

genome [49]. Different degrees of degeneracy of both the coding region and the flanking REs were observed in all the duplicated loci examined. This is consistent with the traditional view of the fate of new duplicons [6,9], which assumes a tendency to be lost because of genetic drift under natural evolution [29,57] while not precluding the possibility for some duplicates to evolve distinct functions either by sub-functionalization or neo-functionalization.

The role of repetitive elements (REs) in dispersed duplication of genomic sequences is fairly documented from previous studies in model organisms [15,20,27,28,30,58,59]. The precise mechanism of this retrotransposon mediated dispersed duplication is not clear but may likely involve RE-mediated DNA level recombination, most likely by non-allelic homologous recombination (NAHR), alternatively called ectopic recombination (see illustration in Additional file 6). Due to their extremely high copy numbers, REs create structural modifications in the genome

by providing the requisite highly similar DNA sequences, initiating recombination between non-allelic elements [20,25,60], the result of which could be deletion, shuffling, duplication or transduction of a genomic DNA segment. Structural modifications introduced in the genome by NAHR mechanism can progress between non-homologous chromosomes (inter-chromosomal), between homologous chromosomes (inter-homologous or intra-chromosomal), between sister chromatids (inter-sister chromatid) or within a chromatid (intra-chromatid); giving rise to dispersed duplication of genomic segments, several forms of deletions or may create isodicentric chromosome by forming a mirrored segment in the chromosome by inversion. See detailed cartoon in Additional file 6. [60].

Many studies in other organisms have elucidated the role of REs in mediating sequence duplication, transduction and other structural variations by ectopic recombination mechanism. Notable among these is the human *Alu* element for which several reports suggest a role in

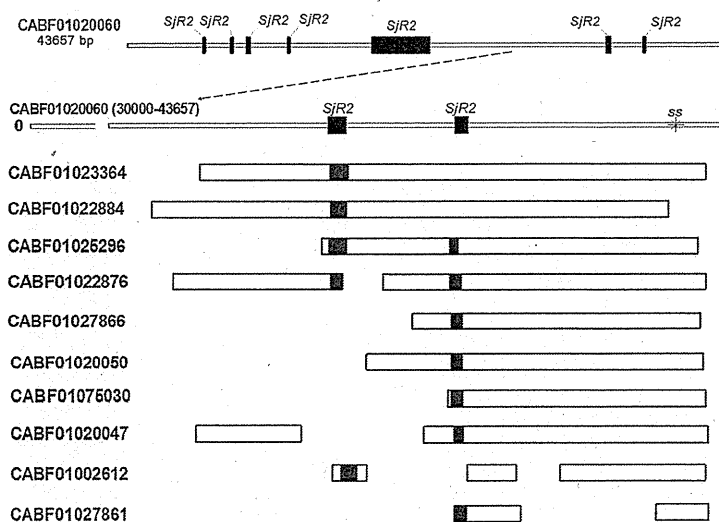


Figure 4 Further evidence that the duplicated genes were duplicons of a single duplication source locus. Apart from the prominent flanking copies of retrotransposons observed around the putative gene duplication source locus [GenBank:CABF01020060], other two short copies of the retrotransposon (*RTE_SJ*) are also found within introns in the coding region. We aligned the source locus with 10 of the duplicons and observed that both the signal sequence and these two partial copies of *RTE_SJ* are relatively aligned at same position, further indicating that the duplicated genes could have originated from a single source locus.

mediating NAHR and other structural modifications in the human genome [7,20,61]. Yang *et al* found an excess of repetitive sequences proximate to the breakpoints of duplicated gene loci in the genome of the fruit fly *Drosophyla melanogaster*, and have suggested that a NAHR mechanism, mediated by REs accounted for the birth of the new duplicons [1,27]. Another study performed on human individuals concluded that NAHR accounted for over 40 % of detected genomic sequence duplications in the human genome [30]. Illegitimate recombination (IR), incomplete crossing over and non-homologous end joining (NHEJ) are other possible mechanisms of gene duplication by DNA-level recombination, but NAHR play a more significant role in producing typical dispersed duplications [1] while the other mechanisms in addition to NAHR are more likely to produce tandem duplicates. Although we could not clearly identify the exact breakpoints of the duplications at both ends still for lack of a reference ancestral homolog and partly due to sequencing gaps, the fact that homology among all the scaffolds examined uniformly terminated at the same point with *Perere* on the 3' end (Figure 3 and Additional file 5), and traces of the observed predominant retrotransposons (*Sjr2*) was found at the exact positions as they occur in the putative source locus (Figure 4) confirm that these gene loci could be products of dispersed duplication from a single genomic source locus.

In addition to RE-mediated DNA-level recombination by NAHR, gene duplication events are also attributable to RE-mediated retrotransduction mechanism either on

the 5' or 3' directions [27]. Xing *et al* and other groups have demonstrated the role of retrotransposons in the duplication of entire genes and creation of previously un-described genes by analyzing SVA (SINE, VNTR and Alu)-mediated retrotransduction events in the human genome [20,29]. However, we did not specifically identify any chimeric duplicon originating via a retrotransduction mechanism among our datasets. Furthermore, retrotransposons including *Sjr2* characteristically encode reverse transcriptase and endonuclease, and can therefore transcribe and 'paste' a gene sequence into new locations in the genome [3,22,62]. However, retrotransposed genes are characteristically intronless since the introns are usually spliced out during the process of retrotransposition. Our duplicons retained their introns, although in some case some portion of the introns may have either degenerated or deleted during duplication and subsequent sequence modifications [3,22,63]. A further evidence that a retrotransposition mechanism is unlikely in our observed cases was that while retrotransposons would not duplicate the promoter regions of duplicated gene based on the process of transcription and insertion of retrocopies [1,57] which leaves the newly retrotransposed sequences to acquire new regulatory sequences from adjacent genes or through mutations in order to be functional [14,19,24]; the protein coding duplicons observed among our duplicated gene loci retained the same or similar core regulatory region and signal sequence as the source locus, suggesting that they may not have been products of retroposition and

may equally explain the parallel assumption of coding potential at their new duplication loci without the need to form chimeric structures with adjacent genes.

Evolution of translatable ORF and evidence of expression of duplicated genes

Some of the duplicons appear degenerative in homology and are relatively shorter than the source locus (Figure 3, also see Additional file 5) thus are consequently redundant and non-coding at the new locations as opined in the canonical view on the fate of new duplicons [6,9]; which assumes a tendency to be lost because of genetic drift under natural evolution [29,57]. However, our data provide evidence that some of the duplicons have evolved into protein coding genes with distinct products at their new loci, the fate of which could tend to either sub-functionalization to the source gene [8,64] or neo-functionalization by acquiring new distinct functions [9,65]. In addition to the two duplicons with alternative splicing variants, which we further explored in the next section, some representatives of the protein coding duplicons were depicted in a supplementary figure (Additional file 7). The nucleotide sequences of these genes are still appreciably similar but accumulation of mutations and other sequence modifications have given rise to novel protein coding ORFs, encoding putatively distinct products. We identified and mapped each cDNA sequence to the genomic contigs using information we generated from GeneMark and GeneQuest gene predictions [66] and confirmed by alignment of the cDNAs to the genomic sequences using NCBI *Splicing* program. This approach was necessitated because the fully mapped and annotated genome of *S. japonicum* is not presently available in the public databases. Intriguingly, our results corroborate the available UniGene and GenBank entries. Nevertheless, it is notable that we only assessed the duplicated copies on the basis of possessing the similar signal sequence. There is possibility that some other duplicons from this source locus could be involved in initiating other forms of structural modifications at other loci when incorporated into the coding region of other genes, but this was not investigated here.

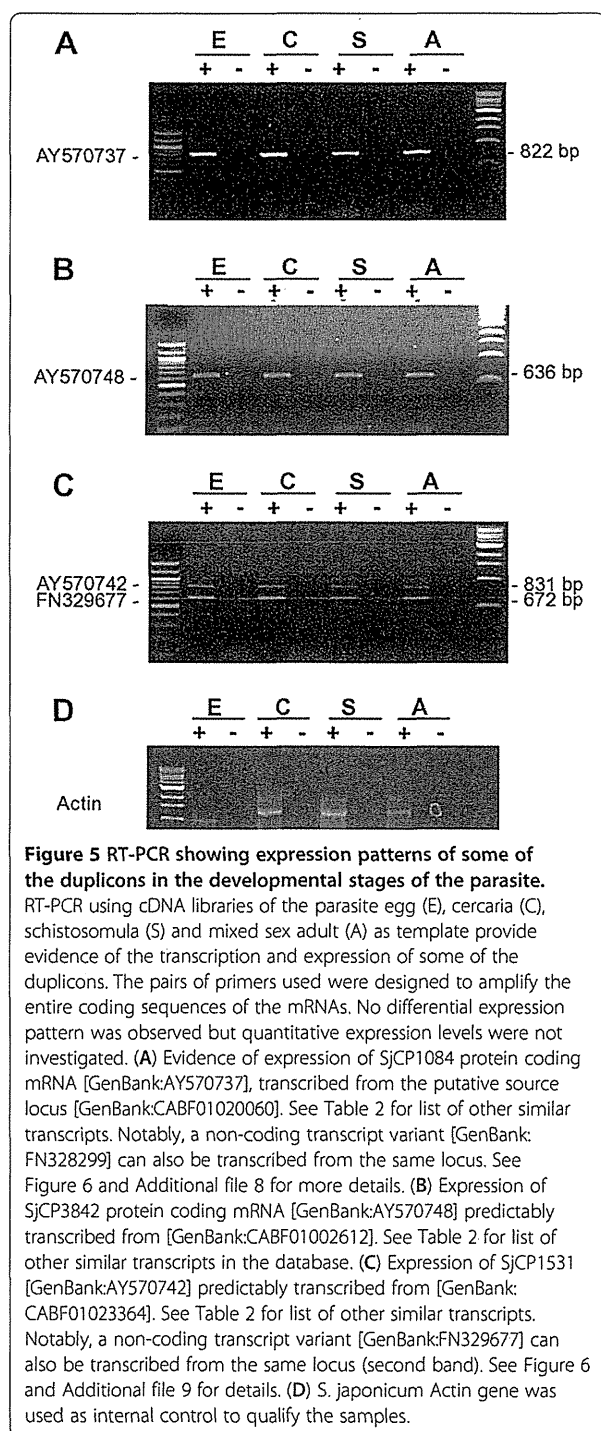
To provide evidence of the transcription and expression of the putative source gene locus and some of the duplicons, we performed developmental stage specific RT-PCR using primers that specifically amplify the coding regions of the candidate genes from the cDNA libraries of each stage of *S. japonicum*. RT-PCR results provide evidence of the transcription of some of the duplicons at their new genomic sites in addition to the source locus (Figure 5). The candidate genes analyzed did not show differential developmental stage specific expression, although we did not perform quantitative estimation of expression levels. It is possible that this

group represents a potential new family of proteins with similar signal peptides in this zoonotic trematode, which possess other extra distinctive characteristics from other members of the genus *Schistosoma*. We are presently undertaking further research to fully characterize the identified novel protein-coding genes to provide insight into the functional and structural significance of this trend in the genome of *S. japonicum*. The protein products of some of these candidate genes have already been expressed in our laboratory and confirmed by the reactivity of the immune sera with the parasite crude antigen preparations. The data will be reported with the molecular and functional characterization information.

Functional selection by alternative splicing

The precise recognition of exon-intron junction in a precursor mRNA (pre-mRNA) by the splicing machinery is central for the production of functional translatable mRNAs. However, there is often uncertainties in the choice of recognizable splice signals, resulting in a process termed alternative splicing [17], which enables the origination of multiple mRNA transcript variants from a single gene locus [67-69]. Alternative splicing mechanism could result in 'intronization' of an exon or 'exonization' of an intronic sequence. Ideally, the creation of an intron from a previously exonic sequence could lead to the loss of an ORF in coding genes. In evolving genes however, functional selection possibly by mutations may evolve the required splice signals and induce the intronization of an exon in a transcribed but non-coding mRNA gene sequence to create a translatable ORF encoding a functional protein. Conversely, while exonization of an intron could disrupt a translatable ORF in a coding gene, selective pressure may also evolve new splice signals within an intron to yield exons that could create a translatable ORF from a previously non-coding gene locus or a chimeric ORF from a protein-coding gene.

These two mechanisms have been shown from our observations to be capable of creating functional coding-genes from previously non-coding albeit transcribed mRNA sequences. We identified at least two classical evidences of alternative splicing and we propose that in addition to increasing coding potential and genomic diversity [68,69], alternative splicing can also be one of the driving forces of adaptive evolution; producing genetic novelties and functional selection. The most prominent example of alternative splicing was observed in the duplication source locus [GenBank:CABF01020060], which was found to be able to produce a protein-coding mRNA [GenBank:AY570737] in addition to a non-coding mRNA transcript variant [GenBank:FN328299] (Figure 6). An alignment of the DNA sequences of these two transcripts with details of this observation is

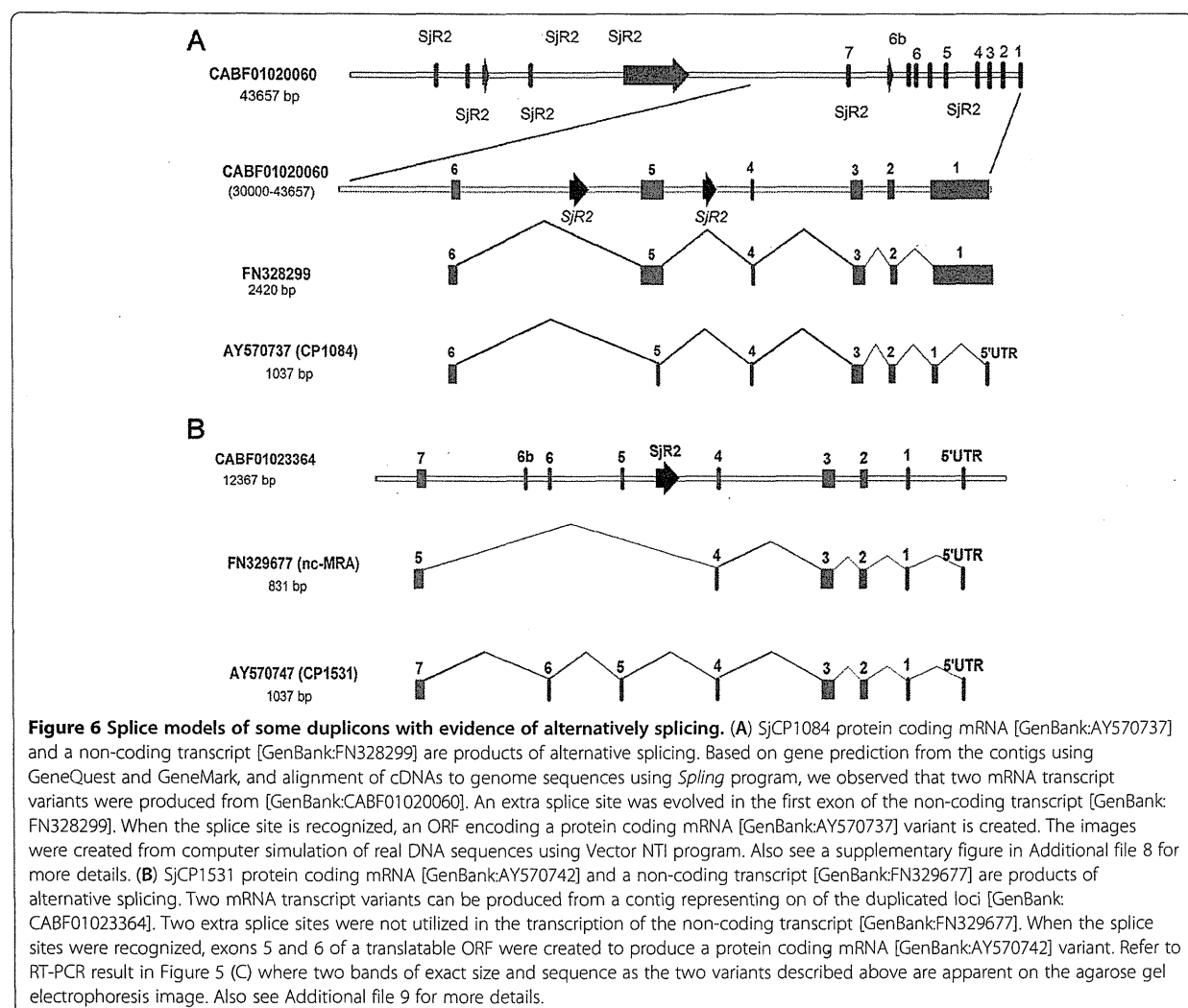


presented in the on-line published supporting information (Additional file 8). UniGene entries also suggest that the two transcripts are from the same locus (Table 2). An extra intron donor and acceptor sites were found within the first exon of the non-coding mRNA transcript

[GenBank:FN328299]. While the transcription model of the non-coding variant did not recognize the extra splice signals and thus retained the intron of about 1 kb, the coding mRNA variant [GenBank:AY570737] recognized the splice sites and created an ORF from the gene by splicing out an intron thereby giving rise to the 5' untranslated region (5'UTR) and the first exon of a protein-coding gene encoding a protein product of 271 residues (SjCP1084). Additionally, another pair of splice acceptor and donor sites evolving at exon 5 of the non-coding variant resulted in the splicing out of a portion of the exon, all contributing in creating a translatable ORF in the protein coding variant (See Additional file 8 for details).

On the other hand, exons 5 and 6 of a coding mRNA variant [GenBank:AY570742] predictably transcribed from one of the progeny loci [GenBank:CABF01023364] were skipped in a non-coding shorter variant [GenBank:FN329677] without a translatable ORF (Figure 6 and Additional file 9). We observed that the sequences of exons 5 and 6 were similar and was repeated five times *in tandem* within this locus, but only two copies of the tandemly duplicated potential exons were incorporated into the coding sequence of the mRNA to create exons 5 and 6 of a protein-coding ORF of 274 codons (SjCP1531). These results represent typical models of alternative splicing by intronization and exonization respectively.

Although in evolutionary perspective, intron retention that creates a translatable ORF is considered more plausible than the reverse process; our data show that both mechanisms are potentially possible. Other groups have also identified intron gains recently in mammalian and rodent retrogenes [68,69]. The identification of non-coding mRNA variant alternatively transcribed from a single gene locus with a protein coding mRNA (Figure 6) is evidence that a novel protein-coding gene can originate from previously transcribed regions that contain the necessary transcription elements and provide RNA material for a protein translation machine [2,39,68]. Exon repetition has also been observed from our data to exist in this organism and could be instrumental in expanding the organism's coding potential. The 'parallel' expression of the non-coding variant alongside the protein-coding transcripts is of significance and could suggest further that the gene may have been recently evolved. Non-coding RNAs have also been shown to perform some regulatory roles at various levels during gene expression [2,68,70]. This could be further explored with our data set. In the two described cases in our analyses, we have treated the non-coding isoforms as evolutionally preceding the coding variants; nevertheless, the reverse could also be the case. In addition to these two cases, we also identified a two-nucleotide insertion into a non-coding mRNA sequence [GenBank:FN330540] that yielded



the coding mRNA of schistosomula protein with the similar signal peptide, with many similar transcripts in the database. However, this last observation could be an artifact from sequencing error since the existence of the non-coding transcript was not traceable to the genomic sequence.

Conclusions

We have passably delineated the possible mechanism that led to the identification of several protein coding genes with similar signal sequence, following lead from our work that isolated secreted proteins candidate genes using SST. A trend was described in the genome of *S. japonicum* whereby a 'newly evolved' gene served as a source locus for dispersed duplication events leading to the formation of several expressed genes with similar transcription core promoter region and signal sequence.

We further found that the duplicated gene locus was flanked by non-long terminal repetitive elements (REs), especially of the *RTE*-like and *Perere* class. We therefore inferred that REs may have played an important role in this dispersed gene duplication by creating the requisite homologous DNA sequence that mediate a DNA-level recombination, most probably by a non-allelic homologous recombination (NAHR) mechanism. Our findings also provide evidence of logical sequential process of novel gene origination by evolution of transcription core elements followed by translatable ORF. While similar RE mediated phenomena had been observed in other organisms, unlike our dataset, most analyses have centered on the model organisms. Our data contribute to the accumulating evidence that REs mediate diverse recombination events leading to novel gene origination and other evolutionary novelties.

Methods

BLAST search

We had earlier identified a particular 81 nucleotides (27 amino acids) sequence, which was commonly utilized as signal sequence by several of our signal sequence trap (SST) isolated *S. japonicum* cDNAs (Table 1) [47]. The sequence of this signal sequence was employed as query to search for matches in the GenBank non-redundant nucleotide sequence database and expressed sequence tags (ESTs) database for all organisms using BLASTN program in National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) [71]. A search on the NCBI UniGene database [72] that provides information on sets of transcript sequences that appear to come from the same transcription locus was performed to ascertain redundancy and group identified transcripts. For genome-wide searches, the same query sequence and program were used to search the WGS reads from all organisms with sequenced genomes deposited in the NCBI genome databases. In a similar search in the protein database, the amino acid and nucleotide sequence of the same signal sequence was used as query for BLASTP and BLASTX searches respectively. Conserved domain architecture searches on all translation products of the SST identified candidate genes were performed using the conserved domain architecture retrieval tool on NCBI website [73] and compared with same analyses on the ProDom database of protein domain families available online at [74].

Multiple alignments

All multiple sequence alignments of DNA and protein sequences were performed in parallel with ClustalW on MegAlign program in Lasergene 7 DNASTAR software, NCBI bl2seq, COBALT multiple alignment programs, and *Multialin* interface software [75]. cDNA-to-genome sequence alignments were computed using the free NCBI *Splign* program [75]. The latest update of the *S. japonicum* genome map is accessible at [52]. Phylogenetic and molecular evolutionary analyses were conducted using *MEGA* version 5 [76].

Gene prediction

Gene predictions were performed using the GeneQuest program (Lasergene 7 DNASTAR) to predict potential coding regions, starts, stops, acceptors and donor sites using Borodovsky matrix files for *Caenorhabditis elegans*; and the results compared with that of the Eukaryotic GeneMark.hmm [66] gene prediction server provided freely on the website of Georgia Institute of Technology, Atlanta, USA.

Repeat masking

The whole sequences of all the genome contigs bearing the similar signal sequence were screened against a reference collection of repetitive DNA elements in the RepBase database available at the Genomic Information Research Institute website, using the CENSOR repeat masking software [77]. Sequence analysis figures were generated using real DNA sequences on Vector NTI Advanced 11.0 (Invitrogen).

Designation of putative duplication source locus and probable breakpoint

Reference to a parent gene is required for accurate determination of duplication breakpoint. However, in absence of a reference homolog, we putatively selected the most prominent contig [GenBank:CABF01020060], the longest among the identified dataset (43.7 kb), which significantly covered the length of the other contigs (Figure 3, also see Additional file 5) as the putative duplication source locus and utilized it as such for most of the analyses performed in this study. When the contigs were aligned with the putative source locus, homology was not lost till the 3' end of the aligned sequences. We therefore recruited two contigs [GenBank:CABF01020061 and GenBank:CABF01020062] downstream of the source locus based on genome assembly information, thereby generating at least 5 kb flanking sequences on either side of the duplication source locus. This sequence was then aligned with the genome contigs and scaffolds to identify the exact point where sequence identity disappeared. This point was arguably chosen as the possible duplication breakpoint and utilized as such in our discussions. We further attempted to identify a recurrent consensus sequence at the breakpoints but this was hampered by several sequencing gaps in the partially assembled scaffolds.

Parasites, genomic DNA and developmental stage mRNA samples

Chinese strain of *S. japonicum* (hereafter abbreviated as *Sj*) was obtained from Jiangsu Provincial Institute of Parasitic Diseases Wuxi, Jiangsu Province, PR China, while the Philippine and Japanese strains of *S. japonicum* in addition to *S. mekongi* (*Smk*) samples, were maintained in the Laboratory of Tropical Medicine and Parasitology, Dokkyo Medical University, Tochigi, Japan. *S. mansoni* (*Sm*) adult worms were maintained by, and kindly provided by the Department of Parasitology, Institute of Tropical Medicine, Nagasaki University, Japan. *S. haematobium* (*Sh*) sample was from Department of Immunology and Parasitology, University of Occupational and Environmental Health, Kitakyushu, Japan. Total genomic DNA was purified from cut tissues of mixed sex adult worms from different species of

Schistosoma using QIAamp DNA Mini Kit (QIAGEN) according to the manufacturer's instructions. Qualification and quantification of genomic DNA extract was assessed by gel electrophoresis and ND-1000 spectrophotometer (NanoDrop, USA). To obtain sufficient amount of genomic DNA for southern hybridization experiments, the whole genome of each sample was amplified using the GenomePhi DNA Amplification Kit (GE Healthcare) according to the manufacturer's instructions. Equally, total RNA was extracted from parasite eggs, cercariae, 24 h cultured schistosomulae and adult worms of *S. japonicum* according to the instruction manual of PureLink Micro-to-midi total RNA Purification System Kit (Invitrogen).

Reverse transcription polymerase chain reaction (RT-PCR)

mRNA from eggs, cercariae, 24 h culture schistosomulae and adult worms of the Chinese strain of *S. japonicum* was used for RT-PCR. The first strand cDNA was synthesized from the total RNA of each developmental stage by using oligo (dT) primer according to the instruction manual of High Capacity cDNA Reverse Transcription Kit (Applied Biosystems) and the resulting cDNA was used as template for RT-PCR. The *S. japonicum* actin gene was used for internal quality assurance. The cDNA sequences of some selected SST identified secreted candidate genes were amplified using pairs of sequence specific primers designed according to the *S. japonicum* transcriptome data [49] in the NCBI public database. All RT-PCR amplicons were analyzed using gel electrophoresis and confirmed by sequencing using the BigDye Terminator v1.1 Cycle Sequencing Kit (Applied Biosystem).

Southern hybridization

Southern hybridization was performed following standard procedures [78] using the DIG nonradioactive labeling and detection system (Roche, Germany). Briefly, the hybridization probe labeled with DIG-dUTP was synthesized using PCR DIG synthesis kit (Roche, Germany) according to the manufacturer's instructions, and labeling was confirmed by size disparity with unlabeled amplicon as a result of slower migration in agarose gel due to digoxigenin labeling. Genomic DNA from different species of *Schistosoma* (*Sh*, *Sm*, *Smk*, *Sj* Japanese (Yamanashi), *Sj* Chinese (Jiangsu) and *Sj* Philippines (Leyte, Mindanao and Mondoro isolates) were double digested with three different pairs of restriction enzymes (*EcoRI* + *EcoRV*, *EcoRI* + *HindIII* and *BamHI* + *HindIII*) to achieve the best possible fragmentation of the genomic DNA. The digested genomic DNA fragments were electrophoresed through 1 % (w/v) agarose gel, depurinated in 250 mM HCl, and denatured by incubating

in two changes of denaturing solution for 15 min each (0.5 M NaOH, 1.5 M NaCl). The gels were then neutralized by incubation in two changes of neutralizing solution (0.5 M Tris-HCl at pH7.5, 1.5 M NaCl) for 15 min each, and DNA was transferred to a positively charged nylon membrane (Roche, Germany) by capillary action overnight using 20x SSC solution (3 M NaCl, 300 mM sodium citrate at pH 7.0). The transferred DNA was fixed to the membrane by baking in an oven at 80 °C for 2 hours after rinsing briefly in 2x SSC. After prehybridizing the membrane in 10 ml hybridization buffer (5x SSC, 0.1 % N-lauroylsarcosine (w/v), 0.02 % SDS (w/v), 1 % blocking solution (Roche, Germany)) for 30mins in a hybridization bag, 5µl of the PCR generated hybridization probe was mixed in 50µl of double deionized water, denatured by boiling for 5mins and introduced into the hybridization bag and incubated overnight with shaking at 50 °C. The membrane was washed in two changes of low stringent wash buffer (2x SSC, 0.1 % SDS) for 5mins each at RT, and twice in high stringent wash buffers (0.5x SSC, 0.1 % SDS) for 15 min each at 65 °C. The hybridized probe was then detected using anti-Digoxigenin antibody (Roche, Germany) using CSPD as the chemiluminiscent substrate according to the manufacturer's instructions. The blot was then visualized by exposing to chemiluminescence for 10 min in a LAS-4000 mini image reader (Fujifilm).

Additional files

Additional file 1: Schematics of some of the mechanisms of novel gene origination. Apart from the pioneering idea of gene duplication [6], there are other mechanisms by which new genes are born. These include but not limited to exon shuffling or exon "scrambling" (a) [4,14-18]; fission or fusion of genes (b) [1,3,22], horizontal gene transfer between organisms (c) [31-33], *de novo* origination of protein coding genes from previously non-coding sequences (d) [2,3,34-40], retrotransposition by retrotransposons yielding intronless chimeric genes (e) [18-25], transduction of adjacent DNA by transposable elements (f) or may involve a repetitive element mediated DNA level recombination by a non-allelic homologous recombination (NAHR) mechanism (g) [7,20,26-30]. The figure was adapted from [3].

Additional file 2: Multiple alignments of signal sequence trap (SST) isolated cDNAs showing similar signal sequence. The similar promoter region including the signal sequence is boxed. The two arrows indicate the 'ATG' start positions utilized in the transcript ORF of the candidate mRNA sequences.

Additional file 3: Phylogenetic tree of the genus *Schistosoma* showing the possible origination point of new duplicated genes. The species phylogeny was adapted from [53] as inferred from DNA sequencing, comparative molecular genomics and karyotyping. This phylogenetic tree was manually simulated and thus the length of the branches does not estimate dates or time scale. The tree shows the *S. japonicum* clade and a representative each of the other clades in the genus including the species that reinvaded Asia from Africa. See review in [53]. Based on the result of the southern hybridization in Figure 2, the species and strains that contain the duplicated genes encoding products with similar signal sequence are colored green and we inferred that the most probable time point estimate (black dot) of the gene's emergence could be after the other species in the *S. japonicum* group (in parenthesis) have diverged.

Additional file 4: Expected fragments on restriction map of genome scaffolds correspond to bands on southern blots. To confirm dispersed duplication hypothesis and to exclude the possibility of overlapping among the loci, the restriction map of six of the genome scaffolds bearing duplicated loci were generated (A). Using same restriction endonuclease enzymes as in the generated maps, we performed southern hybridization using restriction digested genomic DNA from *S. japonicum* species and strains, and were able to match the expected fragment sizes with the observed bands on the hybridization blots. The contigs and the expected probe binding sites were labeled followed by their sequence ranges. We denoted the respective restriction digested fragments with probe binding site using alphabets with their expected restriction digestion product sizes in parenthesis (*E + E: EcoRI + EcoRV; E + H: EcoRI + HindIII; B + H = BamHI + HindIII*). As shown in (B), we were able to match the expected fragment with the southern blot bands, labeled using their corresponding alphabetic codes. Probe binding site on the positive strand were colored 'green' while the antisense sites were colored 'red'. The tiny vertical lines on the graphics represent the cutting sites of the selected restriction enzymes. The restriction map and the image were generated using DNADynamo sequence analysis software.

Additional file 5: Simulations using our raw data to show DNA-Level recombination mediated by REs by NAHR mechanism. The movie created from a Powerpoint presentation (Additional file 10) represents the basic approach we utilized in our analysis to show evidence of DNA level recombination by a non-allelic homologous recombination mechanism. The raw data obtained from BLAST searches and RepBase repetitive element prediction report was used to present a simulation that demonstrates that the duplicated locus is flanked on 5' and 3' ends by retrotransposons of the classes *RTE_SJ* and *Perere* respectively. We proposed that these repetitive elements could have provided the requisite homologous stretch of DNA that is required for such DNA level recombination. NAHR can be inter-chromosomal, intra-chromosomal, inter-sister chromatid, or intra-chromatid to give rise to disperse duplicates of the intervening genomic locus. This movie was created from an original Powerpoint presentation (Additional file 10)

Additional file 6: A simplified illustration of repetitive element mediated DNA level non-allelic homologous recombination (NAHR). Repetitive elements provide the requisite homologous DNA sequence for DNA level recombination between non-allelic pairs by a NAHR mechanism. NAHR can occur within a chromosome (intra-homologous chromosomal), between chromosomes (inter-chromosomal), between sister-chromatids or within a chromatid to give rise to disperse duplicates of the intervening genomic locus. The figure was adapted from [60]. Also see Additional file 1 for a cartoon of NAHR and other mechanisms of new gene origination, and [20,26-29,60] for review.

Additional file 7: Splicing models of some protein-coding representatives of the young duplicons. Based on gene prediction from the contigs using GeneQuest and GeneMark and alignment of cDNAs to genome sequences using *Splicing* program, we married the predicted products to the transcriptome database of this parasite and found that some of the duplicons are able to code for distinct gene products. Some of the transcription loci can encode two mRNA transcript variants. The significance of this was further explored in Figure 6.

Additional file 8: SjCP1084 protein coding mRNA [GenBank: AY570737] and a non-coding transcript [GenBank:FN328299] are products of alternative splicing Based on gene prediction from the contigs using GeneQuest and GeneMark, and alignment of cDNAs to genome sequences using *Splicing* program, we observed that two mRNA transcript variants were produced from [GenBank:CABF01020060]. This figure is same as Figure 6 (A) but we have in addition presented the aligned sequence of the two transcripts showing details of alternative splicing. An extra splice site was evolved in the first exon of the non-coding transcript [GenBank:FN328299]. When the splice site is recognized, an ORF encoding SjCP1084 protein coding mRNA [GenBank:AY570737] variant is created.

Additional file 9: SjCP1531 protein coding mRNA [GenBank: AY570742] and a non-coding transcript [GenBank:FN329677] are products of alternative splicing. Based on gene prediction from the

contigs using GeneQuest and GeneMark, and alignment of cDNAs to genome sequences using *Splicing* program, we observed that two mRNA transcript variants were produced from [GenBank:CABF01023364]. This figure is same as Figure 6 (B) but we have in addition provided the aligned sequence of the two transcripts showing details of alternative splicing. Two extra splice sites were not utilized in the transcription of the non-coding transcript [GenBank:FN329677]. When the splice sites were recognized, exons 5 and 6 of a translatable ORF were created to produce SjCP1531 protein coding mRNA [GenBank:AY570742] variant. Refer to RT-PCR result in Figure 5 (C) where two bands (exact size and sequence as the two variants described above) are seen on the agarose gel electrophoresis image.

Additional file 10: Simulations using our raw data to show DNA-Level recombination mediated by REs by NAHR mechanism. This Powerpoint presentation was used to create the movie in Additional file 5.

Abbreviations

SST: Signal sequence trap; NAHR: Non-allelic homologous recombination; DLR: DNA level recombination; NHEJ: Non-homologous end joining; IR: Illegitimate recombination; RE: Repetitive element; ORF: Open reading frame; WGS: Whole genome shotgun.

Competing interests

The authors declare that they have no competing interests.

Author contributions

ECM participated in the conception and design of the study, in-silico analyses, molecular experiments, data analysis and interpretation and drafted the manuscript. YC carried out the signal sequence trap (SST) and participated in in-silico analyses. MK participated in the design of the study, SST, in-silico analyses, molecular experiments and data interpretation. MNS participated in in-silico analyses, molecular experiments, data interpretation and revised the manuscript. DB participated in molecular experiments and data analyses. MK₂, NH, YC₂ and YO maintained parasite life cycle and participated in molecular experiments. SH participated in data interpretation, supervision and revised manuscript for intellectual content. KH participated in the conception and design of the study, SST, in-silico analyses, data interpretation, revised the manuscript and general coordination. All authors approved final version of the manuscript.

Acknowledgements

We would like to thank Prof. Kenji Hirayama's lab members (Department of Immunogenetics, Institute of Tropical Medicine, Nagasaki University) for insightful discussions. ECM is a recipient of the Japanese Government Ministry of Education, Culture, Sports, Science and Technology (MEXT) PhD fellowship. This study was supported in part by the Global Center of Excellence (GCOE) Program (2008-2011); Grant-in-Aid for 21st century COE program (2003-2008), Nagasaki University; and Grant-in-Aid for Scientific Research B (22406009) and C (23590489) from the Japanese Government Ministry of Education, Culture, Sports, Science and Technology (MEXT). The funding agency played no role in conducting the study, drafting the manuscript and the decision to publish.

Author details

¹Department of Immunogenetics, Institute of Tropical Medicine (NEKKEN), and Global COE Program, Nagasaki University, 1-12-4 Sakamoto, 852-8523, Nagasaki, Japan. ²Laboratory on Technology for Parasitic Disease Prevention and Control, Jiangsu Institute of Parasitic Diseases, 117 Yangxiang, Meiyuan, Wuxi 214064, People's Republic of China. ³Laboratory of Tropical Medicine and Parasitology, Dokkyo Medical University, Tochigi, Japan. ⁴Department of Immunology and Parasitology, The University of Occupational and Environmental Health, Kitakyushu, Japan. ⁵Department of Parasitology, Institute of Tropical Medicine (NEKKEN), and Global COE Program, Nagasaki University, 1-12-4 Sakamoto, 852-8523, Nagasaki, Japan. ⁶Department of Parasitology and Entomology, Faculty of Bioscience, Nnamdi Azikiwe University, P.M.B. 5025, Awka, Nigeria.

Received: 2 February 2012 Accepted: 11 June 2012

Published: 20 June 2012

References

- Zhou Q, Wang W: On the origin and evolution of new genes—a genomic and experimental perspective. *Journal of Genetics and Genomics* 2008, 35(11):639–648.
- Cai J, Zhao R, Jiang H, Wang W: De novo origination of a new protein-coding gene in *Saccharomyces cerevisiae*. *Genetics* 2008, 179(1):487–496.
- Long M, Betran E, Thornton K, Wang W: The origin of new genes: glimpses from the young and old. *Nat Rev Genet* 2003, 4(11):865–875.
- Long M, Deutsch M, Wang W, Betran E, Brunet FG, Zhang J: Origin of New Genes: Evidence from Experimental and Computational Analyses. *Genetica* 2003, 118(2):171–182.
- Kaessmann H: Origins, evolution, and phenotypic impact of new genes. *Genome Res* 2010, 20(10):1313–1326.
- Ohno S: *Evolution by gene duplication*. New York: Springer-Verlag; 1970.
- Bailey JA, Liu G, Eichler EE: An Alu transposition model for the origin and expansion of human segmental duplications. *Am J Hum Genet* 2003, 73(4):823–834.
- Gao X, Lynch M: Ubiquitous internal gene duplication and intron creation in eukaryotes. *Proc Natl Acad Sci* 2009, 106(49):20818–20823.
- Katju V, Lynch M: On the formation of novel genes by duplication in the *Caenorhabditis elegans* genome. *Mol Biol Evol* 2006, 23(5):1056–1067.
- Gu Z, Rifkin SA, White KP, Li W-H: Duplicate genes increase gene expression diversity within and between species. *Nat Genet* 2004, 36(6):577–579.
- Innan H, Kondrashov F: The evolution of gene duplications: classifying and distinguishing between models. *Nat Rev Genet* 2010, 11(2):97–108.
- Osada N, Innan H: Duplication and gene conversion in the *Drosophila melanogaster* genome. *PLoS genetics* 2008, 4(12):e1000305.
- Zhou Q, Zhang G, Zhang Y, Xu S, Zhao R, Zhan Z, Li X, Ding Y, Yang S, Wang W: On the origin of new genes in *Drosophila*. *Genome Res* 2008, 18(9):1446–1455.
- Long M, Langley CH: Natural selection and the origin of jingwei, a chimeric processed functional gene in *Drosophila*. *Science* 1993, 260(5104):91–95.
- Moran JV, DeBerardinis RJ, Kazazian HH Jr: Exon shuffling by L1 retrotransposition. *Science* 1999, 283(5407):1530–1534.
- Patthy L: Genome evolution and the evolution of exon-shuffling—a review. *Gene* 1999, 238(1):103–114.
- Shao X, Shepelev V, Fedorov A: Bioinformatic analysis of exon repetition, exon scrambling and trans-splicing in humans. *Bioinformatics* 2006, 22(6):692–698.
- Elrouby N, Bureau TE: Bs1, a New Chimeric Gene Formed by Retrotransposon-Mediated Exon Shuffling in Maize. *Plant Physiol* 2010, 153(3):1413–1424.
- Wang W, Brunet FG, Nevo E, Long M: Origin of sphinx, a young chimeric RNA gene in *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* 2002, 99(7):4448–4453.
- Cordaux R, Batzer MA: The impact of retrotransposons on human genome evolution. *Nat Rev Genet* 2009, 10(10):691–703.
- Courseaux A, Nahon JL: Birth of two chimeric genes in the Hominidae lineage. *Science* 2001, 291(5507):1293–1297.
- Kaessmann H, Vinckenbosch N, Long M: RNA-based gene duplication: mechanistic and evolutionary insights. *Nat Rev Genet* 2009, 10(1):19–31.
- Shapiro J: Transposable elements as the key to a 21st century view of evolution. *Genetica* 1999, 107(1):171–179.
- Okamura K, Nakai K: Retrotransposition as a Source of New Promoters. *Mol Biol Evol* 2008, 25(6):1231–1238.
- Hancks DC, Goodier JL, Mandal PK, Cheung LE, Kazazian HH Jr: Retrotransposition of marked SVA elements by human L1s in cultured cells. *Hum Mol Genet* 2011, 20(17):3386–3400.
- Johnson ME, Cheng Z, Morrison VA, Scherer S, Ventura M, Gibbs RA, Green ED, Eichler EE: Recurrent duplication-driven transposition of DNA during hominoid evolution. *Proc Natl Acad Sci U S A* 2006, 103(47):17626–17631.
- Yang S, Arguello JR, Li X, Ding Y, Zhou Q, Chen Y, Zhang Y, Zhao R, Brunet F, Peng L, et al: Repetitive Element-Mediated Recombination as a Mechanism for New Gene Origination in *Drosophila*. *PLoS genetics* 2008, 4(1):e3.
- Shapiro JA: A 21st century view of evolution: genome system architecture, repetitive DNA, and natural genetic engineering. *Gene* 2005, 345(1):91–100.
- Xing J, Wang H, Belancio VP, Cordaux R, Deininger PL, Batzer MA: Emergence of primate genes by retrotransposon-mediated sequence transduction. *Proc Natl Acad Sci* 2006, 103(47):17608–17613.
- Kidd JM, Cooper GM, Donahue WF, Hayden HS, Sampas N, Graves T, Hansen N, Teague B, Alkan C, Antonacci F, et al: Mapping and sequencing of structural variation from eight human genomes. *Nature* 2008, 453(7191):56–64.
- Eugene V, Koonin KSM, Aravind L: Horizontal Gene Transfer in Prokaryotes: Quantification and Classification. *Annu Rev Microbiol* 2001, 55:709–742.
- Hotopp JCD, Clark ME, Oliveira DCSG, Foster JM, Fischer P, Torres MCM, Giebel JD, Kumar N, Ishmael N, Wang S, et al: Widespread Lateral Gene Transfer from Intracellular Bacteria to Multicellular Eukaryotes. *Science* 2007, 317(5845):1753–1756.
- Keeling PJ, Palmer JD: Horizontal gene transfer in eukaryotic evolution. *Nat Rev Genet* 2008, 9(8):605–618.
- Yang Z, Huang J: De novo origin of new genes with introns in *Plasmodium vivax*. *FEBS Lett* 2011, 585(4):641–644.
- Levine MT, Jones CD, Kern AD, Lindfors HA, Begun DJ: Novel genes derived from noncoding DNA in *Drosophila melanogaster* are frequently X-linked and exhibit testis-biased expression. *Proc Natl Acad Sci* 2006, 103(26):9935–9939.
- Volff J-N: Turning junk into gold: domestication of transposable elements and the creation of new genes in eukaryotes. *BioEssays* 2006, 28(9):913–922.
- Begun DJ, Lindfors HA, Kern AD, Jones CD: Evidence for de Novo Evolution of Testis-Expressed Genes in the *Drosophila yakuba*/*Drosophila erecta* Clade. *Genetics* 2007, 176(2):1131–1137.
- Knowles DG, McLysaght A: Recent de novo origin of human protein-coding genes. *Genome Res* 2009, 19(10):1752–1759.
- Wu DD, Irwin DM, Zhang YP: De novo origin of human protein-coding genes. *PLoS genetics* 2011, 7(11):e1002379.
- Guerzoni D, McLysaght A: De novo origins of human genes. *PLoS genetics* 2011, 7(11):e1002381.
- Snel B, Bork P, Huynen M: Genome evolution. Gene fusion versus gene fission. *Trends Genet* 2000, 16(1):9–11.
- Wilson MS, Mentink-Kane MM, Pesce JT, Ramalingam TR, Thompson R, Wynn TA: Immunopathology of schistosomiasis. *Immunol Cell Biol* 2006, 85(2):148–154.
- Cass CL, Johnson JR, Califf LL, Xu T, Hernandez HJ, Stadecker MJ, Yates JR III, Williams DL: Proteomic analysis of *Schistosoma mansoni* egg secretions. *Mol Biochem Parasitol* 2007, 155(2):84–93.
- Edungbola LD, Cha YN, Bueding E, Schiller EL: Granuloma formation around exogenous eggs of *Schistosoma mansoni* and *Schistosoma japonicum* in mice. *Afr J Med Med Sci* 1982, 11(2):75–79.
- Burke ML, Jones MK, Gobert GN, Li YS, Ellis MK, McManus DP: Immunopathogenesis of human schistosomiasis. *Parasite Immunology* 2009, 31(4):163–176.
- McManus DP, Loukas A: Current Status of Vaccines for Schistosomiasis. *Clin Microbiol Rev* 2008, 21(1):225–242.
- Yu C, Zhang F, Yin X, Kikuchi M, Hirayama K: Isolation of the cDNAs encoding secreted and membrane binding proteins from egg of *Schistosoma japonicum* (Chinese strain). *Acta Parasitologica* 2008, 53(1):110–114.
- Spanu PD, Abbott JC, Amselem J, Burgis TA, Soanes DM, Stüber K, Loren van Themaat EV, Brown JKM, Butcher SA, Gurr SJ, et al: Genome Expansion and Gene Loss in Powdery Mildew Fungi Reveal Tradeoffs in Extreme Parasitism. *Science* 2010, 330(6010):1543–1546.
- Feng Liu YZ, Wang Zhi-Qin, Lu Gang, Zheng Huajun, Brindley Paul J, McManus Donald P, Blair David, Zhang Qin-hua, Zhong Yang, Wang Shengyue, Han Ze-Guang, Chen Zhu: The *Schistosoma japonicum* genome reveals features of host-parasite interplay. *Nature* 2009, 460:345–351 (16 July 2009).
- Hu W, Brindley PJ, McManus DP, Feng Z, Han Z-G: Schistosome transcriptomes: new insights into the parasite and schistosomiasis. *Trends in Molecular Medicine* 2004, 10(5):217–225.
- Hu W, Yan Q, Shen DK, Liu F, Zhu ZD, Song HD, Xu XR, Wang ZI, Rong YP, Zeng LC, et al: Evolutionary and biomedical implications of a *Schistosoma japonicum* complementary DNA resource. *Nat Genet* 2003, 35(2):139–147.
- Gene DB; [http://www.genedb.org/Homepage/Sjaponicum].
- Lawton SP, Hirai H, Ironside JE, Johnston DA, Rollinson D: Genomes and geography: genomic insights into the evolution and phylogeography of the genus *Schistosoma*. *Parasit Vectors* 2011, 4:131.
- Lockyer AE, Olson PD, Ostergaard P, Rollinson D, Johnston DA, Attwood SW, Southgate VR, Horak P, Snyder SD, Le TH, et al: The phylogeny of the Schistosomatidae based on three genes with emphasis on the interrelationships of *Schistosoma* Weinland, 1858. *Parasitology* 2003, 126(Pt 3):203–224.

55. Le TH, Humair PF, Blair D, Agatsuma T, Littlewood DT, McManus DP: Mitochondrial gene content, arrangement and composition compared in African and Asian schistosomes. *Mol Biochem Parasitol* 2001, **117**(1):61–71.
56. Chen ST, Cheng HC, Barbash DA, Yang HP: Evolution of hydra, a recently evolved testis-expressed gene with nine alternative first exons in *Drosophila melanogaster*. *PLoS genetics* 2007, **3**(7):e107.
57. Szak ST, Pickeral OK, Landsman D, Boeke JD: Identifying related L1 retrotransposons by analyzing 3' transduced sequences. *Genome Biol* 2003, **4**(5):R30.
58. Wang W, Zheng H, Fan C, Li J, Shi J, Cai Z, Zhang G, Liu D, Zhang J, Vang S, et al: High Rate of Chimeric Gene Origination by Retroinsertion in Plant Genomes. *The Plant Cell Online* 2006, **18**(8):1791–1802.
59. Thomas JH, Schneider S: Coevolution of retroelements and tandem zinc finger genes. *Genome Res* 2011, **21**(11):1800–1812.
60. Sasaki M, Lange J, Keeney S: Genome destabilization by homologous recombination in the germ line. *Nat Rev Mol Cell Biol* 2010, **11**(3):182–195.
61. Cordaux R: The human genome in the LINE of fire. *Proc Natl Acad Sci U S A* 2008, **105**(49):19033–19034.
62. Laha T, Brindley PJ, Smout MJ, Verity CK, McManus DP, Loukas A: Reverse transcriptase activity and untranslated region sharing of a new RTE-like, non-long terminal repeat retrotransposon from the human blood fluke, *Schistosoma japonicum*. *Int J Parasitol* 2002, **32**(9):1163–1174.
63. Zhang YE, Vrbancan MD, Krinsky BH, Long M: A cautionary note for retrocopy identification: DNA-based duplication of intron-containing genes significantly contributes to the origination of single exon genes. *Bioinformatics* 2011, **27**(13):1749–1753.
64. Lynch M, Force A: The probability of duplicate gene preservation by subfunctionalization. *Genetics* 2000, **154**(1):459–473.
65. Conant GC, Wolfe KH: Turning a hobby into a job: How duplicated genes find new functions. *Nat Rev Genet* 2008, **9**(12):938–950.
66. Lomsadze A, Ter-Hovhannisyan V, Chernoff YO, Borodovsky M: Gene identification in novel eukaryotic genomes by self-training algorithm. *Nucleic Acids Res* 2005, **33**(20):6494–6506.
67. Modrek B, Lee C: A genomic view of alternative splicing. *Nat Genet* 2002, **30**(1):13–19.
68. Keren H, Lev-Maor G, Ast G: Alternative splicing and evolution: diversification, exon definition and function. *Nat Rev Genet* 2010, **11**(5):345–355.
69. Szczesniak MW, Ciomborowska J, Nowak W, Rogozin IB, Makalowska I: Primate and rodent specific intron gains and the origin of retrogenes with splice variants. *Mol Biol Evol* 2011, **28**(1):33–37.
70. Mattick JS, Makunin IV: Non-coding RNA. *Hum Mol Genet* 2006, **15**(Spec No 1):R17–R29.
71. NCBI: *Basic Local Alignment Search Tool (BLAST)*: [<http://blast.ncbi.nlm.nih.gov/Blast.cgi>].
72. NCBI: *UniGene Database*: [[http://www.ncbi.nlm.nih.gov/unigene?term=txid6182\[organism\]](http://www.ncbi.nlm.nih.gov/unigene?term=txid6182[organism])].
73. NCBI: *Conserved Domain Architecture Retrieval Tool*: [<http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml>].
74. *ProDom Database*: [<http://prodom.prabi.fr/prodom/current/html/home.php>].
75. Corpet F: Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res* 1988, **16**(22):10881–10890.
76. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S: MEGA5: Molecular Evolutionary Genetics Analysis Using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol Evol* 2011, **28**(10):2731–2739.
77. Kohany O, Gentles AJ, Hankus L, Jurka J: Annotation, submission and screening of repetitive elements in Repbase: RepbaseSubmitter and Censor. *BMC Bioinforma* 2006, **7**:474.
78. Southern E: Southern blotting. *Nat Protoc* 2006, **1**(2):518–525.

doi:10.1186/1471-2164-13-260

Cite this article as: Mbanefo et al.: Origin of a novel protein-coding gene family with similar signal sequence in *Schistosoma japonicum*. *BMC Genomics* 2012 **13**:260.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit



トリパノソーマ症・リーシュマニア症

Trypanosomiasis and leishmaniasis

濱野 真二郎*

トリパノソーマ症もリーシュマニア症も単細胞真核生物である原虫によって引き起こされる寄生虫感染症である。いずれも「顧みられない熱帯病 (neglected tropical diseases : NTDs)」を代表する疾患であり、輸入感染症としても注意を要する。近年、アフリカトリパノソーマに対してニフルチモックス経口剤とエフロルニチン注射剤を組み合わせた併用療法 (NECT) が有効であることが判明した。わが国では東京大学の北グループがアスコフラノンというきわめて有望な抗トリパノソーマ薬を見出している。リーシュマニアに対してはミルテフォシン経口剤やアンホテリシン B のリポソーム製剤が著効するなど、一条の光が見えてきた。

Key Words : アフリカ睡眠病 (アフリカトリパノソーマ症)

シャーガス病 (アメリカトリパノソーマ症) / リーシュマニア症 / 昆虫媒介性疾患
顧みられない熱帯病 (neglected tropical diseases : NTDs)

I はじめに

熱帯は複雑なひとつの混沌であり、不可避的に病原体伝播の温床となってきた。特に貧困に喘ぐ熱帯地域ではその自然・社会環境と相まって、感染症が猛威を振るい続けている。そこには病原体と「疾病伝播の全構造」が無造作にある。熱帯感染症の多くは、患者の大部分が先進国ではなく途上国に生きる貧しい人々であるために顧みられることも少なく、「顧みられない熱帯病 (neglected tropical diseases : NTDs)」と呼ばれている。

トリパノソーマ症もリーシュマニア症も単細胞真核生物である原虫によって引き起こされる寄生虫感染症である。いずれも昆虫というベクターによって媒介される典型的な風土病であり、また、上述の NTDs を代表する疾患でもある。これら

NTDs は生命の脅威となるばかりでなく、多くは慢性の経過をたどり勤労意欲を失わせたり、障害を残したりするため、甚大な社会経済的損失を生み出し、地域や国が疲弊していく一因となる。貧困が NTDs の土壌となり、NTDs が貧困を増長する。まさに負の連鎖であり、そのような悪循環を断ち切るためにも国際的な寄生虫対策が喫緊の課題となっている。

II アフリカトリパノソーマ症 (睡眠病)

アフリカトリパノソーマ症はサハラ砂漠以南のアフリカ諸国で蔓延し、ツェツェバエ (tsetse fly) によって媒介される致死性の原虫感染症である。ヒトでは、*Trypanosoma brucei gambiense*, *T. b. rhodesiense* がアフリカ睡眠病 (眠り病) を引き起こし、家畜では、*T. congolense*, *T. vivax*, *T. b. brucei*

*長崎大学・熱帯医学研究所・寄生虫学分野 教授 Shinjiro Hamano

がナガナ病を引き起こす。この疾患の存在はアフリカに住むアラブ人やヨーロッパ人のあいだでは14世紀頃より、アフリカの眠り病、黒人の眠り病という名前で知られていた。睡眠・覚醒のリズムが乱れ、眠り続けて、最後は死亡する病気である。

19世紀の終わり、当時まだ暗黒大陸と呼ばれていたアフリカの植民地化が進み、ヒトや物資の移動が盛んになった。それとともに睡眠病も広がり、限られた地域に蔓延していた睡眠病がコンゴ中央部にまで広がった。1896～1906年にかけて、ヴィクトリア湖周辺では25万人以上が死亡し、コンゴではその倍以上のヒトが死亡するという大流行となった（コンゴとウガンダで80万人以上の方が亡くなったと推定されている）¹⁾。植民地時代はこの疾患をコントロールするためにあらゆる努力が傾注され、1960年代までに一旦沈静化したものの、独立後の1970年代から本症はふたたび流行し、1998年のピーク時には約4万人の新たな感染が報告された²⁾。この疾患は僻地で

蔓延する代表的なNTDsであり、実際にはその5倍以上の人々が感染し、無治療のまま死亡したと考えられる。

現在、年間およそ6,000万人が感染の危険に曝され、そのDALYs (disability adjusted life years) は135万人・年に上る。同症はBSE (牛海綿状脳症) や新型インフルエンザなどと同様、人獣共通感染症であり、ヒトのみならず、ナガナ病として家畜にも甚大な被害を与えてきた。結果として、1億5千万頭の牛が放牧可能なアフリカ大草原での畜産を阻まれており、経済的被害もきわめて深刻である¹⁾。

アフリカトリパノソーマがツェツェバエの吸血を通して宿主に感染すると、末梢血の環境に適応する中で全長約30 μ m前後の血流型へと変態する(写真)。血流型は宿主の感染防御機構から逃れるために細胞表面全体をvariant surface glycoprotein (VSG) という糖タンパクで覆い、2分裂をくり返して活発に増殖する。しかし、これらVSG

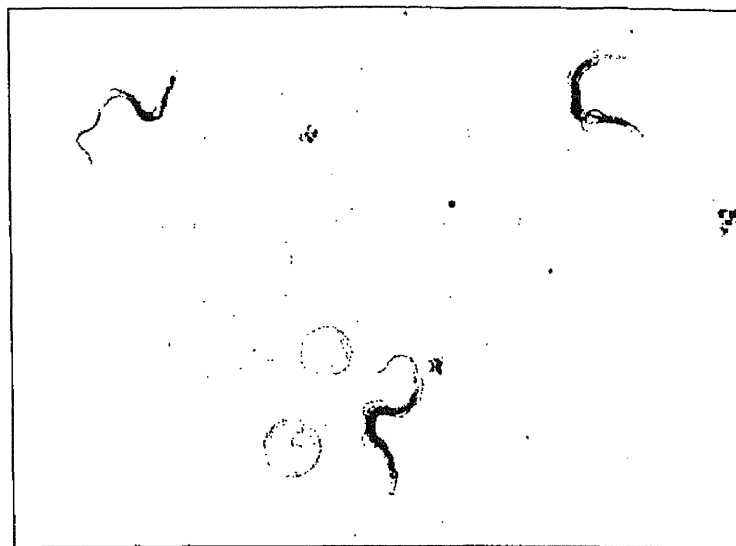


写真 血液中のアフリカトリパノソーマ

末梢血の薄層塗抹標本をギムザ染色したもの。血液中にアフリカトリパノソーマの血流型を認める。

(帯広畜産大学・原虫病研究センター：井上昇博士提供)

NTDs (neglected tropical diseases ; 顧みられない熱帯病)

DALYs (disability adjusted life years)

BSE (牛海綿状脳症)

VSG (variant surface glycoprotein)

はB細胞が産生する抗体の格好のターゲットとなり、たとえばAというVSGを発現する血流型が増殖すると、Aに対する抗体が産生されるようになる。すると、Aを発現する血流型は血中から急速に消失する。ところが、アフリカトリパノソーマは抗原性を異にする1,000にも及ぶVSG遺伝子を有しているため、A以外のVSGに覆われた血流型が増殖してくる。そのように発現するVSG遺伝子を順次切り替えることで宿主の特異的な免疫応答による攻撃を巧みに回避し、慢性感染へ移行する³⁾。

睡眠病の初期症状は間歇熱・頭痛・関節痛・リンパ節腫脹などであり、マラリアなど他の熱性疾患との鑑別が難しく、また、症状が自覚されない場合も多い(第1期)。原虫が血液脳関門を通過し中枢神経系へ侵入すると傾眠傾向から昏睡を経て死に至る(第2期)。たとえ第2期に治療したとしても不可逆的な神経障害が残るため、感染早期に確定診断を行い、治療を開始することが重要である。

医学・生命科学領域におけるトリパノソーマの研究成果は華々しいものであったが、宗主国から独立した後の疾病対策は21世紀に入るまで置き去りにされてきた。それは、一義的には貧しい人々の病気NTDsである故だが、そこから派生した詳細な理由を以下に記す。①感度と特異性に優れた安価で簡易な診断法がない(トリパノソーマカード凝集簡易試験は精度が低く、2ドル/検体と高価である)、②旧来の薬剤は激しい副作用、ときに死を引き起こす。したがって、③入院管理体制が不可欠であり、④高い再燃率のため治療後のフォローが必要であるものの、疾患は主として僻地で蔓延しており、インフラやアクセスが決定的に不足している。現在のところ、⑤有効なワクチンはなく、副作用の大きい薬剤に頼らざるを得ない。WHOを中心とした国際活動はそのような旧態依然とした診断・治療法と移動医療ケアチームなどを総動員して、ある程度の成果を収めたが、

僻地までは手が届かず、度重なる紛争は流行のリスクを増大している。ナガナ病もまた手つかずの状態である⁴⁾。

診断のゴールドスタンダードは血中にアフリカトリパノソーマを同定することにある。しかしながら、顕微鏡による検出感度が低いため感染初期における診断はきわめて困難である。Loop-mediated isothermal amplification (LAMP) 法は特殊機器を必要としないわが国発の遺伝子増幅法である。栄研化学(株)はすでにLAMPを用いたトリパノソーマの高感度検出に成功しており、感度・特異性ともに高く、都市部から遠い僻地でも使用可能な、安価・安定で、かつ簡便な診断法として期待されている。

寄生虫は人間と同じ真核生物であり、原核生物である細菌を標的にした抗生物質は通常無効である。また、寄生虫に対する既存の薬剤には薬剤耐性病原体の出現や重篤な副作用など憂慮すべき点も多く、これは抗アフリカトリパノソーマ薬にも当てはまる。第1期に使用されるスラミンでは、発熱、発疹、消化器症状などが認められ、第2期に使用されるメラルソプロールは、毒性が強く、脳症など激しい副作用を引き起こし、治療を受けた患者の5%が死亡する。近年、ニフルチモックス経口剤とエフロルニチン注射剤を組み合わせた併用療法(NECT)が有効で比較的安全であることが判明した⁵⁾。

2009年、医薬品メーカーのサノフィ・アベンティス社がエフロルニチンを、バイエル社がニフルチモックスを、それぞれWHOに寄付するという福音が届き、進行期(第2期)にあるアフリカ睡眠病の新しい治療選択肢となった。しかしながら、副作用や頻回の静脈内注射など多くの問題が残り、医療スタッフやインフラが決定的に不足する僻地での使用を阻んでいる。

わが国では東京大学の北グループがアスコフラノンというきわめて有望な抗トリパノソーマ薬(リード化合物)を見出している。この薬剤はトリ

LAMP (loop-mediated isothermal amplification)

TAO (trypanosome alternative oxidase ; シアン耐性酸化酵素)

パノソーマ特異的なシアン耐性酸化酵素 (trypanosome alternative oxidase: TAO) をきわめて低濃度 (nM オーダー) で選択的に阻害し原虫を殺滅するために、副作用が最小限に留まる。また、単回投与 30 分後にほとんどの原虫が血中から消失するため医療サポートも最小限で済み、僻地での使用が可能となる¹⁾。抗原虫薬と同様、ワクチン開発も重要であり、その基盤となる有望な研究が帯広畜産大学の井上を中心に進められている³⁾。

III シャーガス病 (アメリカトリパノソーマ症)

シャーガス病 (Chagas' disease) はクルーズトリパノソーマ (*Trypanosoma cruzi*) という鞭毛をもつ原虫によって引き起こされる人獣共通感染症である。中南米を中心に蔓延し、アメリカトリパノソーマ症とも呼ばれる。感染は吸血性の昆虫であるサシガメ (triatomine bugs, kissing bugs) によって媒介され、ヒトに加えて、イヌ、ネコ、トリなどの家畜、アルマジロ、オポッサムなどの野生動物も感染し保虫宿主となる。慢性感染では、心臓伝導障害や不整脈、心筋症、心不全や巨大消化管 (食道・結腸) 症など多彩な病態が引き起こされる。

中南米ではおよそ 1,000 ~ 1,500 万人の人々が同原虫に感染しており、同地域に蔓延する NTDs の中でも死亡率や DALYs を押し上げる最大の要因となっている⁶⁾。ベクターであるサシガメは藪や土壁でできた家屋の屋根や壁の隙間に生存する。南米でも、アルゼンチン、チリ、ウルグアイなど経済的に豊かな国においては、家屋の改善やサシガメのコントロールを通してトリパノソーマの感染率は激減した。一方、経済的に立ち後れたボリビアなどの国ではいまだに高い感染率が認められる。「貧しい人々の病」と言われる所以である。

世界のグローバル化にともない、米国においてもテキサス州やメキシコ湾岸を中心に、およそ 30 ~ 100 万人の感染者が存在すると推定されており、これらの人々が輸血や臓器提供を介して感染源となる可能性がある⁶⁾。胎盤を介した母子感

染も報告されている。また、カナダ・ヨーロッパ・オーストラリアにも数千人の感染者がいると考えられている⁶⁾。わが国においても中南米からの出稼ぎ者や青年海外協力隊員などに感染が見つかった⁵⁾。

感染したサシガメは吸血時に終末錐鞭毛型を含む糞を皮膚上に排泄する。この終末錐鞭毛型は皮膚の刺咬傷や粘膜から宿主内へ侵入する。クルーズトリパノソーマは、脾臓や肝臓などの網内系、心筋、神経系の細胞を中心にあらゆる有核細胞に感染し、細胞内に寄生することで免疫系の直接的な攻撃から逃れる。細胞の中では無鞭毛型に変態し、ファゴゾームから細胞質内に脱出し寄生・増殖する。細胞内で増殖した原虫は鞭毛を有する錐鞭毛型に変態し細胞外へ脱出する。それら錐鞭毛型は新たに次の細胞に侵入するか、もしくはサシガメによる吸血を経て生活環を一巡させる。細胞内に寄生する無鞭毛型の増殖を抑制・排除するためには、宿主は細胞性免疫を適切に活性化する必要がある。この急性期の症状は感染後 1 週間~数カ月間みられ、自然寛解することが多い。慢性感染に移行した患者の一部においては前述した重篤な心臓や消化管の症状が認められる。

診断は血中に存在する原虫を顕微鏡下に検出することによる。ヘマトクリット管で遠沈して buffy coat 層に動く錐鞭毛型を観察する方法や、厚層塗沫標本をギムザ染色して観察する方法がある。慢性期には検出感度も低下するため、各種 DNA 診断や未感染のサシガメに吸血させ、一定期間後に糞内もしくは腸内のトリパノソーマを検出する xenodiagnoses が用いられる。補助診断としては各種血清診断が用いられる。クルーズトリパノソーマに対する特效薬はない。急性期に効果があるものとしてニフルチモックス (Lampit[®]) が、急性期のみならず慢性期にも抗原虫作用を示すものとしてベンズニダゾール (Radanil[®]) が知られているが、現行の治療薬の治癒率は満足できるものではなく、しばしば深刻な副作用を呈する⁴⁾。このような理由から、新たな診断法・治療薬の開発が切望されている。

IV リーシュマニア症

リーシュマニア症は世界中に1,200万人以上の患者がいる寄生虫病である。その病型は多彩であるが大きく2つに分類される。未治療のまま放置すると死に至る内臓型リーシュマニア症（カラ・アザール）と、皮膚や粘膜の肉芽や潰瘍形成を引き起こす皮膚（粘膜）型リーシュマニア症である。皮膚（粘膜）型リーシュマニア症はたとうまく治療できても顔などに醜い瘢痕を残し、心にも大きな傷痕を残す場合がある。同症を引き起こす原虫はトリパノソーマ科リーシュマニア属 *Leishmania* spp. に属し、20種以上が報告されている。いずれもサシチョウバエ (sand fly) によって媒介され、旧世界では *Phlebotomus* 属、新世界では *Lutzomyia* 属が主たるベクターとなる。ヒトに加えて、イヌ、齧歯類、および他の脊椎動物も保有宿主となりうる人獣共通感染症である。輸入感染症としても注意を要する感染症のひとつである。

内臓型、そして皮膚（粘膜）型リーシュマニア症ともに風土病の色彩が濃いNTDsである。内臓型リーシュマニア症は、*L. donovani*, *L. infantum* や *L. chagasi* に起因し、その90%は、インド、バングラデシュ、ネパール、スーダン、エチオピア、ブラジルに分布している。肝脾腫、汎血球減少、るいそうなどをともなう熱性疾患であり、発症後治療せずに放置すると死に至る。内臓型の不十分な治療例では post-kala-azar-dermal leishmaniasis (PKDL) がみられる。治療の後、数カ月～数年後に皮膚にさまざまな皮疹が出現するものであり、その出現頻度は治療薬によっても大きく異なる。

皮膚（粘膜）型リーシュマニア症は、*L. major*, *L. tropica*, *L. mexicana*, *L. amazonensis* や *L. braziliensis* などに起因し、その90%は、アフガニスタン、アルジェリア、イラン、サウジアラビア、シ

リア、ブラジル、コロンビア、ペルー、ポリビアに認められる。サシチョウバエに刺された部位に無痛性の肉芽腫性病変が形成され、中心部が潰瘍化し、適切な免疫応答が誘導されると次第に治癒へ向かう。皮膚粘膜型はしばしば鼻口腔粘膜を侵し、鼻中隔や鼻翼の欠損など、ときに重篤な後遺症を残す。

サシチョウバエ雌の吸血に際してリーシュマニアがヒトや動物の体内に侵入すると、それは感染局所でマクロファージに貪食される。そこで無鞭毛型へと形態変化し、ファゴライソゾームの中でも生存し増殖を続ける。病原体に対する防御免疫にはさまざまな要素が関与するが、その中でも特にCD4陽性ヘルパーT細胞(Th)が重要な役割を果たす。抗原に未曝露のナイーブCD4陽性細胞が樹状細胞などの抗原提示細胞 (antigen presenting cells: APCs) 上の主要組織適合抗原 (major histocompatibility complex: MHC) 上に提示される病原体由来ペプチドを認識し活性化すると、Th1 および Th2 という2つの異なる性格をもつ細胞群に分化していく⁷⁾。この現象は当初、リーシュマニアの実験的感染モデルの研究から見出された知見であり、免疫学全般に大きなインパクトを与えた。マクロファージの中におけるリーシュマニアの増殖抑制の主体は主としてTh1細胞から産生されるサイトカインIFN- γ (インターフェロン γ) である。もしも何らかの原因で適切なTh1応答を誘導できない場合、感染マクロファージを十分に活性化することができずに原虫の増殖を許容し、結果的にリーシュマニア症を発症することとなる。

リーシュマニア症の確定診断は顕微鏡下にリーシュマニアを検出することによる。内臓型の場合は、血液、脾臓もしくは骨髄中に、皮膚（粘膜）型の場合は病変部の皮膚・粘膜の生検組織などに

PKDL (post-kala-azar-dermal leishmaniasis)

APCs (antigen presenting cells ; 抗原提示細胞)

MHC (major histocompatibility complex ; 主要組織適合抗原)

IFN- γ (インターフェロン γ)

PCR (polymerase chain reaction ; ポリメラーゼ連鎖反応)

ICT (immunochromatography test)

原虫を検出することによる。培養による検出もしばしば試みられる。近年はPCR (polymerase chain reaction: ポリメラーゼ連鎖反応) による診断、種の同定も行われている。また、補助診断としての血清学的検査や ICT (immunochromatography test) も多用される。リーシュマニア症の治療には長いあいだ、ヒ素製剤であるステイボ・グルコネートが使用されてきた。近年はミルテフォシン経口剤やアンホテリシンBリポソーム製剤(アムビゾーム®)が内臓型リーシュマニア症に著効することが判明し、流行地での第一選択薬となるなど一条の光が見えてきた。皮膚(粘膜)型への効果に関しても次第に知見が集まりつつある。

文 献

- 1) 山内一也, 北深: <眠り病>は眠らない—日本発! アフリカを救う新薬. 岩波科学ライブラリー 2008年.
- 2) Simarro PP, et al: Eliminating human African trypanosomiasis: where do we stand and what comes next? PLoS Med 5: e55, 2008.
- 3) 櫻井達也, 井上昇: アフリカトリパノソーマ症—獣医寄生虫学の挑戦—. 北海道獣医師会雑誌 55 (6): 211-217, 2011.
- 4) ヒューマンサイエンス振興財団政策創薬総合研究事業, 熱帯病治療薬研究班: 寄生虫症薬物治療の手引き 改訂第7.0版. 2010年.
- 5) Priotto G, Kasparian S, Ngouama D, et al: Multicenterclinicaltrial of nifurtimox-eflornithine combination therapy for second-stage sleeping sickness. XVII th International Congress for Tropical Medicine and Malaria Jeju, Korea, September 2008 (http://www.who.int/selection_medicines/committees/expert/17/application/PriottoNECT.pdf).
- 6) Hotez PJ, et al: Chagas disease: "the new HIV/AIDS of the Americas". PLoS Negl Trop Dis 6: e1498, 2012.
- 7) 濱野真二郎, 吉田裕樹: 寄生虫感染と免疫応答. 感染症 40: 205-211, 2010.

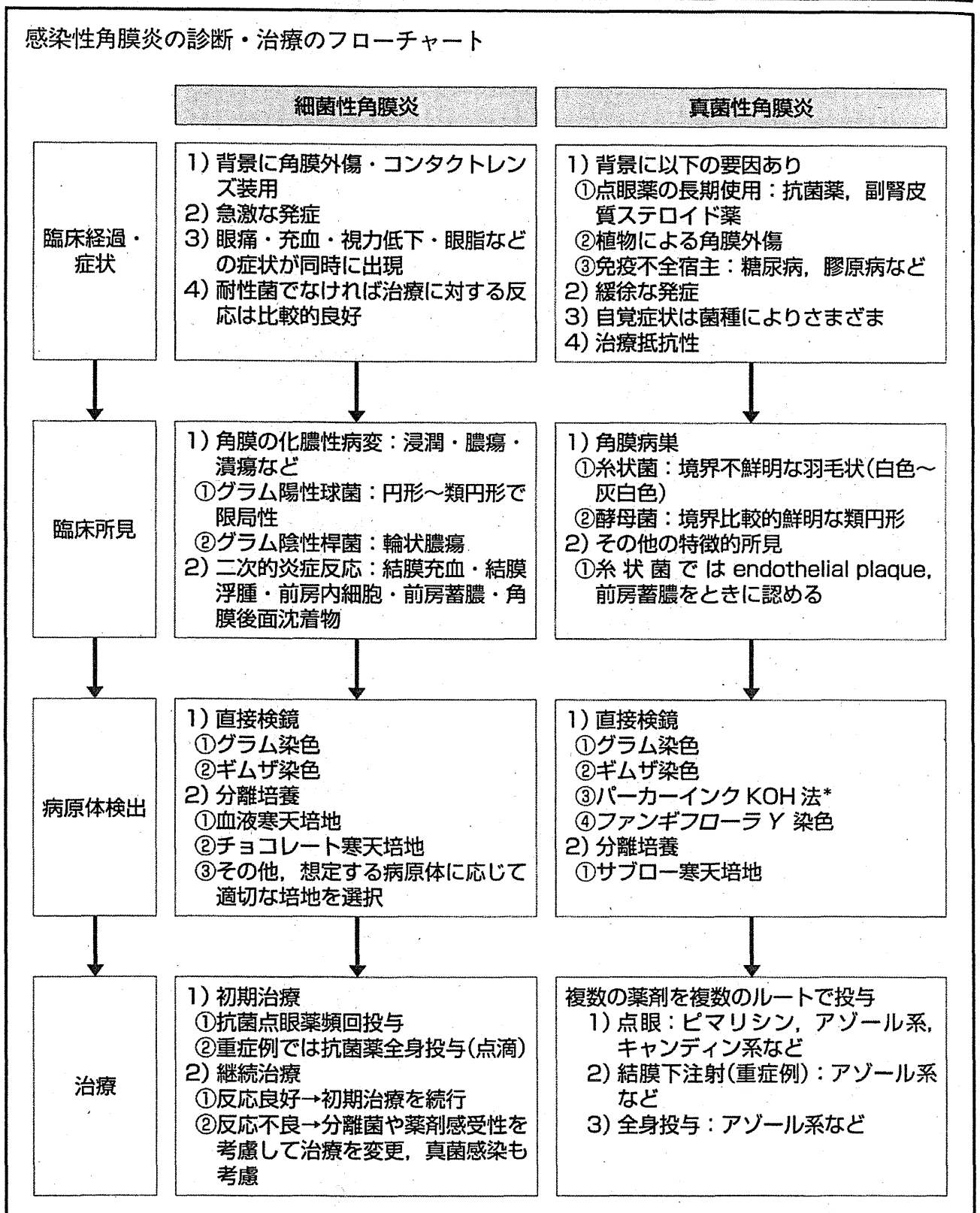
Ⅱ-5 角膜感染症

感染性角膜炎診療ガイドライン(2007)

鳥取大学医学部視覚病態学教授

井上 幸次

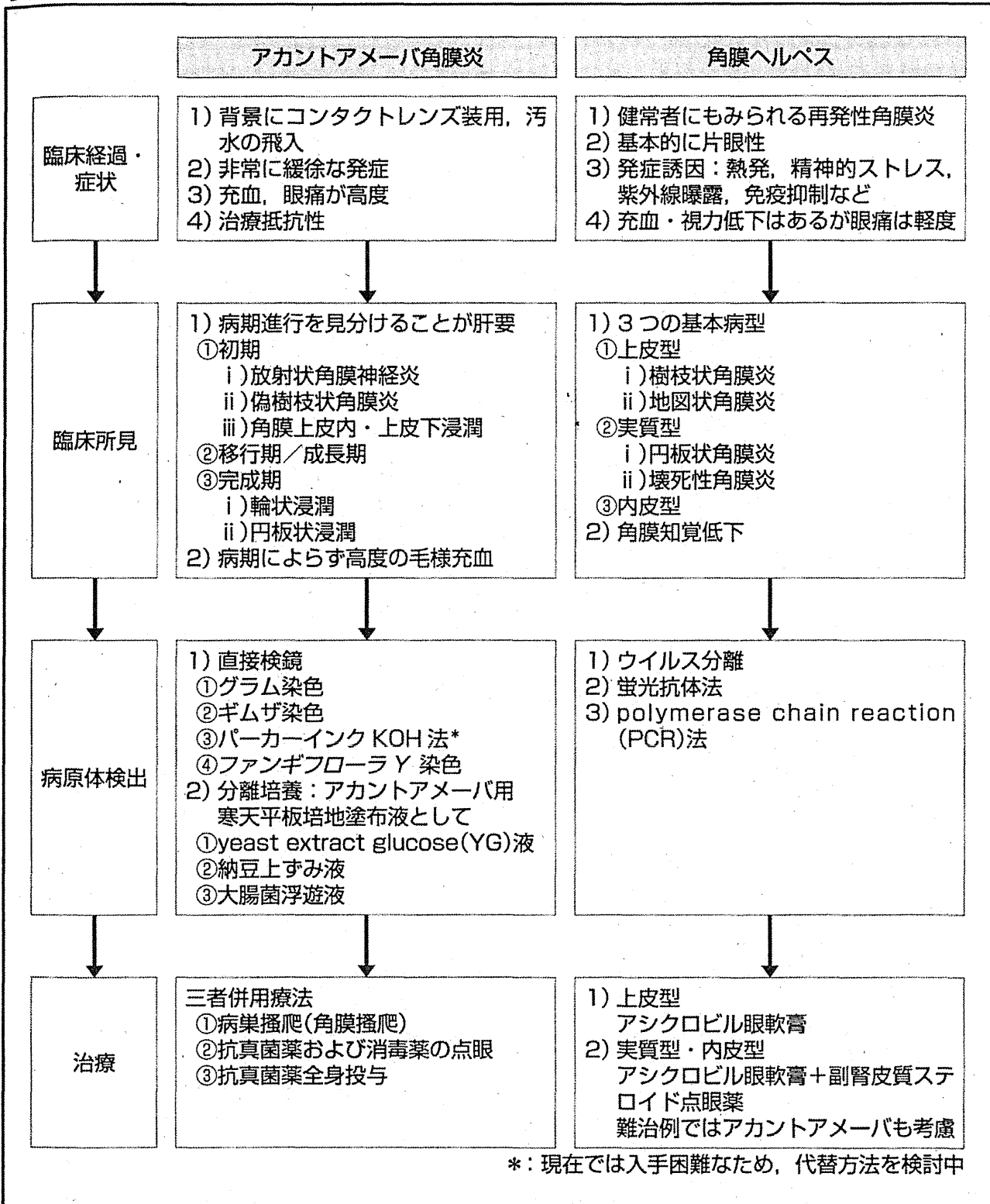
アルゴリズム



次頁につづく

つづき

眼科疾患



(日本眼感染症学会感染性角膜炎診療ガイドライン作成委員会 編: 日眼会誌 111, 770-809, 2007 より引用改変)

総説

感染性角膜炎診療ガイドラインの記載に合わせて、各疾患を分けずに概説する。各疾患別の解説は「アルゴリズム」と「具体的処方」を参照のこと。

定義

角膜に微生物が直接感染して増殖した病態。本来はそれに対してホストの免疫反応が生じて炎症を生じるのが「角膜炎」であるが、最近では医原性に免疫が抑制される状態や種々のバイオマテリアルの導入により、菌が増殖しているにもかかわらず、炎症反応を伴わない特殊な病態もみられるようになってきている。広義には、それも含めて感染性角膜炎と考えられる。

診断

◆病歴(問診)

角膜感染症においても問診が重要なのはいうまでもない。特に何が感染のきっかけになったかを知ることが重要である。最近では外傷よりも、コンタクトレンズ(CL)の重要性が増しており、その種類(商品名も含む)、装用日数・時間(誤用がなかったか)、CLの洗浄をしていたか、CLのこすり洗いをしてきたか、CLレンズケースの定期交換をしていたか、消毒の種類 [特にマルチ・パーパス・ソリューション(MPS)の商品名]、水道水の使用はなかったか、CL装用時に手洗いをしてきたか、などについて問診する。

◆臨床所見

①細隙灯顕微鏡所見

①上皮病変(樹枝状病変, 地図状病変, 星芒状病変)

- 1) 角膜に樹枝状病変をみた場合、まず、これが単純ヘルペスウイルス(HSV)による樹枝状角膜炎(dendritic keratitis)であるかどうかを見定めることが重要である。樹枝状角膜炎の特徴として、末端膨大部(terminal bulb)の存在、上皮内浸潤の存在、ある程度の幅があること、病変部以外の上皮は正常であることがポイントとなる。
- 2) 偽樹枝状病変を示す疾患として、眼部帯状疱疹、薬剤毒性角膜炎(epithelial crack line)、再発性角膜びらん(RCE)、アcantアメーバ角膜炎が重要である。
- 3) 地図状角膜炎は樹枝状角膜炎よりも鑑別が難しく、外傷などによる単純性角膜上皮欠損(単純性角膜びらん)、細菌・真菌感染に伴う角膜上皮欠損、遷延性角膜上皮欠損、栄養障害性角膜潰瘍、シールド潰瘍などとの鑑別が必要である。
- 4) 星芒状角膜炎は樹枝状角膜炎が非常に小規模で発症し、星形と表現したほ

うが合致する病変を呈したもので、眼部帯状疱疹に伴う星芒状角膜炎、Thygeson 点状表層角膜炎などと鑑別を要する。

②実質病変：感染性角膜炎を疑わせる実質病変に、浸潤、膿瘍、潰瘍がある。

- 1) 浸潤は角膜上皮あるいは実質に生じる好中球やリンパ球を主体とする細胞集積像の総称であり、角膜炎における代表的臨床所見の1つである。一般に、中央部に生じた場合は感染性、周辺部に生じた場合は非感染性のことが多い。
- 2) 膿瘍は角膜内に侵入した細菌や真菌に対して主として好中球が集簇したものである。炎症細胞内に含まれる蛋白分解酵素や活性酸素などにより組織破壊が生じる。治癒後には通常、組織の菲薄化が生じる。
- 3) 潰瘍は角膜上皮全層および実質に欠損が生じた状態をいい、多くは浸潤から発展する。典型的な感染症のパターンでは、好中球やリンパ球を主体とした炎症細胞の集積を角膜実質内に伴う。中央部の潰瘍は感染や神経麻痺に、周辺部の潰瘍は自己免疫疾患や感染アレルギーに起因することが多い。

③その他注意すべき所見

- 1) 感染の場合、角膜病巣に目を奪われがちだが、充血・前房内細胞・前房蓄膿・角膜後面沈着物・角膜浮腫・角膜穿孔などの副次的所見が細隙灯顕微鏡にて観察され、診断・治療の上で重要なヒントとなる。
- 2) 細隙灯顕微鏡検査を行う前に眼瞼浮腫・眼瞼発赤・眼脂・流涙などの肉眼所見にも注意を払う。

②角膜知覚検査

Cochet-Bonnet 型角膜知覚計を用いて、臨床の現場で簡単に行える。古典的な方法ではあるが、特に角膜ヘルペスの診断において必要不可欠である。

◆塗抹検鏡

- ①角膜感染症の診断において、塗抹検鏡は非常に重要である。ガイドラインに記載された方法が唯一の方法ではないので、その施設に応じたやり方を工夫していけばよいと思われる。たとえば、ガイドラインでは擦過に、Kimura spatula を推奨しているが、自分が使いやすい器具であればスパーテルでも、27ゲージ針でもよい。また、採取したサンプルの塗抹にあたって、サンプル量が比較的多ければ転がすように塗抹し、少なければスタンプを押すようにすると書かれているが、角膜炎の場合、多くはサンプル量が十分でなく、スタンプ法を選択せざるをえない。また、角膜炎ではサンプル量が少ないためにギムザ染色とグラム染色の両者を行うことは実際には難しいことが多いので、どちらか一方をとる場合はグラム染色を優先する。

- ②真菌については、パーカーインク KOH 法が基本であるが、現在パーカーインクは入手困難なため、ファンギフローラ Y 染色が役立つ。このほうが検鏡に不慣れな検者でも真菌と判定しやすい。

◆臨床検査

- ①細菌培養・感受性検査、真菌培養・感受性検査、アカントアメーバ培養、