Samia cynthia versus *Bombyx mori*: Comparative gene mapping between a species with a low-number karyotype and the model species of Lepidoptera.

Atsuo Yoshido[a*], Yuji Yasukochi[b], Ken Sahara[a]

[a]Laboratory of Applied Molecular Entomology, Graduate School of Agriculture, Hokkaido University, N9, W9, Kita-ku, Sapporo 060-8589, Japan

[b]Insect Genome Research Unit, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki 305-8634, Japan

Running title: Gene-based mapping in *Samia cynthia*

Key words: comparative gene mapping; fosmid-FISH; Lepidoptera; sex chromosome; conserved synteny.

[*]**corresponding author**: Atsuo Yoshido

Laboratory of Applied Molecular Entomology, Graduate School of Agriculture,

Hokkaido University, N9, W9, Kita-ku, Sapporo 060-8589, Japan

E-mail: yoshido@abs.agr.hokudai.ac.jp Phone & Fax: +81-11-706-2402

Abstract

We performed gene-based comparative FISH mapping between a wild silkmoth, *Samia cynthia* ssp. with a low number of chromosomes (2n=25-28) and the model species, *Bombyx mori* (2n=56), in order to identify the genomic components that make up the chromosomes in a low-number karyotype. Mapping of 64 fosmid probes containing orthologs of *B. mori* genes revealed that the homologues of either two or four *B. mori* chromosomes constitute the *S. c. ricini* (Vietnam population, 2n=27♀/28♂, Z0/ZZ) autosomes. Where tested, even the gene order was conserved between *S. c. ricini* and *B. mori*. This was also true for the originally autosomal parts of the neo-sex chromosomes in *S. c. walkeri* (Sapporo population, 2n=26♀/26♂, neo-Wneo-Z/neo-Zneo-Z) and *S. cynthia* subsp. indet. (Nagano population, 2n=25♀/26♂, neo-$WZ_1Z_2/Z_1Z_1Z_2Z_2$). The results are evidence for an internal stability of lepidopteran chromosomes even when all autosomes had undergone fusion processes to form a low-number karyotype.

1. Introduction

Genome rearrangements among species have been recently revealed by comparative genomic analyses. These showed that mammalian genomes have changed by reshuffling chromosomal segments from the common ancestral karyotype (Ferguson-Smith and Trifonov, 2007), while birds show a high degree of collinearity among species (Ellegren, 2010). In insects, comparative genome analysis among genome-sequenced 12 *Drosophila* species revealed that many orthologous genes mapped to the corresponding chromosomal arms but gene orders were scrambled between species (Ranz et al., 2001; Bhutkar et al., 2008). The availability of fully sequenced genomes provides an opportunity to study chromosomal rearrangements and evolutionary relationship among related species in detail.

Lepidoptera, moths and butterflies, consist of more than 150,000 species. They are the second largest order of animals (Kristensen and Skalski, 1999; Grimaldi and Engel, 2005). Lepidoptera have holokinetic chromosomes like aphids and bugs, and share the sex chromosome system of female heterogamety with caddis flies (Trichoptera) (Traut et al., 2007). The most common chromosome numbers of Lepidoptera range from n=28 to n=32 among the more than 1,000 species investigated

(Robinson, 1971). There are, however, also species with lower or higher number karyotypes in Lepidoptera. Species with low or high-number karyotypes are thought to have evolved by chromosomal fusion and fission from the putative ancestor with n=31 chromosome number (Lukhtanov, 2000; Marec et al., 2010).

Since the first lepidopteran genome, that of *Bombyx mori*, has been accessible in public databases, comparative mapping of genes was carried out against *B. mori* (n=28) by either linkage analysis (Beldade et al., 2009) or BAC-FISH (Yasukochi et al., 2009). These authors detected a high degree of conserved synteny and gene order between *B. mori* and two other species with the same chromosome number, *Bicyclus anynana* and *Manduca sexta*. Pringle et al. (2007) detected conserved gene order in *Heliconius melpomene* and suggested simple fusion events to account for the reduced chromosome number (n=21) of this species. These studies suggested an internal stability of lepidopteran chromosomes. DNA sequencing of each 15 selected BACs from *Helicoverpa armigera* and *Spodoptera frugiperda* also revealed a high degree of conserved synteny with only a few rearrangements between *B. mori* and the two noctuid moths, both of them having the supposed ancestral chromosome number (n=31) of Lepidoptera (d'Alençon et al., 2010).

Samia cynthia, a wild silkmoth, belongs to the family Saturniidae which is

closest to the family Bombycidae including *B. mori*. *S. cynthia* has a reduced chromosome number, ranging from 2n=25 to 2n=28 and is about half that of *Bombyx mori*. The variation is due to the variable sex chromosome constitution among geographic subspecies (Yoshido et al., 2005b). Previous study proposed repeated autosome-sex chromosome fusions resulted in the variable sex chromosome constitution as found in *S. cynthia* subspecies (Yoshido et al., 2010). The sex chromosome constitution in *S. c. walkeri* (the ailanthus silkworm, Sapporo population) females is designated as neo-W and neo-Z chromosomes, which originated by fusion of the ancestral W and Z with an autosome pair (A_1). Then sex chromosome constitution (designated as neo-WZ_1Z_2) in *S. cynthia* subsp. indet. (the shinju silkworm, Nagano population) females has been formed by next evolutionary step, in which neo-W chromosome fused with an autosome (A_2) and, consequently, its unfused homologue became a Z_2 chromosome. In *S. c. ricini* (the Eri silkworm, Vietnam population), no such fusion of sex chromosomes with autosomes occurred and, the sex chromosome constitution in females is Z0, which arose from ancestral WZ by a loss of the W chromosome. Hence, the chromosome number is 2n=27/28 in *S. c. ricini* with a Z0/ZZ, 2n=26/26 in *S. c. walkeri* with neo-Wneo-Z/neo-Zneo-Z, and 2n=25/26 in *S. cynthia* subsp. indet. with neo-WZ_1Z_2/$Z_1Z_1Z_2Z_2$ sex chromosomes in female/males.

We show here by comparative gene mapping between *S. cynthia* subspecies and *B. mori* the internal stability of lepidopteran chromosomes even when low-chromosome-number karyotypes evolve by chromosome fusion. We constructed a fosmid-library of *S. cynthia* and carried out gene-based comparative FISH mapping between *B. mori* and the three *S. cynthia* subspecies. Sixty-four fosmid probes which contain orthologs of *B. mori* genes, cytogenetically identified all chromosomes of the *S. c. ricini*. Furthermore, fosmid-FISH mapping identified the gene order of the neo-sex chromosomes in *S. c. walkeri* and *S. cynthia* subsp. indet.

2. Materials and Methods

2.1. Insects

Samia cynthia was originally collected at three different locations: *S. c. walkeri* in Sapporo, Japan, *S. cynthia* subsp. indet. in Nagano, Japan, and *S. c. ricini* in Vietnam (for details, see Yoshido et al., 2005b). For rearing, we released the hatched larvae on an *Ailanthus altissima* tree in a field of the Field Science Center for Northern Biosphere, Hokkaido University (Sapporo, Japan).

2.2. Construction of a *Samia cynthia* fosmid library

A fosmid library was constructed from female pupae of *S. c. walkeri* (Sapporo population). High molecular weight DNA was physically sheared and fractionated by pulsed-filed gel electrophoresis. Blunt-ended DNA fragments collected from the 33 to 48kb fraction were cloned into pCC1FOS vector. The ligates were packaged into *Escherichia coli* (EPI300) by *in vitro* packaging kit of MaxPlax Lambda packaging Extract (EPICENTER, Madison, WI, USA). A total of 35,712 single colonies were picked up and stocked into 93 of 384-well microplates.

2.3. Fosmid screening for FISH mapping

Sequence information of *S. cynthia* orthologs of *B. mori* genes was acquired mainly from the *S. c. ricini* EST database (http://silkbase.ab.a.u-tokyo.ac.jp/cgi-bin/index.cgi) (Arunkumar et al., 2008) and NCBI Genbank. Because the orthologs of *BR-C*, *laccase2*, *PKGIb* and *egg specific protein* have not yet been recorded in any public database, we amplified segments of them by DOP-PCR. Degenerated oligonucleotide primers (Table S1) were designed from regions conserved between *B. mori* and either *Drosophila melanogaster* and/or other lepidopteran species. The amplified fragments from *S. c.*

walkeri genomic DNA were cloned into pGEM T-easy vector (Promega KK, Tokyo, Japan) and sequenced with an ABI PRISM3100 Genetic Analyzer (Applied Biosystems, Tokyo, Japan) according to the protocol recommended by the manufacturer. Tblastn searches by KaikoBlast (http://sgp.dna.affrc.go.jp/KAIKObase/) showed 81.25%, 73.43%, 98.48% and 58.38% identities for the orthologs of *BR-C*, *laccase2*, *PKG-Ib* and *egg specific protein* respectively.

We then designed sequence-tagged site (STS) primers (Table S1) in *S. c. walkeri* in order to isolate clones carrying the respective orthologs from the fosmid library. For the isolation we followed the 3-step method described by Yasukochi (2002). Briefly, the first PCR screening was done against DNA pools from 93 plates, a mixture of 384 fosmids from each plate. The second step was applied only to positive plates with the same size of PCR products as from genomic DNA. We performed PCRs against fosmid-DNA pools from 24 columns and 16 rows to determine the coordinates of the positive clone. In the last step, we individually repeated the PCR to confirm the positive selection. We also confirmed the positive selection by sequencing the PCR products.

2.4. Chromosome preparations

Spread chromosome preparations from females and males were obtained according to methods described in Sahara et al. (1999) and Yoshido et al. (2005b). Briefly, gonads of the last instar larvae were dissected in a saline solution and fixed for 10-15min in Carnoy's fixative (ethanol, chloroform, acetic acid, 6:3:1). Testes but not ovaries were swollen for 10-15 min in a hypotonic solution (75mM KCl) before fixation. Cells were dissociated in 60% acetic acid and spread on a heating plate at 50°C. Then preparations were passed through a graded ethanol series (70%, 80% and 98%) and stored at -30°C until further use.

2.5. Fosmid-FISH mapping

Fosmid-FISH was carried out according to the BAC-FISH method described in Yoshido et al. (2005a) with slight modifications. Fosmid-DNA was extracted with a Plasmid Midi kit (Qiagen GmbH, Hilden, Germany). DNA labeling was done by nick translation using Nick Translation Mix (Roche Diagnostics GmbH, Mannheim, Germany) with Green-dUTP, Orange-dUTP, and Red-dUTP (Abbott Molecular Inc., Des Plaines, IL, USA), and Cy5-dUTP (GE Healthcare UK, Buckinghamshire, UK). We used the reprobing protocol for Lepidoptera (Shibata et al., 2009) for karyotyping and, if more

than four fosmid probes were mapped to a single chromosome. The *S. cynthia* ortholog of *RpL4* was recovered by PCR with primers devised from the EST database (SilkBase) because we could not isolate a suitable fosmid clone in our library. Similarly, 18S rDNA was recovered by PCR. To prepare FISH probes, both were labeled with Orange-dUTP by PCR according to the method described in Yoshido et al. (2010).

Denaturation of preparations was carried out at 72°C for 3.5 min in 70% formamide, 2×SSC, after the slides were passed through an ethanol series and air-dried. For one preparation, we used a probe cocktail with 0.1-0.4μg each labeled DNA (see Table S2) and 3-10μg of unlabeled sonicated *S. cynthia* male genomic DNA in 10μl hybridization solution (50% formamide, 10% dextran sulfate, 2×SSC). Hybridization in a moist chamber at 37°C for 3 days was followed by washing at 62°C in 0.1×SSC, 1% Triton X-100. The slides were counterstained and mounted in 30μl VECTORSHIELD with DAPI (Vector Laboratories, Inc., Burlingame, USA). Counterstained chromosome and hybridization signals were captured with a DFC350FX B&W CCD camera (Leica Microsystems Inc., Tokyo, Japan) described in Sahara et al. (2007). The signal and chromosome images were processed by Adobe Photoshop according to Yoshido et al. (2005a).

3. Results

3.1. Isolation of *S. cynthia* fosmid clones containing orthologs of *B. mori* genes

We isolated a total of 71 fosmid clones by PCR-based screening using 66 STS primer sets (Table S1). The fosmid clones carrying the orthologs of *B. mori* genes were selected from each of the 27 *B. mori* autosomes and the Z chromosome (Table 1).

3.2. Identification of individual chromosomes in *S. cynthia ricini*

To identify individual bivalents in *S. c. ricini*, we carried out FISH mapping on pachytene nuclei using the *S. c. walkeri* fosmid probes. In total, 64 fosmid clones mapped to single locations of *S. c. ricini* chromosomes (Fig. 1, Table 1). Although the clones used were from *S. c. walkeri*, the hybridization signals were strong enough in the other *S. cynthia* subspecies. This enabled us to identify the 13 autosomal bivalents and the Z chromosome of the *S. c. ricini* karyotype (Fig. 1) and their correspondence to *B. mori* (Table 1). Three of the fosmid probes produced multiple signals (Table S2) and, hence were discarded from further experiments.

The Z chromosome of *S. c. ricini* was recognized only by probes which

contained orthologs of *B. mori* Z chromosomal genes. Of the 13 *S. c. ricini* autosomes, ten corresponded to two *B. mori* different autosomes each (Fig. 1, Table 1). There were three exceptions. One was chromosome 1 of *S. c. ricini* which corresponded to four *B. mori* chromosomes, 2, 20, 26, and 27. The others were chromosomes 11 and 12. Chromosome 11 of *S. c. ricini* corresponded to *B. mori* chromosome 24 and a part of *B. mori* chromosome 11. *S. c. ricini* chromosome 12 corresponded to *B. mori* chromosome 21 and the other part of *B. mori* chromosome 11. In those cases where we used more than two probes (*B. mori* chromosomes 8, 11, 12, 15 and 19), even the gene order was conserved between the component parts of *S. c. ricini* chromosomes and the *B. mori* chromosomes. We confirmed that *S. c. ricini* chromosome 11 was the chromosome carrying nucleolar organizer region (NOR) by mapping the 18S rDNA fragment to that chromosome (Fig. 1). Our previous data of *B. mori* BAC-FISH revealed that the *B. mori* NOR located somewhere between *RpL4* and *Topo II* (Yoshido et al., 2010). In *S. c. ricini*, however, the signal of 18S rDNA (white signal of Chr11 in Fig. 1) was detected in the region between the *RpL4* ortholog and the *P450* ortholog. The result reveals that a limited repositioning of the NOR has occurred in the *B. mori* or *S. c. ricini* phylogenetic lineages.

3.3. Fosmid-FISH karyotype of *Samia cynthia*

For karyotyping, i. e. identifying all chromosomes of a set, we used a total of 22 fosmid probes (Fig. 2, Table S2). We identified 11 autosomal bivalents (Chromosome 2-10, 12 and 13) in a first round of FISH. They were distinguished by the color combination of the probes, or by the different signal position when they had the same color combination (chromosome 2 and 6, 3 and 5, 9 and 12). To recognize the remaining chromosomes 1, 11 and Z, we reprobed the slides in a second round of FISH. In addition, we used two probes for chromosomes 2 and 3 to confirm results of the first round. In this manner, we identified all 13 autosomal bivalents and the Z univalent, and hence definitively karyotyped the *S. c. ricini* pachytene complement (Fig. 2).

3.4. Sex chromosome variation in *S. cynthia* subspecies

We also carried out fosmid-FISH mapping in the other two *S. cynthia* subspecies, *S. c. walkeri* (Sapporo population) and *S. cynthia* subsp. indet. (Nagano population). The homologue of *S. c. ricini* chromosome 13 forms the neo-sex chromosomes (the segment of autosomal origin) of *S. c. walkeri* and *S. cynthia* subsp. indet. (Fig. 3b, c). In *S. cynthia* subsp. indet., the homologue of *S. c. ricini* and *S. c. walkeri* chromosome 12 contribute an additional segment to the neo-W and the Z_2 chromosome (Fig. 3a, c). The

results confirm the step-by-step evolution of sex chromosome recently proposed by Yoshido et al. (2010) in these *S. cynthia* subspecies. Our results also show that the gene order is strictly conserved between the respective segments of autosomal origin in neo-sex chromosomes from the Japanese subspecies of *S. cynthia* and the homologous autosomes of *S. c. ricini* and *B. mori* (compare Figs. 1 and 3).

4. Discussion

We show here that each chromosome of *S. cynthia* can be reliably identified by FISH, using fosmid probes. A selected set of 22 probes from a genomic fosmid library was sufficient to karyotype *S. cynthia* (Fig. 2). Fosmid clones have been used as reliable cytogenetical markers for FISH karyotyping in animals and plants (Zhang et al., 2008; Dalzell et al., 2009; Liu et al., 2010). However, this is the first application of fosmid-FISH in Lepidoptera and - as far as we know - in all insect. Because constructing a fosmid library is easier than a BAC library, FISH with fosmid clones will accelerate gene mapping in Lepidoptera.

Fosmid-FISH clearly revealed that gene order of *S. c. ricini* chromosomes 12

and 13 was conserved in the corresponding parts of the neo-sex chromosomes of *S. c. walkeri* and *S. cynthia* subsp. indet. (Fig. 3). Consequently, the step-by-step evolution of the sex chromosome system as proposed by Yoshido et al. (2010) in the three *S. cynthia* subspecies has been fully confirmed in the present study. In addition, comparative gene mapping of their chromosomes shows a high degree of collinearity between the *S. cynthia* subspecies and *B. mori* (Fig. 3).

The supposed ancestral chromosome number in Lepidoptera is n=31 (Robinson, 1971; Lukhtanov, 2000; De Prins and Saitoh, 2003). *B. mori* (n=28) is rather close to the ancestral chromosome number. We show here that the autosomes of *S. cynthia* are composed of either two or four *B. mori* homologues, apparently complete the karyotype, with one exception (see next paragraph). Thus the low number-karyotype of *S. cynthia* was indeed formed by multiple chromosome fusions. The same explanation was also proposed by Pringle et al. (2007) for the decreased chromosome number (n=21) of the *H. melpomene* karyotype. The fusion partners, however, were different in the two species (Fig. 1, Table 1). Hence, the fusion events in *S. cynthia* have happened independently from those in *H. melpomene*.

The exceptional autosome, *B. mori* chromosome 11, appeared to have split to contribute to two *S. cynthia* chromosomes, *S. c. ricini* chromosome 11 and 12. In fact, *B.*

mori chromosome 11 itself is probably the fusion product of two separate ancestral chromosomes, as it was found homologous to two autosomes in yet another species, *Manduca sexta* (Yasukochi et al., 2009). Thus, *S. c. ricini* chromosome 11 and 12 are probably also the fusion products of two entire ancestral chromosomes each.

B. mori has a single NOR which is located in chromosome 11 (Yoshido et al., 2005a). According to our BAC-FISH mapping, the rDNA repeats in *B. mori* locate somewhere between *RpL4* (position 2.86 Mb in Kaikobase) and *Topo II* (position 8.6 Mb in Kaikobase). Although a single NOR was also observed in *S. c. ricini* (Fig. 1), it maps to a region between the orthologs of *B. mori RpL4* and *P450* (in *B. mori* at position 2.86 Mb and 1.27 Mb of chromosome 11, respectively) (Table 1). The inverted order of *RpL4* and NOR in *S. c. ricini* may be due to a rearrangement accompanying the fusion process or to NOR transposition. In *M. sexta*, the single NOR also had an altered position relative to *B. mori* (Yasukochi et al., 2009). Transposition or translocation of rDNA appears to be a rather common event in insects (Roy et al., 2005; Cabrero and Camacho, 2008; Nguyen et al., 2010) and plants (Schubert and Wobus, 1985; Dubcovsky and Dvořák, 1995; Datson and Murray, 2006).

Gene order is well conserved between *B. mori* chromosomes and the corresponding chromosome segments in *S. cynthia*. We detected no change in the order

of orthologous genes in all cases in which the order could be inferred from mapping more than two probes (see Fig. 1, chromosome 5, 7, 12 and 13, and Table 1), the only exception being the rDNA locus. Gene-based comparative mapping has so far revealed that the gene order in *B. mori* (n=28) and either *H. melpomene* (n=21) (Pringle et al., 2007), *B. anynana* (n=28) (Beldade et al., 2009) or *M. sexta* (n=28) (Yasukochi et al., 2009) was rather well conserved too, with only a few translocations, inversions and fusion/fission events. In *S. cynthia,* even though the autosomes had undergone fusion events in their evolutionary history, gene order was conserved in the corresponding chromosome parts.

In contrast, the relationships of genomic components among dipteran species (fly and mosquito) based on genomic sequence resources have revealed higher degree of rearrangements. Comparative genome analysis in dipterans species, *D. melanogaster*, *Anopheles gambiae*, *Aedes aegypti* showed shuffling of gene order with only conservation of microsynteny among species (Zdobnov et al., 2002; Nene et al., 2007). Among more closely related species in genus *Drosophila* (12 species) or in genus *Anopheles* (2 species), their gene order has been also shuffled along corresponding chromosomal arms (Ranz et al., 2001; Sharakhov et al., 2002; Bhutkar et al., 2008). Taken together, in dipteran species, inter- and/or intra-chromosomal rearrangements

may have been frequently occurred in evolutionary history. Hence, higher degree of internal stability in the lepidopteran chromosomes is rather surprising.

Both phenomena, conserved synteny of whole chromosomes and conserved gene order, were also observed in birds. Conserved gene order in the avian large autosomes, the so-called macrochromosomes, has been revealed by gene-based molecular linkage analysis (Backström et al., 2006, 2008; Stapley et al., 2008). Collinearity of the small chromosomes, the so-called microchromosomes, between chicken and turkey (*Meleagris gallopavo*) has also been detected (Griffin et al., 2008). A recent genome sequence comparison between chicken and zebra finch (*Taeniopygia guttata*) disclosed high homology between 6 chicken and 8 zebra finch macrochromosomes (Ellegren, 2010). Although inverted order of chromosomal segments and fusion/fission events are apparent between the distantly related bird species, the degree of synteny is high among birds. Ellegren (2010) attributed the stability of chromosomes to the lower density of interspersed repetitive elements in birds, approximately 10% in the chicken genome compared to 40-50% in other sequenced animal genomes. The lower density is supposed to reduce the possibility of chromosomal rearrangements mediated by repeat elements.

This explanation does not hold true for the conserved synteny and gene order in

Lepidoptera. The *B. mori* genome contains 43.6% interspersed repetitive sequences (The International Silkworm Genome Consortium, 2008). There is also another reason to expect even more rearrangements than in other animals. The holokinetic chromosome structure of Lepidoptera (Murakami and Imai, 1974; Traut, 1986; Wolf, 1996) should stand more chromosomal rearrangements than the common monocentric structure (see discussion in Yasukochi et al., 2009). Since fragments of holokinetic chromosomes can survive during cell cycles (Fujiwara et al., 2000) or even several generations (Rathjens, 1974; Marec et al., 2001), the risk for lethality is much lower. It is the more surprising that Lepidoptera exhibit conserved synteny and high stability of gene order within chromosomes. It is open to which property of the genome or chromosome structure the Lepidoptera chromosomes owe their stability.

Acknowledgements

We thank W. Traut (Lübeck, Germany) for critical reading the manuscript and valuable suggestions. Our thanks also go to Y. Yamada (Sapporo, Japan) for supporting to rear *Samia cynthia*, T. Fujii (Tokyo, Japan) for providing *S. c. ricini*, and M.

Tanaka-Okuyama, H. Ohta, E. Igari, H. Hoshida (Tsukuba, Japan) for preparing fosmid library. A.Y. received a grant 19･1114 of Japan Society for the Promotion of Science (JSPS). We acknowledge the research fellowship 21-7147 given to K.S. from JSPS.

References

Arunkumar, K.P., Tomar, A., Daimon, T., Shimada, T., Nagaraju, J., 2008. WildSilkbase: An EST database of wild silkmoths. BMC Genomics 9, 338.

Backström, N., Brandström, M., Gustafsson, L., Qvarnström, A., Cheng, H., Ellegren, H., 2006. Genetic mapping in a natural population of collared flycatchers (*Ficedula albicollis*): conserved synteny but gene order rearrangements on the avian Z chromosome. Genetics 174, 377-386.

Backström, N., Karaiskou, N., Leder, E.H., Gustafsson, L., Primmer, C.R., Qvarnström, A., Ellegren, H., 2008. A gene-based genetic linkage map of the collared flycatcher (*Ficedula albicollis*) reveals extensive synteny and gene-order conservation during 100 million years of avian evolution. Genetics 179, 1479-1495.

Beldade, P., Saenko, S.V., Pul, N., Long, A.D., 2009. A gene-based linkage map for *Bicyclus anynana* butterflies allows for a comprehensive analysis of synteny with the lepidopteran reference genome. PLoS Genet. 5, e1000366.

Bhutkar, A., Schaeffer, S.W., Russo, S.M., Xu, M., Smith, T.F., Gelbart, W.M., 2008. Chromosomal rearrangement inferred from comparison of 12 *Drosophila* genomes. Genetics 179, 1657-1680.

Cabrero, J., Camacho, J.P.M., 2008. Location and expression of ribosomal RNA genes in grasshoppers: Abundance of silent and cryptic loci. Chromosome Res. 16, 595-607.

d'Alençon, E., Sezutsu, H., Legeai, F., Permal, E., Bernard-Samain, S., Gimenez, S., Gagneur, Z., et al., 2010. Extensive synteny conservation of holocentric chromosomes in Lepidoptera despite high rates of local genome rearrangements. Proc. Natl. Acad. Sci. USA 107, 7680-7685.

Dalzell, P., Miles, L.G., Isberg, S.R., Glenn, T.C., King, C., Murtagh, V., Moran, C., 2009. Standardized reference ideogram for physical mapping in the saltwater crocodile (*Crocodylus porosus*). Cytogenet. Genome Res. 127, 204-212.

Datson, P.M., Murray, B.G., 2006. Ribosomal DNA locus evolution in *Nemesia*: transposition rather than structural rearrangement as the key mechanism? Chromosome Res. 14, 845-857.

De Prins, J., Saitoh, K., 2003. Karyology and sex determination, in: Kristensen, N.P. (Ed.), Lepidoptera, moths and butterflies. 2. Morphology, physiology, and development. Walter de Gruyter, Berlin & New York, pp. 449-468.

Dubcovsky, J., Dvořák, J., 1995. Ribosomal RNA multigene loci: nomads of the *Triticeae* genomes. Genetics 140, 1367-1377.

Ellegren, H., 2010. Evolutionary stasis: the stable chromosomes of birds. Trends Ecol. Evol. 25, 283-291.

Ferguson-Smith, M.A., Trifonov, V., 2007. Mammalian karyotype evolution. Nat. Rev. Genet. 8, 950-962.

Fujiwara, H., Nakazato, Y., Okazaki, S., Ninaki, O., 2000. Stability and telomere structure of chromosomal fragments in two different mosaic strains of the silkworm, *Bombyx mori*. Zool. Sci. 17, 743 –750.

Griffin, D.K., Robertson, L.B., Tempest, H.G., Vignal, A., Fillon, V., Croolimans, R.P., et al., 2008. Whole genome comparative studies between chicken and turkey and their implications for avian genome evolution. BMC Genomics 9, 168.

Grimaldi, D.A., Engel, M.S., 2005. Evolution of the insects. Cambridge Univ. Press, New York.

Kristensen, N.P., Skalski, A.W., 1999. Phylogeny and palaeontology, in: Kristensen, N.P. (Ed.), Lepidoptera, moths and butterflies. 1. Evolution, systematics, and biogeography. Walter de Gruyter, Berlin & New York, pp. 7-25.

Liu, C., Liu, J., Li, H., Zhang, Z., Han, Y., Huang, S., Jin, W., 2010. Karyotyping in melon (*Cucumis melo* L.) by cross-species fosmid fluorescence *in situ* hybridization. Cytogenet. Genome. Res. 129, 241-249.

Lukhtanov, V.A., 2000. Sex chromatin and sex chromosome systems in non-ditrysian Lepidoptera (Insecta). J. Zool. Syst. Evol. Res. 38, 73-79.

Marec, F., Tothová, A., Sahara, K., Traut, W., 2001. Meiotic pairing of sex chromosome fragments and its relation to atypical transmission of a sex-linked marker in *Ephestia kuehniella* (Insecta: Lepidoptera). Heredity 87, 659-671.

Marec, F., Sahara, K., Traut, W., 2010. Rise and Fall of the W chromosome in Lepidoptera, in: Goldsmith, M.R., Marec, F. (Eds.), Molecular Biology and Genetics of the Lepidoptera. CRC Press, Boca Raton, FL, USA, pp. 49-63.

Murakami, A., Imai, H.T., 1974. Cytological evidence for holocentric chromosomes of the silkworms, *Bomby mori* and *B. mandarina,* (Bombycidae, Lepidoptera). Chromosoma 47, 167-178.

Nene, V., Wortman, J.R., Lawson, D., Haas, B., Kodira, C., Tu, Z.J., et al., 2007. Genome sequence of *Aedes aegypti*, a major arbovirus vector. Science 316, 1718-1723.

Nguyen, P., Sahara, K., Yoshido, A., Marec, F., 2010. Evolutionary dynamics of rDNA clusters on chromosomes of moths and butterflies (Lepidoptera). Genetica 138, 343-354.

Pringle, E.G., Baxter, S.W., Webster, C.L., Papanicolaou, A., Lee, S.F., Jiggins, C.D.,

2007. Synteny and chromosome evolution in the Lepidoptera: evidence from mapping in *Heliconius melpomene*. Genetics 177, 417-426.

Ranz, J.M., Casals, F., Ruiz, A., 2001. How malleable is the eukaryotic genome? Extreme rate of chromosomal rearrangement in the genus *Drosophila*. Genome Res. 11, 230-239.

Rathjens, B., 1974. Zur Funktion des W-chromatins bei *Ephestia kuehniella* (Lepidoptera). Isolierung und Charakterisierung von W-Chromatin-Mutanten. Chromosoma 47, 21-44.

Robinson, R., 1971. Lepidoptera genetics. Pergamon, Oxford.

Roy, V., Monti-Dedieu, L., Chaminade, N., Siljak-Yakovlev, S., Aulard, S., et al., 2005. Evolution of the chromosomal location of rDNA genes in two *Drosophila* species subgroups: *ananassae* and *melanogaster*. Heredity 94, 388-395.

Sahara, K., Marec, F., Traut, W., 1999. TTAGG telomeric repeats in chromosomes of some insects and other arthropods. Chromosome Res. 7, 449-460.

Sahara, K., Yoshido, A., Marec, F., Fuková, I., Zhang, H.B., Wu, C. C., et al., 2007. Conserved synteny of genes between chromosome 15 of *Bombyx mori* and a chromosome of *Manduca sexta* shown by five-color BAC-FISH. Genome 50, 1061-1065.

Schubert, I., Wobus, U., 1985. In situ hybridization confirms jumping nucleolus organizing regions in *Allium*. Chromosoma 92, 143-148.

Sharakhov, I.V., Serazin, A.C., Grushko, O.G., Dana, A., Lobo, N., Hillenmeyer, M.E., et al., 2002. Inversions and gene order shuffling in *Anopheles gambiae* and *A. funestus*. Science 298, 182-185.

Shibata, F., Sahara, K., Naito, Y., Yasukochi, Y., 2009. Reprobing of multicolour FISH in preparations of lepidopteran chromosomes. Zool. Sci. 26, 187-190.

Stapley, J., Birkhead, T.R., Burke, T., Slate, J., 2008. A linkage map of the zebra finch *Taeniopygia guttata* provides new insights into avian genome evolution. Genetics 179, 651-667.

The International Silkworm Genome Consortium, 2008. The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. Insect Biochem. Mol. Biol. 38, 1036-1045.

Traut, W., 1986. A genetic linkage study of W-chromosome-autosome fusions, breakage, and kinetic organization of chromosomes in *Ephestia* (Lepidoptera). Genetica 69, 69-79.

Traut, W., Sahara, K., Marec, F., 2007. Sex chromosomes and sex determination in Lepidoptera. Sex. Dev. 1, 332-346.

Wolf, K.W., 1996. The structure of condensed chromosomes in mitosis and meiosis of

insects. Int. J. Insect Morphol. Embryol. 25, 37-62.

Yasukochi, Y., 2002. PCR-based screening for bacterial artificial chromosome libraries. Methods Mol. Biol. 192, 401-410.

Yasukochi, Y., Ashakumary, L.A., Baba, K., Yoshido, A., Sahara, K., 2006. A second-generation integrated map of the silkworm reveals synteny and conserved gene order between lepidopteran insects. Genetics 173, 1319-1328.

Yasukochi, Y., Tanaka-Okuyama, M., Shibata, F., Yoshido, A., Marec, F., Wu, C., et al., 2009. Extensive conserved synteny of genes between the karyotypes of *Manduca sexta* and *Bombyx mori* revealed by BAC-FISH mapping. PLoS One 4, e7465.

Yoshido, A., Bando, H., Yasukochi, Y., Sahara, K., 2005a. The *Bombyx mori* karyotype and the assignment of linkage groups. Genetics 170, 675-685.

Yoshido, A., Marec, F., Sahara, K., 2005b. Resolution of sex chromosome constitution by genomic *in situ* hybridization and fluorescence *in situ* hybridization with (TTAGG)*n* telomeric probe in some species of Lepidoptera. Chromosoma 114, 193-202.

Yoshido, A., Sahara, K., Marec, F., Matsuda, Y., 2010. Step-by-step evolution of neo-sex chromosomes in geographical populations of wild silkmoths, *Samia cynthia* ssp. Heredity doi: 10.1038/hdy.2010.94.

Zhang, L., Bao, Z., Wang, S., Hu, X., Hu, J., 2008. FISH mapping and identification of zhikong scallop (*Chlamys farreri*) chromosomes. Marine Biotechnol. 10, 151-157.

Zdobnov, E.M., von Mering, C., Letunic, I., Torrents, D., Suyama, M., Copley, R.R., Christophides, G.K., et al., 2002. Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*. Science 298, 149-159.

Figure legends

Fig. 1. FISH identification of individual pachytene bivalents in *Samia cynthia ricini* and their comparison with *Bombyx mori* chromosomes. Note that the lengths of chromosome bivalents cannot be compared in this figure as they are derived from different pachytene complements and/or preparations in both sexes. Vertical bars represent corresponding parts of *B. mori* chromosomes (black italic numbers) drawn to relative scales in Mb taken from Kaikobase. *Samia* chromosomes were counterstained with DAPI (light blue). For details of fosmid probes used and their labeling, see M&M section, Table 1 and Table S2.

Fig. 2. Fosmid-FISH karyotype of *Samia cynthia ricini*. (**a**) A female postpachytene complement (most probably developing to a nurse cell) of 13 autosomal bivalents and a Z-chromosome univalent, each identified by 1-3 pseudocolored hybridization signals of fosmid probes. Chromosomes were counterstained with DAPI (light blue). A total of 22 fosmid were hybridized to this complement (for details, see Table 1 and Table S2). (**b**) Thirteen autosomal bivalents and the Z-chromosome univalent of the same postpachytene complement as in (**a**) arranged according to their chromosome numbers

(see Table 1). Bar = 10μm.

Fig. 3. Conserved synteny of genes between the sex chromosomes and its autosomal homologs in *Samia cynthia* subspecies. Signal orders of probes on the chromosome Z, 12 and 13 of *S. c. ricini* (drawing in the left side) are conserved both in an autosomal bivalent (**a**) and neo-sex chromosomes (**b**) of *S. c. walkeri*, and neo-sex chromosomes of *S. cynthia* subsp. indet. (**c**). Note that the distal end of the original Z chromosome parts in the neo-sex bivalent (**b**) and neo-sex trivalent (**c**), respectively, remained unpaired most probably because of length differences between paring partners. Also note the asymmetrical signals of each 32B23 (magenta) and 56J22 (red) fosmid probes on the *S. cynthia* subsp. indet. neo-sex chromosome trivalent, most probably resulting from unequal length of paired chromosomes. The sex chromosome system of each subspecies (populations) are described below the sex chromosome drawing and figures; Z0 for *S. c. ricini* (Vietnam), neo-Wneo-Z for *S. c. walkeri* (Sapporo) and neo-WZ_1Z_2 for *S. cynthia* subsp. indet. (Nagano). See Table S2 for details of probe information.

Table 1 Summary of FISH mapping of *Samia cynthia* fosmids, containing orthologs of *Bombyx mori* genes; the fosmids are listed acccording to the order of corresponding hybridization signals in a particular chromosome (see Fig. 1).

S. cynthia					*B. mori*	
Chr No.	fosmid code	Accession No.	EST (SilkBase*)	putative function	Chr No.	Location in Kaikobase**
1	39N17	DQ465407	S06A01NCLL0005_L15	3-hydroxy-3-methylgultary-CoA reducase	2	916,956-919,459
1	9G21		I10A02NGRL0003_G06	ribosomal protein S21	2	9,846,943-9,848,321
1	28O11		S06A01NCLL0014_E20	eukaryotic translation initiation factor 2α	20	7,111,198-7,115,385
1	63H6		I10A02NGRL0008_J19	ribosomal protein S20	20	13,287,125-13,290,285
1	57F7		S13A02NGRL0004_B05	hypothetical protein	26	741,104-745,803
1	42E11		S13A01NGRL0005_O18	THO complex subunit 3	26	2,003,467-2,008,338
1	62D4	AB190810	not found	vitellogenin	27	10,593,279-10,604,046
1	21F6		S13A02NGRL0008_K14	HSP70	27	12,563,862-12,565,843
2	47N14		S06A01NCLL0006_I20	ribosomal protein L15	6	4,970,835-4,972,168
2	5C24		I09A02NGRL0003_D23	attacin	6	18,597,002-18,597,814
2	51F19	AB564748	not found	laccase2	10	4,836,432-4,840,937
2	20E12		S13A01NGRL0005_E18	diapause bioclock protein	10	17,019,682-17,023,915
3	6P17		S13A01NGRL0008_G03	serpin-4A	28	10,271,213-10,272,445
3	35P19		I10A02NGRL0006_O06	ribosomal protein S11	28	8,034,435-8,039,353
3	60P13		S13A01NGRL0009_H12	ribosomal protein S27	7	8,032,934-8,033,836
3	55M14	AB201280	not found	chitinase	7	14,117,862-14,119,517
4	20J22		I10A02NGRL0006_L21	cytoplasmic actin (A4)	17	2,936,956-2,938,173
4	33F8		I10A02NGRL0008_D04	ribosomal protein S24	17	16,084,868-16,085,379
4	12E18		I09A02NGRL0002_J15	ribosomal protein L19	5	3,573,371-3,577,651
4	25G4		I09A02NGRL0007_G07	elongation factor 1α	5	17,106,041-17,107,432
5	61P10		S06A01NCLL0022_K14	ribosomal protein L5	15	5,494,680-5,499,312
5	63B20		S06A01NCLL0009_G10	ribosomal protein S5	15	7,586,991-7,587,650
5	32P7		S13A01NGRL0001_M19	ribosomal protein S8	15	13,998,126-14,000,450
5	27A12		S13A02NGRL0011_B17	bmtub4	15	16,664,500-16,666,892
5	18O10		S06A01NCLL0013_L06	Ala-tRNA synthetase	4	8,169,535-8,171,127
5	45O20	AF015065	S13A01NGRL0011_F03	dopa decarboxylase	4	12,504,927-12,521,471
6	29M1		S06A01NCLL0021_G23	ribosomal protein L6	16	1,241,297-1,245,534
6	52K2		I10A02NGRL0008_D12	ribosomal protein P2	16	12,100,458-12,102,307
6	47I5		I09A02NGRL0007_E20	ribosomal protein S30	23	12,462,114-12,465,725
6	56E18		S13A02NGRL0010_N13	storage protein (SP1)	23	16,906,448-16,911,216
7	44L5		S06A01NCLL0024_M07	elongation factor 1 gamma	18	12,666,084-12,671,258
7	49D20	AB026557	not found	sensory neuron membrane protein-1 (snmp)	18	not mapped***
7	13B10		I09A02NGRL0002_I11	ribosomal protein S9	19	1,216,844-1,218,990
7	52I13		S06A01NCLL0007_L10	translation elongation factor 2	19	12,802,871-12,811,573
7	17F23	AB564749	not found	egg-specific protein	19	13,154,521-13,156,200
8	34H14	L25668	not found	PTTH	22	10,390,124-10,391,823
8	12L19		I10A02NGRL0007_G03	calreticulin	22	4,685,110-4,689,288
8	23N7		I09A02NGRL0004_J13	ribosomal protein S25	25	6,334,666-6,349,673
8	46C5		S13A02NGRL0004_F13	prophenoloxidase activating factor 3	25	8,534,344-8,540,906
9	59L14		S13A02NGRL0001_I05	arylphorin (SP2)	3	8,650,251-8,654,561
9	52F14	AB564750	not found	PKGIb	3	8,060,428-8,060,994
9	55G12		I09A02NGRL0002_B18	ribosomal protein L31	13	9,474,668-9,476,195
9	24K19		I09A02NGRL0007_H06	ribosomal protein L21	13	16,214,502-16,216,987

Table 1 (continued)

S. cynthia					*B. mori*	
Chr No.	fosmid code	Accession No.	EST (SikBase*)	putative function	Chr No.	Location in Kaikobase**
10	54I24		I09A02NGRL0001_C05	allatostatin preprohormone	14	205,746-208,559
10	26D8		S13A01NGRL0004_L17	hypothetical protein	14	14,055,940-14,059,769
10	53P19		S13A02NGRL0002_A22	ribosomal protein S14	9	14,970,011-14,971,055
10	3A24		I10A02NGRL0001_K04	LCP18	9	18,065,584-18,066,900
11	11P18		S06A01NCLL0011_D13	ADP/ATP translocase	24	15,066,947-15,067,789
11	15C21	AB220992	I09A02NGRL0003_K24	hemolin	24	13,879,059-13,884,127
11	32H9		S06A01NCLL0008_E19	P450	11	1,265,031-1,270,289
11	not found	AB543313, AB543314	S13A01NGRL0010_F15	ribosomal protein L4	11	2,852,029-2,862,665
12	11O1	AB543310	not found	topoisomerase II (Topo II)	11	8,589,427-8,602,332
12	44L2		S06A01NCLL0020_P12	ribosomal protein L18	11	9,399,580-9,401,164
12	22B11		I09A02NGRL0005_B07	ribosomal SOP2	11	19,584,552< <23,949,056***
12	17P9		S13A02NGRL0011_J07	sorbitol dehydrogenase	21	1,680,348-1,690,455
12	38O13		S06A01NCLL0012_N21	UBI3 mRNA for polyubiquitin	21	11,614,170-11,616,911
13	56J8		S13A01NGRL0004_C20	hypothetical protein	8	205,759-211,533
13	32B23		S06A01NCLL0014_M19	glycin rich protein	8	4,927,144-4,928,289
13	56J22	AB564751	not found	broad-complex (BR-C)	8	18,077,270-18,081,471
13	75O1	AB048258	I10A02NGRL0003_P23	lysozyme	12	3,309,259-3,312,565
13	21P14		I10A02NGRL0003_B23	p109	12	4,989,327-5,006,556
13	44E23		S13A02NGRL0001_G21	J domain containing protein (JDCP)	12	7,728,258-7,729,559
13	14J3	AB543315	I09A02NGRL0003_I02	xanthine dehydrogenase I (XDH I)	12	8,493,509-8,511,254
Z	45A6	AB543309	not found	kettin	1	6,505,696-6,533,895
Z	19B8		S06A01NCLL0014_I05	hypothetical protein	1	11,424,630-11,431,079

*see http://silkbase.ab.a.u-tokyo.ac.jp/cgi-bin/index.cgi

**see http://kaikoblast.dna.affrc.go.jp/

*** The *B. mori* sensory neuron membrane protein-1 and ribosomal SOP2 mapped in the contigs 18_a and 11_3, respectively (Yasukochi et al., 2006). Only the location of ribosomal SOP2 can be determined from the primer sequence of the contigs 11_2 and 11_6 described in Tables S1 and S2 in Yasukochi et al. (2006).

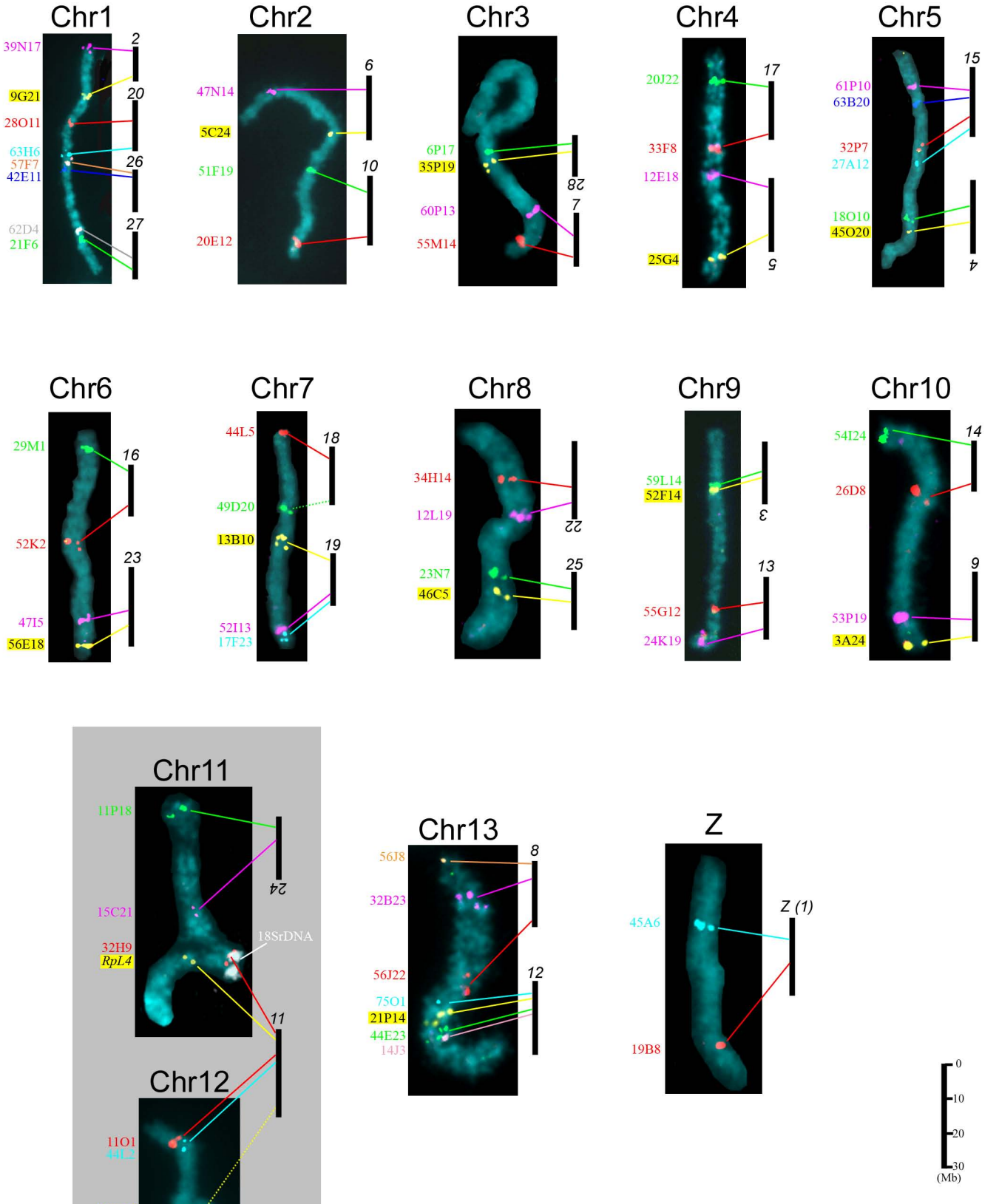
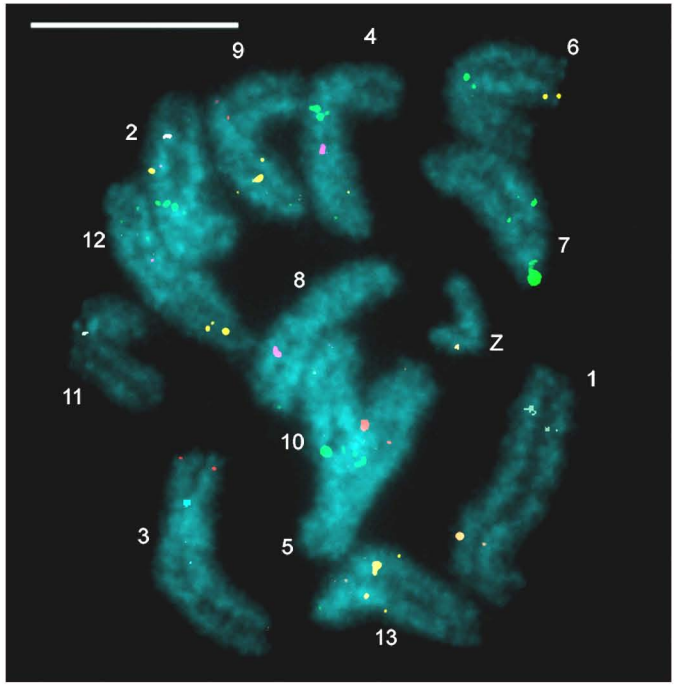

Chr1
39N17
9G21
28O11
63H6
57F7
42E11
62D4
21F6
2
20
26
27
Chr2
47N14
5C24
51F19
20E12
6
10
Chr3
6P17
35P19
60P13
55M14
28
7
Chr4
20J22
33F8
12E18
25G4
17
5
Chr5
61P10
63B20
32P7
27A12
18O10
45O20
15
4
Chr6
29M1
52K2
47I5
56E18
16
23
Chr7
44L5
49D20
13B10
52I13
17F23
18
19
Chr8
34H14
12L19
23N7
46C5
22
25
Chr9
59L14
52F14
55G12
24K19
3
13
Chr10
54I24
26D8
53P19
3A24
14
9
Chr11
11P18
15C21
32H9
RpL4
18SrDNA
24
11
Chr12
11O1
44I2
22B11
17P9
38O13
21
Chr13
56J8
32B23
56J22
75O1
21P14
44E23
14J3
8
12
Z
45A6
19B8
Z (1)
0
10
20
30
(Mb)

Fig. 1

a


9
4
6
2
12
7
8
Z
1
11
10
3
5
13

b

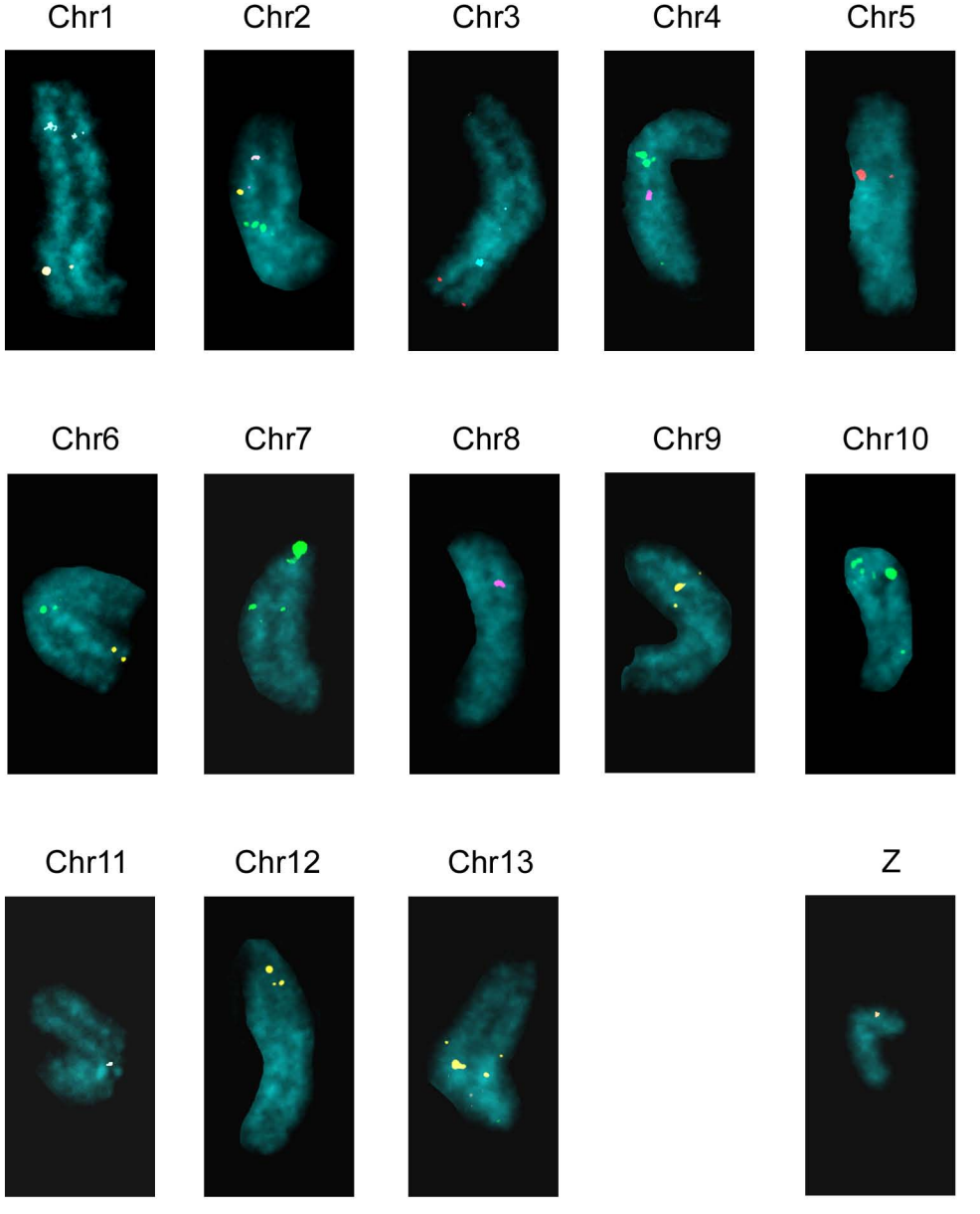
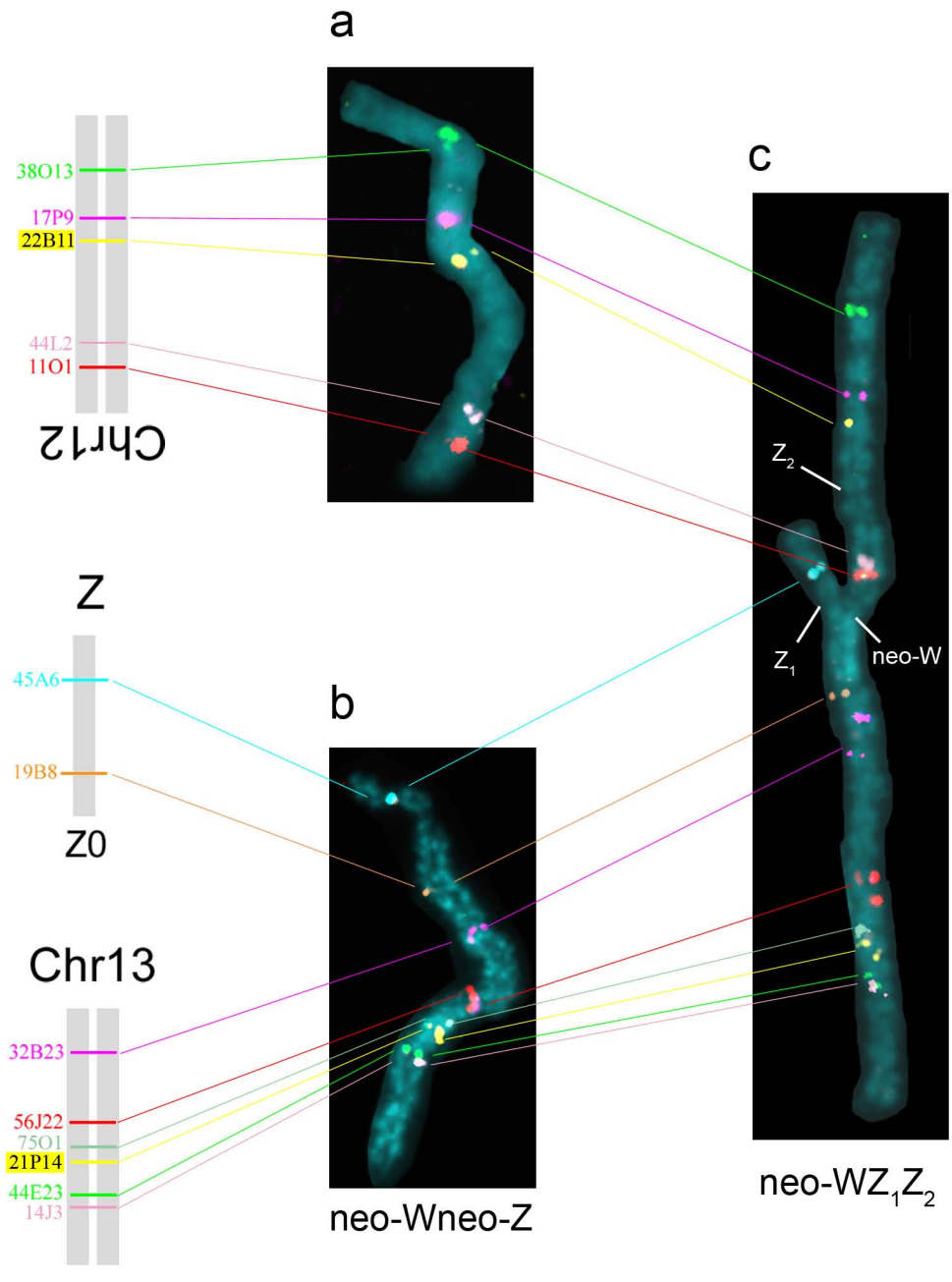

Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chr10
Chr11
Chr12
Chr13
Z

Fig. 2



S. c. ricini **(Vietnam)**

S. c. walkeri **(Sapporo)**

S. cynthia **subsp. indet. (Nagano)**

Fig. 3