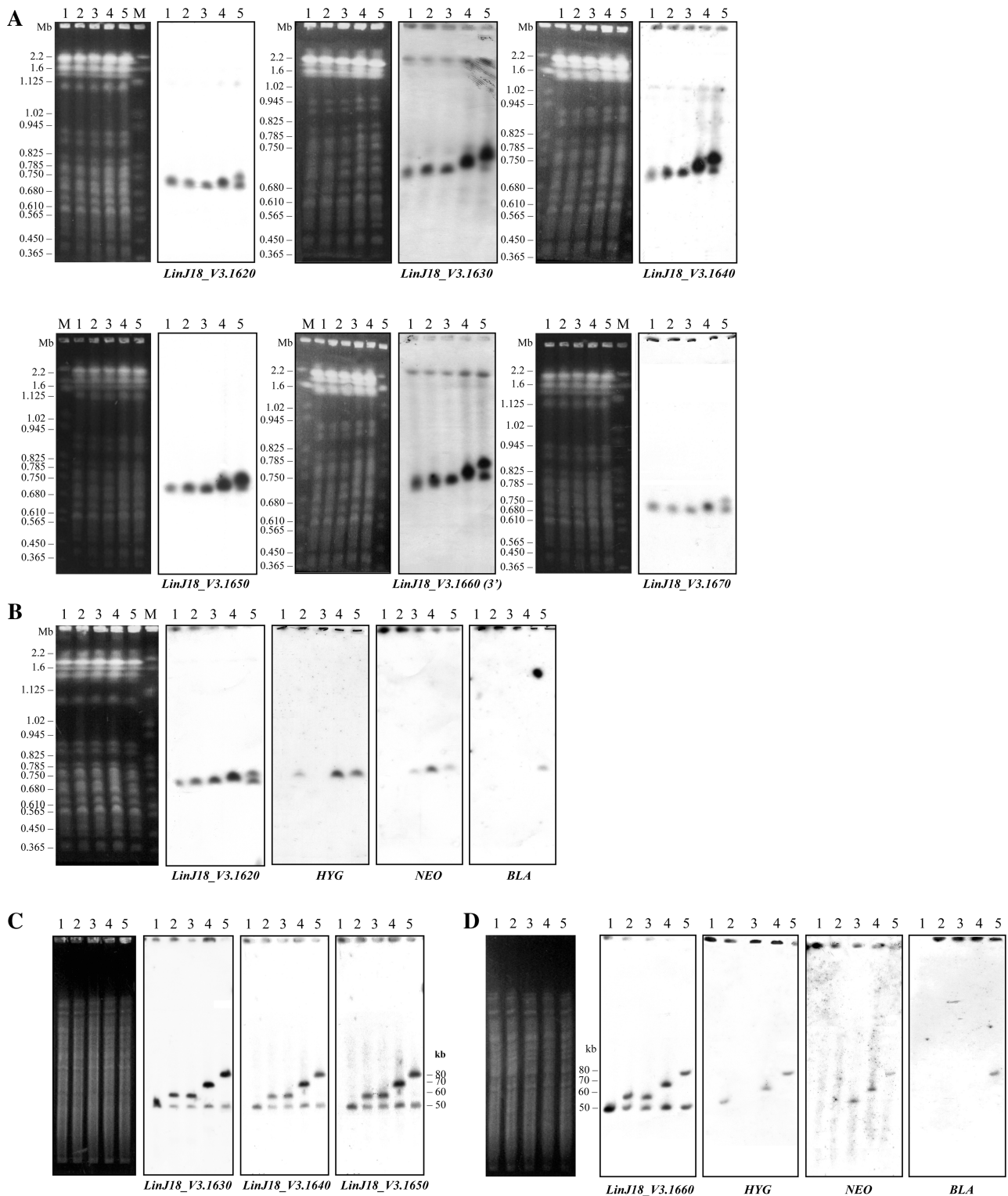
Figure 5. Analysis of chromosomal sized DNA upon inactivation of *GSH1*. (**A** and **B**) Chromosomes from the WT and recombinants were separated by PFGE using a separation range of 500 and 1000 kb and hybridized with specific probes. (A) Probes recognizing genes outside the repeated sequences, *LinJ18_V3.1620* and *LinJ18_V3.1670* and genes within the repeated sequences (*LinJ18_V3.1630* to *LinJ18_V3.1660*) were used. (B) Probes specific for *LinJ18_V3.1620*, *HYG*, *NEO* and *BLA* were used to hybridize the separated chromosomes. (**C** and **D**) Agarose blocks containing total DNA was digested with HpaI and DNA was separated by PFGE using a separation range of 25–100 kb and hybridized with probes recognizing genes within the repeated sequences (*LinJ18_V3.1630* to *LinJ18_V3.1650*) (C) or *GSH1*, *HYG*, *NEO* and *BLA* (D). 1, WT *L. infantum, GSH1/GSH1*; 2, *GSH1/GSH1/HYG*; 3, *GSH1/GSH1/NEO*; 4, *GSH1/GSH1/NEO/HYG*; 5, *GSH1/GSH1/NEO/HYG/BLA*.
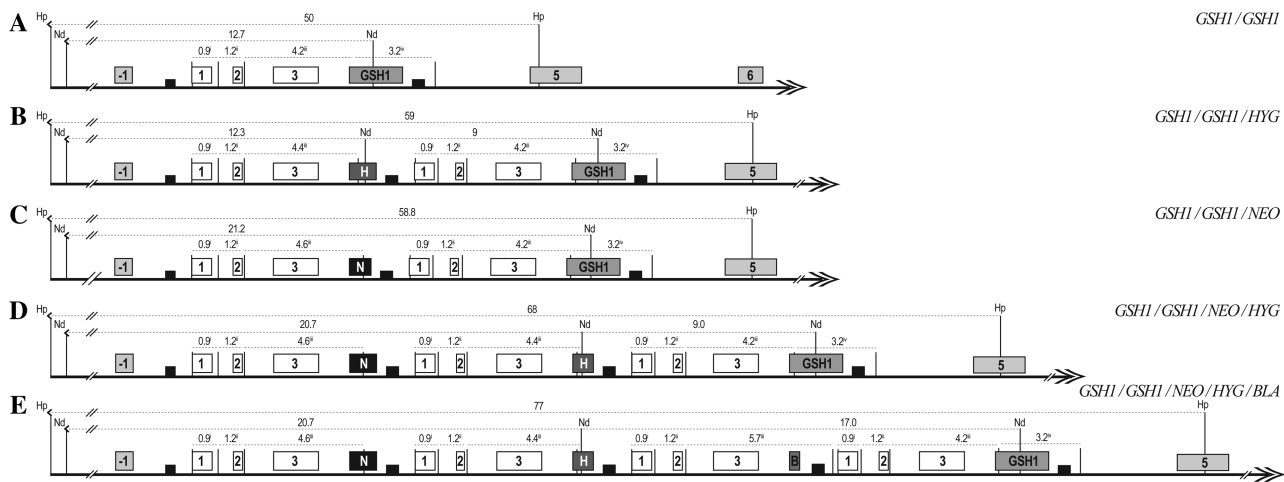
Figure 6. Tandem duplication of the *GSH1* locus by homologous recombination between direct repeats. (**A**) Schematic diagram of the *GSH1* locus in WT *L. infantum*. Only one of the chromosomal homologs of the disomic chromosome is shown. Numbers in boxes correspond to the genes in vicinity of *GSH1*. The estimated size of restriction fragments following NcoI, NdeI (Nd) and HpaI (Hp) digestions are shown with dotted lines along with the length of the restriction fragments. Superscripts denote the panels in Figure 4 where those restriction fragments are demonstrated by Southern blots. (**B–E**) The rearrangement at the genomic locus of the mutants when *GSH1* was targeted, either with *HYG* to form *GSH1/GSH1/HYG* (B) or *NEO* to generate *GSH1/GSH1/NEO* (C). The *GSH1/GSH1/NEO* targeted with *HYG* to form *GSH1/GSH1/NEO/HYG* (D); and the *GSH1/GSH1/NEO/HYG* targeted with *BLA* to generate *GSH1/GSH1/NEO/HYG/BLA* (E). H, hygromycin phosphotransferase B; N, neomycin phosphotransferase; B, blasticidin deaminase.

we first digested intact chromosomes in agarose blocks with HpaI, an enzyme which does not cut within the *GSH1* locus (Figure 6). These digests were run on CHEF and hybridized to genes located within the repeated sequences. Consistent with the scenario shown in Figure 6, all markers integrated into the same allele and increased the size of one chromosomal homolog by about 10 kb (the size of the *GSH1* locus between the 466-bp repeats) after each round of targeting, keeping the other homolog intact (Figure 5C, lanes 2–5). Probing with the *GSH1* open reading frame (ORF) revealed an expected equal intensity in both chromosomal homologs in all mutants (Figure 5D). Hybridization with *HYG*, *NEO* and *BLA* confirmed that all the markers integrated into the same allele (Figure 5D).

The four-gene locus between the 466-bp repeats has thus duplicated in tandem each time *GSH1* was targeted. In order to show that duplication is ordered, we digested the DNA of all the recombinants with NdeI, which cuts within the *HYG* and *GSH1* but neither within *NEO* and *BLA* nor elsewhere within the locus (Figure 6). Using a probe covering the first 1000 bp of *GSH1* (before the NdeI site), a NdeI digest should lead to a fragment of 12.7 kb in WT cells (Figure 6A) and, indeed, this was observed (Figure 7, lane 1). The same band was also observed in all recombinants, consistent with at least one intact *GSH1* allelic locus, (Figure 7A, lanes 2–5). Integration of *HYG* or *NEO* in the locus hybridized to the same *GSH1* probe should lead to 9-kb and 21-kb NdeI fragments (Figure 6B and C) only if the WT locus duplicated downstream of integration and this was indeed observed experimentally (Figure 7A, lanes 2 and 3). If the order of duplication mimics the order of inactivation attempts we should obtain the intermediate structure shown in Figure 6D and the final product shown in Figure 6E.

Digestion with NdeI and hybridization with the same *GSH1* probe would lead to hybridizing fragments of 9 kb and 17 kb in the *GSH1/GSH1/NEO/HYG* and *GSH1/GSH1/NEO/HYG/BLA* recombinants, respectively (Figure 6D and E) and this is what we have observed (Figure 7A, lanes 4 and 5). Hybridization of these NdeI-digested DNA with a probe derived from the first 434 bp of *HYG* (before the NdeI site) would yield a 12-kb fragment in *GSH1/GSH1/HYG* cells but a 21-kb fragment in the *GSH1/GSH1/NEO/HYG* and *GSH1/GSH1/NEO/HYG/BLA* cells (Figure 6B, D and E) and this was indeed observed (Figure 7B). Hybridization data of the same digests with *NEO* and *BLA* probes (Figure 7C and D) were also totally consistent with the maps shown in Figure 6. This suggests that upon each round of targeting, the WT allele is the one targeted and duplicated; so, the most recently targeted marker will be just upstream of the WT allele while additional markers will be more internal, in the order in which they were sequentially targeted.

The data are thus consistent with the duplication of the four gene locus containing *GSH1* that are within the 466-bp direct repeated sequences. This duplication occurred a first time upon the integration of either the *HYG* (Figure 6B) or *NEO* marker (Figure 6C), a second time upon the integration of the *HYG* marker into a *GSH1/GSH1/NEO* line (Figure 6D) and a third time upon integration of the *BLA* marker when *GSH1* was targeted into a *GSH1/GSH1/NEO/HYG* line (Figure 6E). All these integrations took place on the same chromosomal homolog explaining the increase in both the size and the hybridization intensity of chromosome 18 as seen in CHEF analysis upon each round of integration (Figure 5C and D). This process always maintains two copies of *GSH1* intact, despite repetitive targeting of *GSH1* and ordered integration of various markers used.
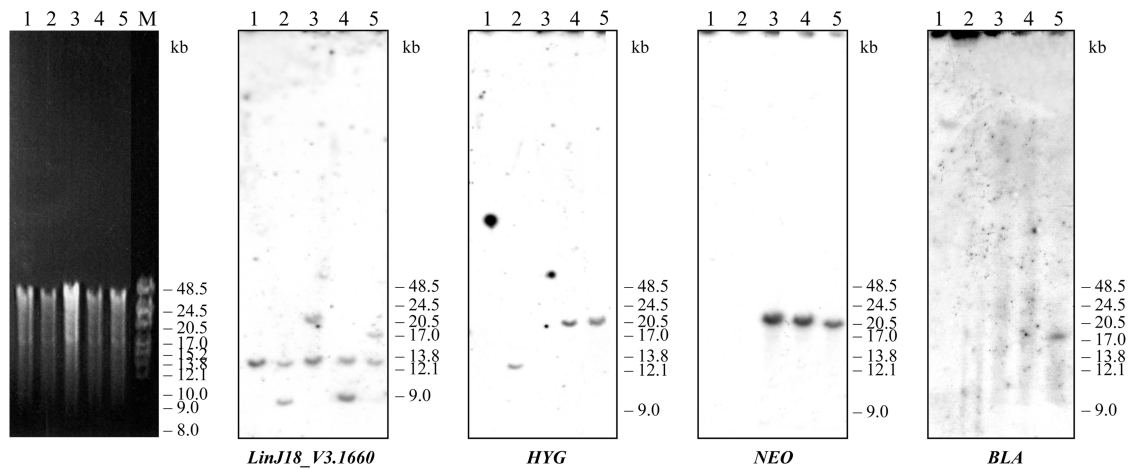
Figure 7. Duplication of the *GSH1* locus as determined by NdeI digests and Southern blot analyses. Leishmania genomic DNA was digested with NdeI and probed with the initial 1000 bp of *GSH1* ORF and 434 bp of *HYG* (up to their NdeI site) or complete ORFs of *NEO* and *BLA*. Molecular weight markers from High Range DNA ladder are indicated on the right. 1, WT *L. infantum, GSH1/GSH1*; 2, *GSH1/GSH1/HYG*; 3, *GSH1/GSH1/NEO*; 4, *GSH1/GSH1/NEO/HYG*; 5, *GSH1/GSH1/NEO/HYG/BLA*.
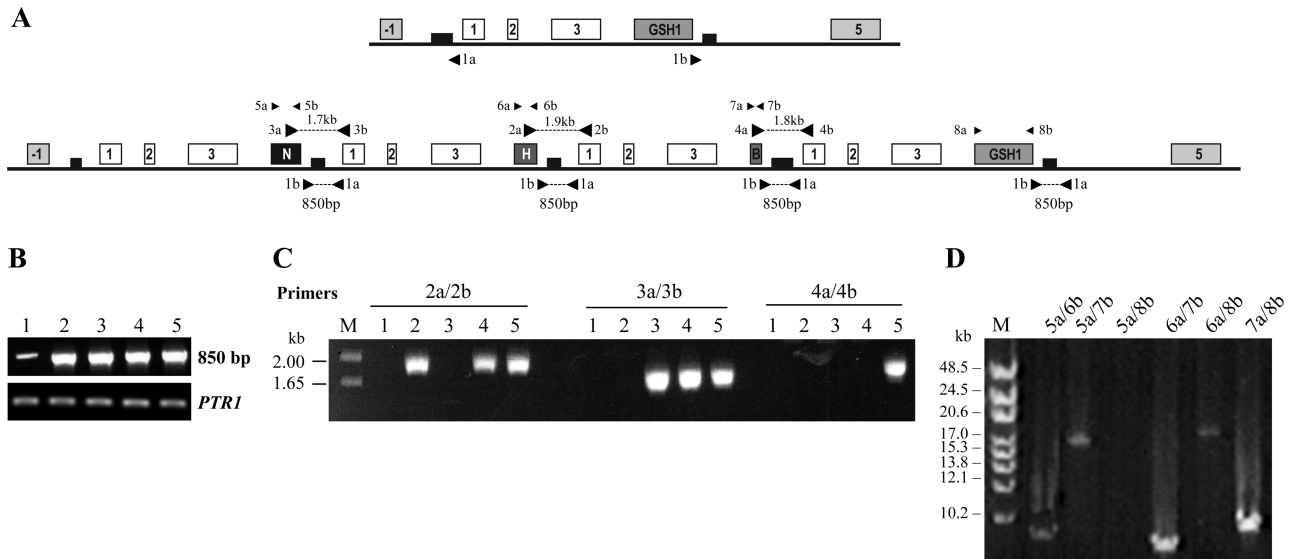
Figure 8. Duplication of *GSH1* locus at the level of direct repeats. (**A**) A pair of primers, 1a and 1b can amplify a region of 850 bp when the *GSH1* locus has duplicated in tandem. To detect the novel junctions and the sequential duplication of the locus, forward primers were also designed before the end of the marker genes (H, N and B) as indicated by 2–4a and reverse primers in the middle of *LinJ18_V3.1630* (box 1), indicated as 2–4b. Other primers 5a/5b to 8a/8b were used for long-range PCR. The genomic locus of the WT and triple mutant (*GSH1/GSH1/NEO/HYG/BLA*) are shown in this figure. (**B**) PCR amplification with the primers 1a and b using genomic DNA from WT and recombinants. *PTR1* was used for normalization of template DNA. (**C**) Detection of novel junctions due to the rearranged genomic locus in the recombinants using primer pairs 2a/2b, 3a/3b and 4a/4b. 1, WT *L. infantum, GSH1/GSH1*; 2, *GSH1/GSH1/HYG*; 3, *GSH1/GSH1/NEO*; 4, *GSH1/GSH1/NEO/HYG*; 5, *GSH1/GSH1/NEO/HYG/BLA*. (**D**) Genomic DNA of the *GSH1/GSH1/NEO/HYG/BLA* was amplified by long-range PCR (DyNAzyme™ EXT DNA Ploymerase, Finnzymes) to further validate the order of marker integration on chromosome 18. Lane 1, High Range DNA ladder; long-range PCR products with primers 5a/6b (2); 5a/7b (3); 5a/8b (4); 6a/7b (5); 6a/8b (6) and 7a/8b (7).

The data led us to suggest that the duplication occurs at the level of the direct repeated sequences.

This was further supported by PCR amplification of the novel junctions created. The pair of primers 1a and 1b (described in Figure 3) amplified a DNA fragment of 850 bp in the WT cells (Figure 8B, lane 1) but with more intensity in the recombinants (Figure 8B, lanes 2–5) with the same amount of template genomic DNA as normalized with amplification of *PTR1*

(*LinJ23_V3.0310*). The presence of this novel junction in WT cells using a highly sensitive PCR assay suggests that recombination between these repeated sequences is a relatively frequent event (see 'Discussion' section).

To confirm the rearrangement by tandem duplication in these *GSH1* recombinants, we designed primers to amplify the novel junctions generated by duplication but by having forward primers before the end of marker genes (Figure 8A) and reverse primers were designed in the

middle of *LinJ18_V3.1630* (Figure 8A). These three sets of primers in the mutants yielded the expected product at the appropriate size by PCR (Figure 8C) and no band was detected using WT genomic DNA (Figure 8C). The PCR products were sequenced and the analysis was consistent with the rearrangements occurring at the level of the repeats (data not shown) and confirmed the process of tandem duplication on chromosome 18. The ordered integration of all markers and duplication of WT locus on the same chromosomal homolog were verified by Southern blot analyses (Figures 5C and D and Figure 7). This was further supported by long-range PCR of the *GSH1/GSH1/NEO/HYG/BLA* recombinants (Figure 8D). Indeed, the *NEO* gene was 9.4 kb apart from *HYG* (Figure 8D, lane 2) and 19 kb from *BLA* (Figure 8D, lane 3) using long-range PCR. We failed, however, under our experimental conditions, to amplify the 27-kb fragment separating *NEO* and *GSH1* (Figure 8D, lane 4). The use of long-range PCR confirmed that the *HYG* gene was 9 kb apart from *BLA* (Figure 8D, lane 5) while the *HYG* and *BLA* genes were respectively 19 kb and 9 kb apart from the *GSH1* gene (Figure 8D, lanes 6 and 7). Other primer combinations did not give rise to PCR products. These results confirmed the structure of chromosome 18 shown in Figures 6E and 8A.

DISCUSSION

We recently demonstrated that *GSH1* is essential where the markers (*HYG*, *NEO* and *BLA*) integrated at the homologous locus but the recombinants invariably retained two WT copies of *GSH1*. However, when an episomal copy of *GSH1* was provided it was possible to generate a *GSH1* chromosomal null mutant (14).

Previous attempts to inactivate essential genes in *Leishmania* resulted in gene rearrangements such as supernumerary chromosomes (2) or genome wide polyploidy (13) or linear amplicons (19). For example, attempts to inactivate the dihydrofolate reductase-thymidylate synthetase (*DHFR-TS*) in *Leishmania* resulted in aneuploid trisomic lines, genomic tetraploids and diploids bearing homologous integration of the targeting fragment without replacement (2). In an attempt to inactivate trypanothione reductase (*TR*), both the WT alleles were disrupted but a third copy of *TR* was generated by genomic rearrangement, involving the translocation of *TR* containing region to a larger chromosome (12). Inactivation of the cdc2-related kinase (*crk1*) led the transfected fragment to form an episome, or the cloned transfectants were genomic triploids or tetraploids (13). In this study, we show a complete different mechanism of regional polyploidy emerging from tandem duplication of a segment bordered by direct repeats (Figure 6).

Homologous recombination between direct repeated sequences is well known for its implication in the formation of extrachromosomal circular elements (3,4,6,7,11). In our loss of heterozygosity attempts of *GSH1*, we observed circular amplification of the selectable markers (*HYG* and *NEO*) in *GSH1/GSH1/NEO/HYG* cells in whose media the selective drugs were raised 10-fold

(Figure 3C). These segments were amplified as extrachromosomal circles through HR between the 466-bp direct repeats bordering *LinJ18_V3.1630* and *GSH1* (*LinJ18_V3.1660*) (Figure 3C). The same 466-bp direct repeated sequences, present on chromosome 18, were used for tandem duplication of a four-gene region upon attempts to inactivate the essential *GSH1* gene. This is a novel strategy for *Leishmania* to generate polyploidy. It is intriguing that in WT cells (Figure 8B, lane 1) we could detect (using a sensitive PCR assay) a new junction, suggesting that this rearrangement occurs continuously in a limited number of cells and is selected when we attempted to disrupt *GSH1*. It is salient to point out, however, that it is not possible to discriminate whether this recombination leads to duplication (as seen for *GSH1*) or as extrachromosomal circles since such events could give rise to the same amplification product (compare Figure 3C and Figure 8A). Ongoing work in the lab suggests that this background amplification in *Leishmania* at the level of repeated sequences is common throughout the *Leishmania* genome (manuscript in preparation). In eubacteria, gene amplification by gene duplication is frequent. In a growing population of bacteria the frequency of spontaneous tandem duplication ranges from 10^{-2} to 10^{-4}, depending on the gene and the genomic region (20). These duplications provide a large reservoir of standing genetic variation from which selective pressure can favor cells with an increased copy number of any specific region. Once a tandem genetic duplication has been formed, selection for increased gene dosage may drive further amplification of the duplicated sequence, increasing the length of the genome by 10–80% in *Salmonella* and *Escherichia coli* (21–24). The mechanism of amplification from an initial duplication involves RecA-mediated nonequal crossing over between sequences on sister chromatids (20,25,26) or a rolling circle type of mechanism resulting in the formation of large tandem arrays (27). Intrachromosomal, clustered and tandem amplicons by palindromic duplication are also one of the somatic rearrangements in cancer cells, suggesting to occur by a breakage–fusion-bridge (BFB) mechanism (28,29).

Successful integration of the targeting fragment at *GSH1* was always accompanied by expansion of the targeted locus to include both the targeting DNA and the WT copy of *GSH1*, and, in every instance, the WT *GSH1* gene was at the telomere proximal side of the targeted locus (Figure 6). Interestingly, this duplication was not observed when a WT copy of *GSH1* was supplied on a plasmid. Gene targeting may thus occur in cells that had previously duplicated the *GSH1* locus by unequal sister chromatid exchange mediated at the level of the direct repeats flanking the locus, and only those cells with pre-existing duplications survive selection for the marker on the targeting DNA. Indeed, unequal sister chromatid exchange has often been invoked for explaining naturally occurring tandem duplications in several organisms (30,31). If this were the case, however, one would expect that the duplicated alleles would be targeted randomly, but, instead, we observed a specific order of integration suggesting that the process of gene
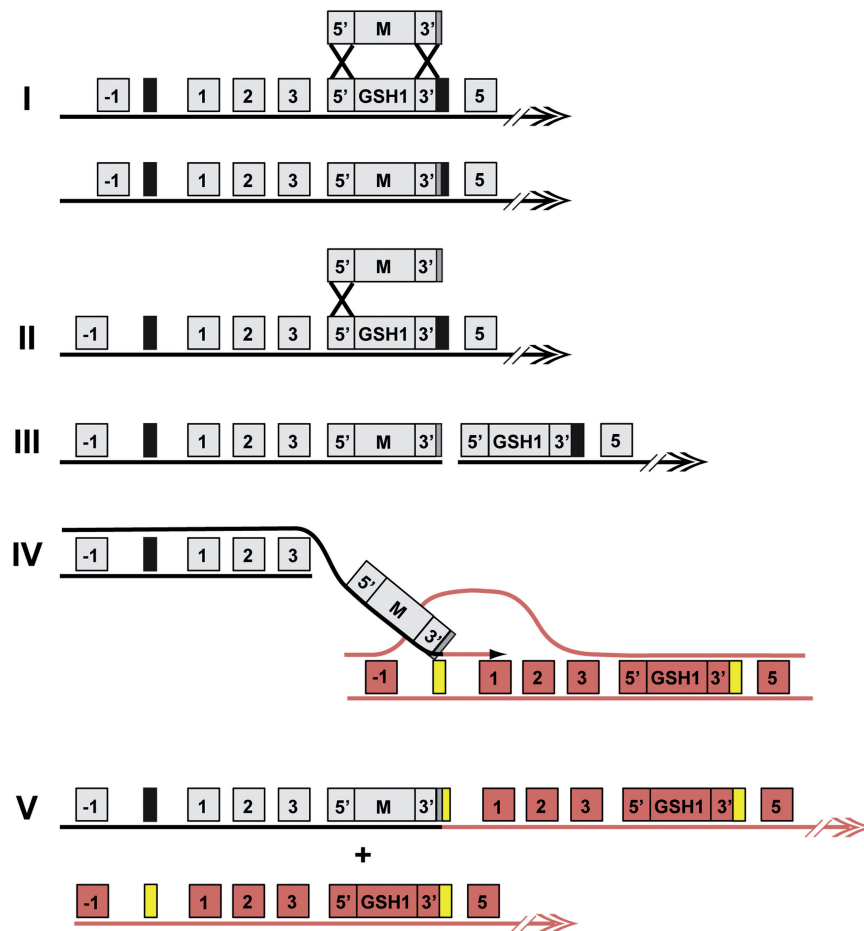
Figure 9. Tandem duplication of a subtelomeric essential gene in *Leishmania* by break induced replication. (I) Inactivation of *GSH1* by gene replacement with a targeting cassette by homologous recombination in the 5′ and 3′ flank involving a double cross over would result in the integration of the marker (M) and disruption of *GSH1*. (II) However, if the integration of the inactivation cassette into the homologous position happens by a single crossover to survive the selection pressure, it creates a double-strand break (DSB) on the chromosome (III). (IV) At the DSB, a 5′→3′ degradation of DNA takes place, exposing single-strand DNA. Single-strand DNA could initiate pairing and strand invasion of homologous duplex DNA from the sister chromatid or the other homologous allele (red). The 3′ flank of *GSH1* containing the first 100 bp of the repeated sequences (black box) could misalign and pair with the repeated sequence of the homologous duplex DNA (yellow box) and synthesis begins. (V) DNA synthesis is extended and replication continues to the end of the chromosome yielding a chromosomal copy with a duplicated locus in tandem and the other one intact. Double arrowhead represents the telomere.

targeting might lead to the duplication rather than simply taking advantage of a pre-existing duplication.

The fact that *GSH1* is flanked by direct repeated sequences and is subtelomeric lead us to propose a model for duplication of the *GSH1* locus involving partial integration of the targeting DNA followed by break-induced replication (BIR), as shown in Figure 9. In the instance where targeted gene replacement occurs, loss of the essential *GSH1* gene leads to cell death, so these events are not observed (Figure 9I). Instead, we propose that one end of the targeting DNA integrates into the chromosome before the other (Figure 9II) and this single crossover creates a double-strand break (DSB) in the chromosome (Figure 9III). The concept of single crossover at one end of the targeting fragment has been previously shown in yeast (32). Double-strand breaks on chromosomes are lethal and must be repaired. Several sub-pathways involving HR have been defined for repair and these include double-strand break repair (DSBR),

synthesis-dependent strand annealing (SDSA) and BIR (33). Repair of this DSB during gene targeting by either SDSA or gene conversion would lead to removal of either the targeting DNA or the *GSH1* gene, so these events would not survive selection. Instead, we propose that the break is repaired by BIR, a mechanism well studied in *Saccharomyces cerevisiae* (34). During BIR, the centromere-proximal end of a break is processed into a 3′ overhang, which can invade the sister chromatid or the other copy of the chromosome using homologous sequences and use it as a template for repair. DNA polymerase-mediated synthesis completes the repair (35). In *Leishmania*, duplication of the *GSH1* locus occurs when the 466-bp direct repeated sequence on the left side of the break invades an identical 466-bp repeat upstream of the *GSH1* locus on the sister chromatid/homologous chromosome (Figure 9IV) and copies to the end of the chromosome, resulting in a characteristic tandem duplication of the locus containing the wild type *GSH1* gene on the

telomere proximal side, and keeping the other allele intact (Figure 9V). If the targeting construct did not contain any part of the 466-bp repeated sequences, we speculate that tandem duplication could still have occurred by BIR with two or more rounds of invasion (Supplementary Figure S2), a phenomenon already described in yeast (36). The presence of the repeats on the targeting DNA probably increases the likelihood of misaligned BIR dramatically because it can occur on the first round of strand invasion and does not involve secondary steps, namely invasion, synthesis, strand displacement and re-invasion. Other mechanisms could also explain the gene duplication observed, notably unequal sister chromatid exchange, but this model parsimoniously explains the duplication and the specific orientation of the two copies of the duplicated locus, and is supported by the presence of the direct repeats flanking *GSH1* and by the subtelomeric location of the *GSH1* gene. The subtelomeric regions of eukaryotes are hotspots of genetic rearrangements and most of these events are best explained by the introduction of a DSB that is repaired by BIR using repetitive sequence elements located in the same or different chromosomes (36,37). Recent work has shown that a programmed double strand break in *Trypanosoma brucei* is an effective, natural trigger of variant surface glycoprotein (VSG) switching and the repair of DSB is probably achieved by BIR and a repetitive sequence adjacent to a VSG gene in the telomeric loci facilitate BIR through homology recognition (37). Also, compelling evidence was given to support the idea of duplicative gene conversion into an active VSG expression site for activation of a new VSG in *T. brucei* (38). BIR is also used for repair of shortened telomeres in the absence of telomerase in numerous instances (39). While these processes are mechanistically distinct from the duplication observed during targeting of the *GSH1* locus, it was the prevalence of repeat-mediated, break-dependent subtelomeric DNA rearrangements in a wide variety of organisms, including *T. brucei*, that first suggested a possible model by which gene targeting might drive the duplications described here.

Leishmania is a diploid organism with considerable genome plasticity. It was soon realized that targeting essential genes can lead to gene rearrangements (2,12,13,19). In this study, we provide evidence for a new type of rearrangement leading to the duplication of a region at the level of repeated sequences. Sequence analysis of the *Leishmania* genome has revealed several gene duplication and it is believed that smaller gene families (<10 members) appears to have arisen from tandem duplication (40). Tandem duplication of a gene cluster (two to three genes) is also found but with a lesser frequency in some chromosomes and it could be possible that some of these events took place by intrachromosomal HR between repeated sequences resulting in repeat expansion. A preliminary analysis partially supports this hypothesis. We found six examples where a gene cluster (two or three genes) is duplicated in tandem intrachromosomally. Two of these clusters were likely duplicated by HR through these repeats. For the other clusters either we had no evidence for this event or they have undergone complex rearrangements which could have led to the loss of

repeated sequences. A group of eight genes involved in lipophosphoglycan biosynthesis or modification (the *SCGR* and *SCA* genes) are clustered on chromosome 2 (41). The arrangement would suggest a duplication of four genes and indeed repeated sequences of 412 bp (98% identity) are strategically positioned to duplicate the locus (Supplementary Figure S3A). The other example we found is that of the tandem duplication of a hypothetical gene and the gene coding for *S*-adenosylmethionine synthetase on chromosome 30 (42). On inspecting the genomic locus we discovered direct repeated sequences of 212 bp (98% identity) at positions consistent with gene duplication through HR at the level of 212-bp repeats (Supplementary Figure S3B). Expansion of gene families by tandem duplication by HR at the level of direct repeats is a potential mechanism by which parasites can increase gene copy number and this was shown experimentally here during gene inactivation attempts of *GSH1*. Further work should allow testing for additional roles for repeated sequences in gene duplication in *Leishmania* and whether unequal sister chromatid exchange and BIR (Figure 9) can explain such duplicative events.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

FUNDING

REFERENCES

1. Scholler,J.K., Reed,S.G. and Stuart,K. (1986) Molecular karyotype of species and subspecies of Leishmania. *Mol. Biochem. Parasitol.*, **20**, 279–293.
2. Cruz,A.K., Titus,R. and Beverley,S.M. (1993) Plasticity in chromosome number and testing of essential genes in Leishmania by targeting. *Proc. Natl Acad. Sci. USA*, **90**, 1599–1603.
3. Leprohon,P., Legare,D., Raymond,F., Madore,E., Hardiman,G., Corbeil,J. and Ouellette,M. (2009) Gene expression modulation is associated with gene amplification, supernumerary chromosomes and chromosome loss in antimony-resistant Leishmania infantum. *Nucleic Acids Res.*, **37**, 1387–1399.
4. Ubeda,J.M., Legare,D., Raymond,F., Ouameur,A.A., Boisvert,S., Rigault,P., Corbeil,J., Tremblay,M.J., Olivier,M., Papadopoulou,B. *et al.* (2008) Modulation of gene expression in drug resistant Leishmania is associated with gene amplification,

gene deletion and chromosome aneuploidy. *Genome Biol.*, **9**, R115.

5. Sterkers,Y., Lachaud,L., Crobu,L., Bastien,P. and Pages,M. (2010) FISH analysis reveals aneuploidy and continual generation of chromosomal mosaicism in Leishmania major. *Cell. Microbiol.*, **13**, 274–283.

6. Beverley,S.M. (1991) Gene amplification in Leishmania. *Annu. Rev. Microbiol.*, **45**, 417–444.

7. Ouellette,M., Hettema,E., Wust,D., Fase-Fowler,F. and Borst,P. (1991) Direct and inverted DNA repeats associated with P-glycoprotein gene amplification in drug resistant Leishmania. *EMBO J.*, **10**, 1009–1016.

8. Papadopoulou,B., Roy,G. and Ouellette,M. (1993) Frequent amplification of a short chain dehydrogenase gene as part of circular and linear amplicons in methotrexate resistant Leishmania. *Nucleic Acids Res.*, **21**, 4305–4312.

9. Borst,P. and Ouellette,M. (1995) New mechanisms of drug resistance in parasitic protozoa. *Annu. Rev. Microbiol.*, **49**, 427–460.

10. Grondin,K., Papadopoulou,B. and Ouellette,M. (1993) Homologous recombination between direct repeat sequences yields P-glycoprotein containing amplicons in arsenite resistant Leishmania. *Nucleic Acids Res.*, **21**, 1895–1901.

11. Grondin,K., Roy,G. and Ouellette,M. (1996) Formation of extrachromosomal circular amplicons with direct or inverted duplications in drug-resistant Leishmania tarentolae. *Mol. Cell. Biol.*, **16**, 3587–3595.

12. Dumas,C., Ouellette,M., Tovar,J., Cunningham,M.L., Fairlamb,A.H., Tamar,S., Olivier,M. and Papadopoulou,B. (1997) Disruption of the trypanothione reductase gene of Leishmania decreases its ability to survive oxidative stress in macrophages. *EMBO J.*, **16**, 2590–2598.

13. Mottram,J.C., McCready,B.P., Brown,K.G. and Grant,K.M. (1996) Gene disruptions indicate an essential function for the LmmCRK1 cdc2-related kinase of Leishmania mexicana. *Mol. Microbiol.*, **22**, 573–583.

14. Mukherjee,A., Roy,G., Guimond,C. and Ouellette,M. (2009) The gamma-glutamylcysteine synthetase gene of Leishmania is essential and involved in response to oxidants. *Mol. Microbiol.*, **74**, 914–927.

15. Papadopoulou,B., Roy,G. and Ouellette,M. (1992) A novel antifolate resistance gene on the amplified H circle of Leishmania. *EMBO J.*, **11**, 3601–3608.

16. Sambrook,J., Fritsch,E.F. and Maniatis,T. (1989) *Molecular Cloning*. Cold Spring Harbour Laboratory, New York.

17. Grondin,K., Haimeur,A., Mukhopadhyay,R., Rosen,B.P. and Ouellette,M. (1997) Co-amplification of the gamma-glutamylcysteine synthetase gene gsh1 and of the ABC transporter gene pgpA in arsenite-resistant Leishmania tarentolae. *EMBO J.*, **16**, 3057–3065.

18. Gueiros-Filho,F.J. and Beverley,S.M. (1996) Selection against the dihydrofolate reductase-thymidylate synthase (DHFR-TS) locus as a probe of genetic alterations in Leishmania major. *Mol. Cell. Biol.*, **16**, 5655–5663.

19. Genest,P.A., ter Riet,B., Dumas,C., Papadopoulou,B., van Luenen,H.G.A.M. and Borst,P. (2005) Formation of linear inverted repeat amplicons following targeting of an essential gene in Leishmania. *Nucleic Acids Res.*, **33**, 1699–1709.

20. Anderson,P. and Roth,J. (1981) Spontaneous tandem genetic duplications in *Salmonella typhimurium* arise by unequal recombination between rRNA (rrn) cistrons. *Proc. Natl Acad. Sci. USA*, **78**, 3113–3117.

21. Edlund,T., Grundstrom,T. and Normark,S. (1979) Isolation and characterization of DNA repetitions carrying the chromosomal beta-lactamase gene of Escherichia coli K-12. *Mol. Gen. Genet.*, **173**, 115–125.

22. Andersson,D.I., Slechta,E.S. and Roth,J.R. (1998) Evidence that gene amplification underlies adaptive mutability of the bacterial lac operon. *Science*, **282**, 1133–1135.

23. Andersson,D.I. and Hughes,D. (2009) Gene amplification and adaptive evolution in bacteria. *Annu. Rev. Genet.*, **43**, 167–195.

24. Edlund,T. and Normark,S. (1981) Recombination between short DNA homologies causes tandem duplication. *Nature*, **292**, 269–271.

25. Haack,K.R. and Roth,J.R. (1995) Recombination between chromosomal IS200 elements supports frequent duplication formation in Salmonella typhimurium. *Genetics*, **141**, 1245–1252.

26. Shyamala,V., Schneider,E. and Ames,G.F. (1990) Tandem chromosomal duplications: role of REP sequences in the recombination event at the join-point. *EMBO J.*, **9**, 939–946.

27. Petit,M.A., Mesas,J.M., Noirot,P., Morel-Deville,F. and Ehrlich,S.D. (1992) Induction of DNA amplification in the *Bacillus subtilis* chromosome. *EMBO J.*, **11**, 1317–1326.

28. Tanaka,H. and Yao,M.C. (2009) Palindromic gene amplification – an evolutionarily conserved role for DNA inverted repeats in the genome. *Nat. Rev. Cancer*, **9**, 216–224.

29. Haber,J.E. and Debatisse,M. (2006) Gene amplification: yeast takes a turn. *Cell*, **125**, 1237–1240.

30. Rauen,K.A., Bitts,S.M., Li,L., Golabi,M. and Cotter,P.D. (2001) Tandem duplication mosaicism: characterization of a mosaic dup(5q) and review. *Clin. Genet.*, **60**, 366–370.

31. Mattina,T., Sorge,G., Milone,G., Garozzo,R. and Conti,L. (1987) Duplication 9p due to unequal sister chromatid exchange. *J. Med. Genet.*, **24**, 303–305.

32. Langston,L.D. and Symington,L.S. (2004) Gene targeting in yeast is initiated by two independent strand invasions. *Proc. Natl Acad. Sci. USA*, **101**, 15392–15397.

33. Llorente,B., Smith,C.E. and Symington,L.S. (2008) Break-induced replication: what is it and what is it for? *Cell Cycle*, **7**, 859–864.

34. Kraus,E., Leung,W.Y. and Haber,J.E. (2001) Break-induced replication: a review and an example in budding yeast. *Proc. Natl Acad. Sci. USA*, **98**, 8255–8262.

35. Holmes,A.M. and Haber,J.E. (1999) Double-strand break repair in yeast requires both leading and lagging strand DNA polymerases. *Cell*, **96**, 415–424.

36. Smith,C.E., Llorente,B. and Symington,L.S. (2007) Template switching during break-induced replication. *Nature*, **447**, 102–105.

37. Boothroyd,C.E., Dreesen,O., Leonova,T., Ly,K.I., Figueiredo,L.M., Cross,G.A. and Papavasiliou,F.N. (2009) A yeast-endonuclease-generated DNA break induces antigenic switching in *Trypanosoma brucei*. *Nature*, **459**, 278–281.

38. Kim,H.S. and Cross,G.A. (2010) TOPO3alpha influences antigenic variation by monitoring expression-site-associated VSG switching in *Trypanosoma brucei*. *PLoS Pathog.*, **6**, e1000992.

39. McEachern,M.J. and Haber,J.E. (2006) Break-induced replication and recombinational telomere elongation in yeast. *Annu. Rev. Biochem.*, **75**, 111–135.

40. Ivens,A.C., Peacock,C.S., Worthey,E.A., Murphy,L., Aggarwal,G., Berriman,M., Sisk,E., Rajandream,M.A., Adlem,E., Aert,R. *et al.* (2005) The genome of the kinetoplastid parasite, Leishmania major. *Science*, **309**, 436–442.

41. Dobson,D.E., Scholtes,L.D., Myler,P.J., Turco,S.J. and Beverley,S.M. (2006) Genomic organization and expression of the expanded SCG/L/R gene family of Leishmania major: internal clusters and telomeric localization of SCGs mediating species-specific LPG modifications. *Mol. Biochem. Parasitol.*, **146**, 231–241.

42. Garcia-Estrada,C., Perez-Pertejo,Y., Ordonez,D., Balana-Fouce,R. and Reguera,R.M. (2007) Analysis of genetic elements regulating the methionine adenosyltransferase gene in Leishmania infantum. *Gene*, **389**, 163–173.