Heredity (2016) 116, 75–83

www.nature.com/hdy

ORIGINAL ARTICLE

# A FISH-based chromosome map for the European corn borer yields insights into ancient chromosomal fusions in the silkworm

Y Yasukochi[1], M Ohno[2], F Shibata[3,5], A Jouraku[1], R Nakano[4,6], Y Ishikawa[4] and K Sahara[2]

A significant feature of the genomes of Lepidoptera, butterflies and moths, is the high conservation of chromosome organization. Recent remarkable progress in genome sequencing of Lepidoptera has revealed that syntenic gene order is extensively conserved across phylogenetically distant species. The ancestral karyotype of Lepidoptera is thought to be $n=31$; however, that of the most well-studied moth, *Bombyx mori*, is $n=28$, and diverse studies suggest that three chromosomal fusion events occurred in this lineage. To identify the boundaries between predicted ancient fusions involving *B. mori* chromosomes 11, 23 and 24, we constructed fluorescence *in situ* hybridization (FISH)-based chromosome maps of the European corn borer, *Ostrinia nubilalis* ($n=31$). We first determined a 511 Mb genomic sequence of the Asian corn borer, *O. furnacalis,* a congener of *O. nubilalis,* and isolated bacterial artificial chromosomes and fosmid clones that were expected to localize in candidate regions for the boundaries using these sequences. Combined with FISH and genetic analysis, we narrowed down the candidate regions to 40 kb–1.5 Mb, in strong agreement with a previous estimate based on the genome of a butterfly, *Melitaea cinxia.* The significant difference in the lengths of the candidate regions where no functional genes were observed may reflect the evolutionary time after fusion events.
*Heredity* (2016) **116,** 75–83; doi:10.1038/hdy.2015.72; published online 12 August 2015

## INTRODUCTION

Evolutionary conservation of karyotypes and internal chromosome structure show high diversity among multicellular animals. A high degree of chromosome rearrangement has been reported in many organisms; yet, conservation of approximate haploid chromosome number and gene contents in parts of chromosomes (synteny blocks) makes it possible to trace the history of rearrangements in mammalian chromosomes (Kemkemer *et al.*, 2009). In contrast, a slow rate of change in karyotype and interchromosomal rearrangements are observed for the vertebrate clade Sauropsida, which includes birds, lizards, snakes and turtles (Alföldi *et al.*, 2011; Pokorná *et al.*, 2012). Among insects, notably Diptera, flies and mosquitoes, comparative genomics reveals a high degree of gene-order shuffling which suggests frequently occurring chromosome rearrangements, although total chromosome numbers remain relatively low (Nene *et al.*, 2007).

Accumulating evidence currently reveals that the gene order and content of each chromosome are highly conserved among lepidopteran species, butterflies and moths. As we first described syntenic gene order on four pairs of chromosomes between the silkworm, *Bombyx mori*, and a butterfly, *Heliconius melpomene* (Yasukochi *et al.*, 2006), a number of reports have demonstrated extensive shared synteny between *B. mori* and other species based on genetic linkage studies (Pringle *et al.*, 2007; Beldade *et al.*, 2009; Baxter *et al.*, 2011; Van't Hof *et al.*, 2013) and chromosomal fluorescence *in situ* hybridization (FISH) analysis (Yasukochi *et al.*, 2009; Yoshido *et al.*, 2011; Sahara *et al.*, 2013).

As $n=31$ is the most common chromosome number observed in distantly related families of Lepidoptera (Robinson, 1971), it is likely that $n=31$ was the karyotype of the common ancestor. Reports using distantly related species having 31 chromosomes consistently show that chromosomes 11, 23 and 24 of *B. mori* ($n=28$) were generated by ancestral chromosomal fusion events (Baxter *et al.*, 2011; Sahara *et al.*, 2013; Van't Hof *et al.*, 2013). Genome sequencing with high resolution assignments to chromosomes enabled more direct and precise comparison between species (International Silkworm Genome Consortium 2008; *Heliconius* Genome Consortium, 2012; Ahola *et al.*, 2014). The genomes of *B. mori* and *H. melpomene* ($n=21$) were revealed to be composed of 31 well-conserved chromosomal segments corresponding to 31 chromosomes of another butterfly, *Melitaea cinxia* (*Heliconius* Genome Consortium, 2012; Ahola *et al.*, 2014). This hypothesis would be greatly strengthened if other species of $n=31$ karyotype that are distantly related to *M. cinxia* are shown to have similarly organized chromosomes.

The genus, *Ostrinia*, belonging to the superfamily Pyraloidea, which is distantly related to *B. mori* and butterflies (Regier *et al.*, 2013), is widely utilized for analyzing evolution in female sex pheromone

[1]Insect Genome Research Unit, National Institute of Agrobiological Sciences, Tsukuba, Ibaraki, Japan; [2]Faculty of Agriculture, Iwate University, Ueda, Morioka, Japan; [3]Graduate School of Agriculture, Hokkaido University, Sapporo, Japan and [4]Department of Agricultural and Environmental Biology, Graduate School of Agricultural and Life Sciences, The University of Tokyo, Tokyo, Japan
[5]Present address: Faculty of Education, Ehime University, Matsuyama 790–8577, Japan.
[6]Present address: Breeding and Pest Management Division, NARO Institute of Fruit Tree Science, Tsukuba, Ibaraki 305-8605, Japan.
Correspondence: Dr Y Yasukochi, Insect Genome Research Unit, National Institute of Agrobiological Sciences, Owashi 1-2, Tsukuba, Ibaraki 305-8634, Japan.
E-mail: yaskoch@affrc.go.jp
Received 18 May 2015; revised 29 June 2015; accepted 29 June 2015; published online 12 August 2015

biosynthesis and male recognition systems during speciation (for review, see Lassance, 2010). Among them, the Asian and European corn borers, *Ostrinia furnacalis* ($n=31$) and *O. nubilalis* ($n=31$), are serious pests of maize. Although a linkage map was constructed for *O. nubilalis* based on amplified fragment length polymorphism and microsatellites, not many genes were mapped on it (Dopman *et al.*, 2004; Streiff *et al.*, 2014). Fine mapping of *O. nubilalis* genes is useful not only for characterization of the genetical basis of complex phenotypes but also for verification of conserved synteny. We previously succeeded in identifying three chromosomes of *O. nubilalis* orthologous to *B. mori* chromosomes 1(Z), 16 and 23 by FISH analysis (Yasukochi *et al.*, 2011a).

Here, we describe construction of a cytogenetic map covering all 31 chromosomes of *O. nubilalis* using 122 bacterial artificial chromosomes (BACs) and eight fosmid clones as FISH probes. We then focused this study on the identification of boundaries between ancient chromosomes within *B. mori* chromosomes 11, 23 and 24. These regions have a highly distinctive feature of proximal and distal localization in *B. mori* and *O. nubilalis* chromosomes, which facilitated the detection of incorrectly isolated probes and improved accuracy of FISH analysis.

## MATERIALS AND METHODS

### Insects

The Z-race of *O. nubilalis,* originally collected in Darmstadt, Germany, was used in this study. *O. furnacalis*, O. *scapulalis, O. latipennis, O. palustralis* and *Pleuroptya ruralis* were from wild populations collected in the Kanto area, Japan. All species were fed an artificial diet, Insecta LFS (Nosan, Yokohama, Japan), in plastic cups. *B. mori* strain p50 was used and reared on mulberry leaves.

### Genomic sequencing of *O. furnacalis*

Genomic DNA was prepared from an *O. furnacalis* male pupa that was a subcultured offspring of wild insects originally collected in Inawashiro, Fukushima, Japan, using a DNeasy Blood & Tissue Kit (Qiagen, Venlo, the Netherlands) according to the manufacturer's protocol. The genomic DNA was then submitted to Otogenetics Corporation (Norcross, GA, USA) for sequencing. An Illumina paired-end library with an average insert size of 196 bp was made from fragmented DNA using NEBNext reagents (New England Biolabs, Ipswich, MA, USA), and sequenced on an Illumina HiSeq 2000 which generated paired-end reads of 100 nucleotides.

### Genome assembly

The raw paired-end reads were trimmed and filtered by Trimmomatic (Bolger *et al.*, 2014) and reads of low quality and those contaminated with adapter sequences (approximately 15%) were removed. Then, genome assembly and scaffolding were carried out using SOAPdenovo2 (version r240) (Luo *et al.*, 2012) with k-mer length 75.

### Construction of an *O. furnacalis* fosmid library

A fosmid library was constructed by Takara Bio (Kyoto, Japan). Genomic DNA was purified from whole male pupae of *O. furnacalis* by the cetyl trimethyl ammonium bromide method (Murray and Thompson, 1980). The DNA was fragmented by sonication and fractionated by pulse-field electrophoresis. Genomic fragments ranging from 33 to 48 kb were purified from agarose gels and cloned into a pCC1FOS vector (Epicentre, Madison, WI, USA). A total of 75 264 single colonies were arrayed into 196 384-well microplates.

### Screening of BAC and fosmid libraries

An *O. nubilalis* BAC library, ON_Ba, was obtained from the Clemson University Genomics Institute (Clemson, SC, USA). Construction of a *B. mori* BAC library is described elsewhere (Wu *et al.*, 1999). Detailed methods for PCR screening of large-insert libraries and FISH analysis are described in Yoshido *et al.* (2014). Briefly, the first screening was carried out against DNA pools representing each microplate containing 384 clones. A second screening was then carried out against DNA pools prepared from the mixture of cell cultures located in the same column or row of each positive microplate. This was followed by single colony isolation from frozen stocks of candidate clones and overnight cultures of the isolated clones for final PCR confirmation. PCR conditions were as follows: 3-min denaturation at 94 °C, followed by 45 cycles with a 1-min denaturation at 94 °C, 2-min annealing at 55 °C, and 3-min elongation at 72 °C, ending with a 5-min final extension at 72 °C. The sequences of primers used in this study are listed in Supplementary Table S1.

### FISH analysis

Chromosome preparation was carried out following the method of Sahara *et al.* (1999) with slight modifications as described in Sahara *et al.* (2013). BAC- or fosmid-DNA extraction and probe labeling were carried out according to the descriptions in Yasukochi *et al.* (2009),Yoshido *et al.* (2011) and Sahara *et al.* (2013). A published reprobe technique for Lepidoptera (Shibata *et al.,* 2009) was used to integrate results obtained by multiple hybridizations with different probes. All FISH and chromosome images were acquired with a DFC350FX CCD camera attached to a DM 6000B microscope (Leica Microsystems Japan, Tokyo, Japan) and processed with Adobe Photoshop ver 7.

### Genotyping

The $F_2$ progeny of *B. mori* and the $BF_1$ progeny between *O. nubilalis* and *O. scapulalis* used in this study were identical to those described in previous reports (Yasukochi *et al.*, 2006; Yasukochi *et al.*, 2011a). Genomic DNAs were individually amplified in the same manner as library screens. The completed reactions were denatured at 95 °C for 5-min, annealed at 55 °C for 15-min to promote heteroduplex formation, and loaded on 0.67 × MDE gels (Cambrex, Rockland, ME, USA) in 0.5 × TBE to detect polymorphisms.

### Search for telomere-associated sequences

*B. mori* genomic sequences (scaffold build2) position 5 400 000–5 800 000 of chromosome 11, 17 000 000–17 400 000 of chromosome 23, and 9 000 000–11 000 000 of chromosome 24, were obtained from a *B. mori* genome database, Kaikobase (Shimomura *et al.*, 2009). The sequences were then searched for 1690 previously identified repeat sequences including 868 transposable elements (Osanai-Futahashi *et al.*, 2008) using the repeat search program, RepeatMasker Open-4.0.5 (Smit *et al.*, 2013–2015).

## RESULTS AND DISCUSSION

### FISH analysis using previously isolated *O. nubilalis* BAC clones

We had previously isolated 163 *O. nubilalis* BAC clones containing single-copy genes which could be used as anchors for comparing genomes (Yasukochi *et al.*, 2011b). In addition, we isolated two BACs, 04I15 and 04K09, which contained *O. nubilalis* orthologues of *B. mori* genes located on the distal part of *B. mori* chromosome 24. Thus, we used these BACs as probes for FISH analysis of *O. nubilalis*. Subsequently, we obtained stable and reproducible signals from 112 BACs harboring 117 conserved genes between *O. nubilalis* and *B. mori* (Supplementary Table S2), which could identify 29 of 31 *O. nubilalis* chromosomes (Figure 1). The gene order on each chromosome was well conserved between *O. nubilalis* and *B. mori* with few exceptions (Figure 1). We designated the same chromosome numbers, 1–31 excluding 11 and 30, as those of orthologous chromosomes in previously characterized species of $n=31$ karyotype (Baxter *et al.*, 2011; Sahara *et al.,* 2013; Van't Hof *et al.,* 2013).

Mating experiments between species or populations belonging to the genus *Ostrinia* have been performed to map traits that might drive speciation by changes in behavior and pheromone composition (Dopman *et al.*, 2005; Streiff *et al.*, 2014; Wadsworth *et al.*, 2015). However, genetic recombination in regions of interest was reported to be suppressed, presumably because of an inversion occurring on the Z chromosome of *O. nubilalis* (Wadsworth *et al.*, 2015). Suppression of genetic recombination makes linkage analysis difficult; and therefore,

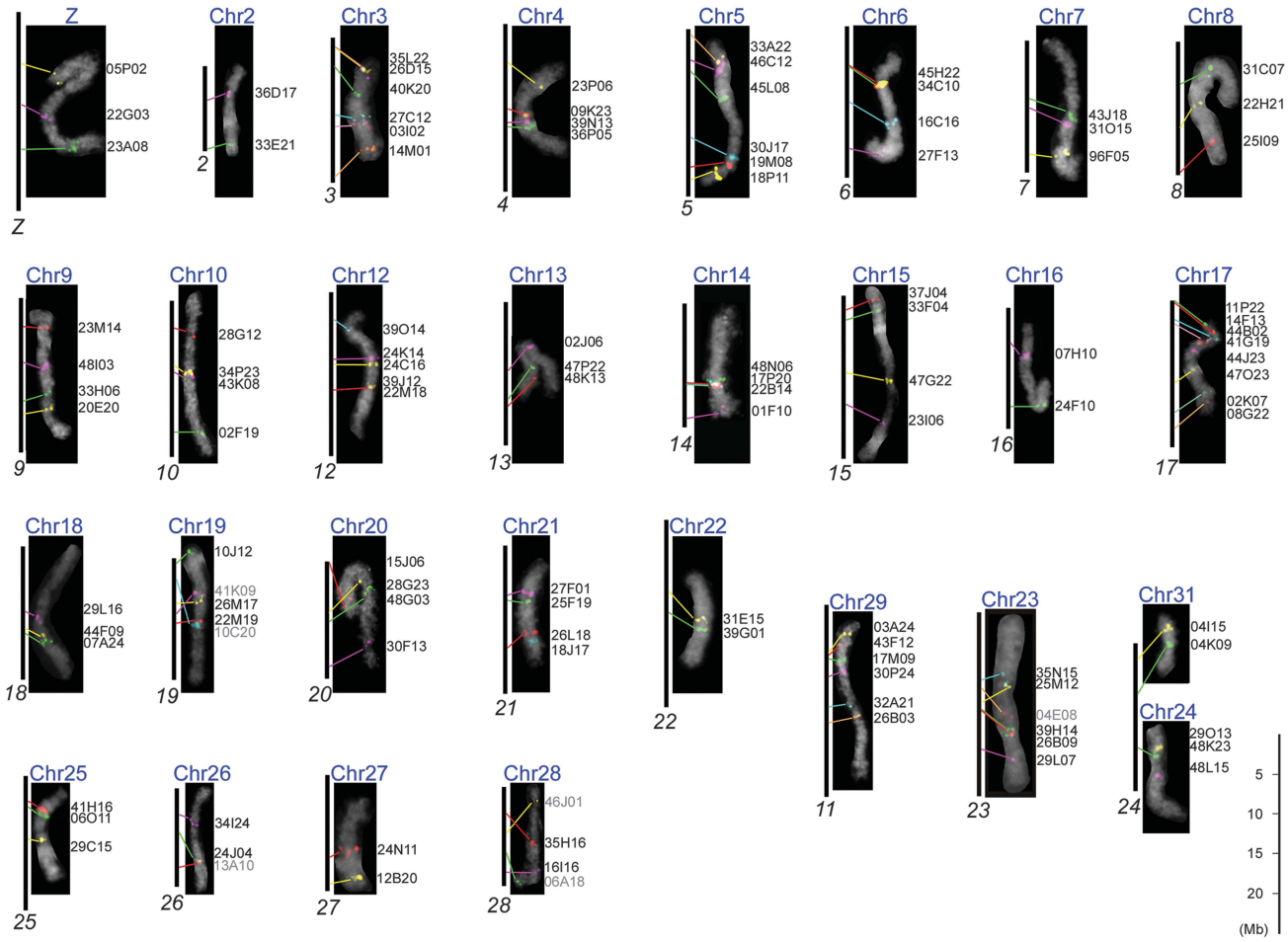


**Figure 1** Comparison of orthologues between individual chromosomes of *O. nubilalis* and *B. mori*. 4',6-diamidino-2-phenylindole-stained images (gray pseudo color) of individual pachytene bivalents of *O. nubilalis* (blue chromosome numbers) show hybridization signals from *O. nubilalis* BAC probes. Black vertical bars represent orthologous *B. mori* chromosomes (italic black numbers) based on locations taken from Kaikobase (http://sgp.dna.affrc.go.jp/KAIKObase/). *O. nubilalis* BAC designations are shown on the right of each chromosome image. Gray letters indicate intrachromosomal rearrangement of localized genes and lack of conserved gene order between *O. nubilalis* and *B. mori*. Colored lines show correspondence of orthologues between *B. mori* and *O. nubilalis* based on the carrier BAC signals. Note that images of *O. nubilalis* chromosomes do not reflect true relative sizes because they were taken from different preparations.

we investigated the possibility of performing cross-species FISH analysis which is able to detect chromosome rearrangements peculiar to species or populations. We used four *Ostrinia* species (*O. scapulalis, O. furnacalis, O. latipennis* and *O. palustralis*), as well as *P. ruralis*, which belongs to the same subfamily, Pyraustinae. Of the four *Ostrinia* moths, *O. scapulalis* is most closely related and *O. latipennis* and *O. palustralis* are most distantly related to *O. nubilalis* (Kim *et al.*, 1999). As shown in Figure 2, using three common *O. nubilalis* BAC probes previously located on chromosome 5, we obtained clear signals in the same order on a single chromosome of these five species.

### Sequencing of the *O. furnacalis* genome

As described above, we did not have any BAC probes ostensibly located on chromosomes 11 and 30. For BAC isolation, we had utilized sequences of known genes and expressed sequence tag in public databases or newly determined sequences of PCR fragments amplified by nested degenerate primers designed from conserved amino acid sequences (Yasukochi *et al.*, 2011b). However, because genes expressed in limited stages and tissues are rarely included in expressed sequence tag, it was very difficult to find an adequate number of conserved genes on such narrow chromosomal regions having no markers. In addition, cDNA sequences lack sequences for introns, which makes it difficult to design PCR primers and to judge whether amplified fragments are genuine or non-specific.

Thus, we decided to obtain genome sequences of an *Ostrinia* species, which would greatly facilitate finding targeted genes and establishing genetic markers in targeted locations. Because *O. nubilalis* does not occur in Japan, strict escape prevention measures are required for rearing it. Consequently, we used a local congener, *O. furnacalis*, as a substitute. We sequenced a paired-end library (average insert size 196 bp) constructed from an *O. furnacalis* male pupa and obtained 511 Mb assembled sequences comprising 745 531 scaffolds and singleton contigs (N50 = 930 bp) (Supplementary Table S3).

### Screening of *O. furnacalis* orthologues

Supplementary Table S4 lists lepidopteran genes that were previously found to be located near the regions orthologous to the boundaries between ancient chromosomes in *B. mori* when we embarked on the second-phase library screening aimed at these regions (Beldade *et al.*, 2009; Yasukochi *et al.*, 2009; Van't Hof *et al.*, 2013;

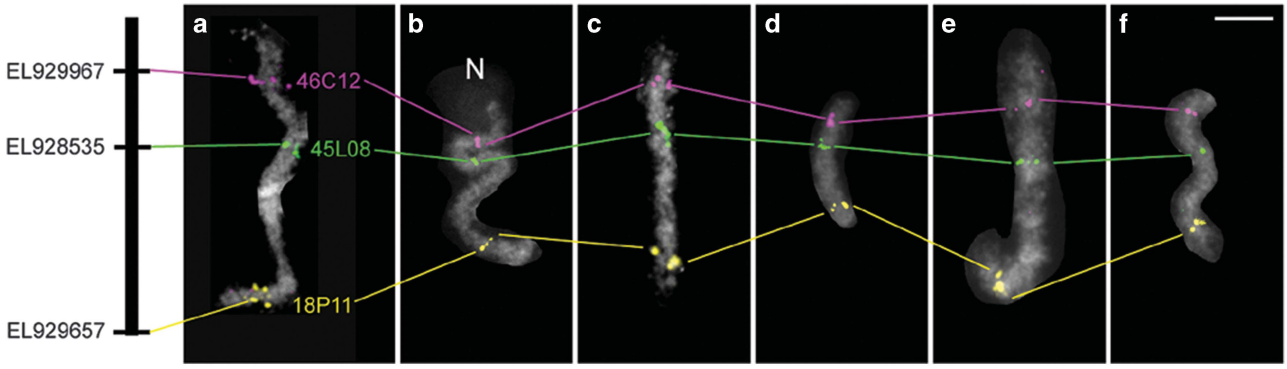


**Figure 2** Conserved gene order on chromosome 5 of five *Ostrinia* species and a close relative based on cross-hybridization of three representative BACs from *O. nubilalis*. (**a**) *O. nubilalis*; (**b**) *O. furnacalis*; (**c**) *O. scapulalis*; (**d**) *O. latipennis*; (**e**) *O. palustralis*; and (**f**) *P. ruralis*. Signals from BAC probes 46C12, 45L08 and 18P11, which carry orthologues EL92996, EL928535 and EL929657, are pseudocolored with magenta, green and yellow, respectively. The black bar represents *B. mori* chromosome 5. N, nucleolus. Scale bar, 5 μm.

**Table 1 Summary of the results onbtained by FISH and genetic analysis in this study**

| B. mori | | | | | *Accession No. of O. furnacalis orthologues* | O. nubilalis *Chr.* | | O. nubilalis *BAC* | O. furnacalis *fosmid* |
|---|---|---|---|---|---|---|---|---|---|
| *Gene* | *Chr.* | *Bm_scaf* | *Location in Kaikobase* | *BAC* | | *Genotyping* | *FISH* | | |
| NM_001043856.1 | 11 | 35 | 2426352-2420617 | | LC002981 | 11 | 11 | 32O01 | |
| BGIBMGA001828 | 11 | 35 | 4003867-4006125 | | LC002982 | | 11 | | 49B19 |
| AK383118 | 11 | 59 | 4988171-4987961 | | LC002983 | 23 | 23 | 02B23 | |
| AK385284 | 11 | 59 | 5050845-5051092 | | LC002984 | 11 | | | |
| AK378226 | 11 | 59 | 5110274-5114967 | | LC002985 | 11 | 11 | 08O23 | |
| NM_001173138.1 | 11 | 59 | 5302237-5342471 | | LC002986 | 11 | | | 03G23 |
| AK385912 | 11 | 59 | 5436505-5445770 | | LC002987 | | 11 | | 13G13 |
| AK378774 | 11 | 59 | 5455029-5460410 | | LC002988 | 11 | | | |
| BGIBMGA011986 | 11 | 59 | 5474195-5467027 | | LC002989 | 11 | | | 07M24 |
| BGIBMGA011911 | 11 | 59 | 5565779-5572726 | | LC002990 | 29 | | | |
| AK378190 | 11 | 59 | 5748073-5732931 | | LC002991 | 29 | 29 | 02G09 | |
| AK384105 | 11 | 59 | 5778613-5759910 | | LC002992 | 29 | | | |
| AK381196 | 11 | 59 | 6133070-6140746 | | LC002993 | | 29 | 32C13 | |
| NM_001043886.1 | 23 | 12 | 17064070-17064217 | | LC002994 | 23 | 23 | 03N16 | 35A13 |
| AK378672 | 23 | 12 | 17155881-17150565 | | LC002995 | | 23 | | |
| AK383286 | 23 | 12 | 17195618-17248107 | | LC002996 | | 30 | | 76F08 |
| AK385089 | 23 | 12 | 17315309-17309972 | | LC002997 | 30 | | | 09P08 |
| AK378640 | 23 | 12 | 17328996-17334405 | | LC002998 | 30 | | | |
| AK384547 | 23 | 12 | 17408923-17419988 | | LC002999 | | | | 63L08 |
| AK386411 | 23 | 12 | 17506663-17510276 | | LC003000 | 30 | | | 39I14 |
| NM_001043790.1 | 23 | 12 | 21055045-21060438 | | LC003001 | | 30 | | 14J14 |
| NM_001135885.1 | 24 | 43 | 7589049-7590339 | | LC003002 | | 31 | | 01J24 |
| NM_001173167.1 | 24 | 43 | 9082682-9097385 | | LC003003 | | 31 | 04D21 | |
| AK378694 | 24 | 43 | 9124973-9128620 | | LC003004 | | 31 | 26P11 | |
| AK386872 | 24 | 43 | 9232080-9237493 | 17J12F | LC003005 | | 31 | | 06H23 |
| AK385865 | 24 | 103 | 10769257-10780986 | 19D05C | LC003006 | 24 | 24 | | 36F10 |
| AK377444 | 24 | 103 | 10967433-10954881 | | LC003007 | 24 | 24 | 09N16 | |
| AK380262 | 24 | 103 | 11063900-11060287 | | LC003008 | 24 | | | |
| NM_001046732.1 | 24 | 103 | 11454798-11469070 | | LC003009 | 24 | | | 04O02 |
| AK385042 | 24 | 208 | 12462114-12471476 | | LC003010 | | 24 | 18O24 | |
| NM_001043588.1 | 24 | not mapped in Kaikobase | | 04L10D | | | | | |
| BP125410 | 24 | 75 | 17956279-17955584 | 04A07H | | | | | |

Abbreviations: BAC, bacterial artificial chromosome; FISH, fluorescence *in situ* hybridization.

Sahara *et al.*, 2013). The boundaries were predicted to be located at position 4 006 128–6 271 907 of chromosome 11, 17 071 002–17 410 092 of chromosome 23, and 8 719 098–11 527 820 of chromosome 24 in Kaikobase. Subsequently, we obtained full-length cDNAs, gene models and known genes that were predicted to localize in these regions from Kaikobase. In addition, we selected several *B. mori* sequences located at position 1–4 006 128 of chromosome 11 and position 17 410 092–23 133 604 of chromosome 23, which correspond to unidentified *O. nubilalis* chromosomes 11 and 30.

Candidate regions of ancient chromosomal fusions were located within single scaffolds, Bm_scaf59 and Bm_scaf12, of *B. mori* chromosomes 11 and 23. In contrast, the candidate region in *B. mori* chromosome 24 was separated into three scaffolds. Furthermore, we could not find any conserved genes in the middle scaffold, Bm_scaf147 (position 9 817 336–10 242 811), because no known genes or full-length cDNAs were located there, and gene models and transcripts assigned to the scaffold did not show significant similarities to available *O. furnacalis* genome sequences. Ultimately, we utilized 43 *B. mori* sequences as queries for a tblastn search against the *O. furnacalis* genome sequences. Thirty of the query sequences showed striking similarities to *O. furnacalis* scaffolds and contigs containing well-conserved single-copy genes (Table 1).

### Screening of *O. nubilalis* BAC and *O. furnacalis* fosmid libraries

We then designed 30 pairs of PCR primers to isolate genomic clones containing candidate genes (Supplementary Table S1). In general, BACs are preferable as FISH probes compared with fosmids, presumably because longer inserts can label longer chromosomal regions more stably. In addition, the chromosome preparations used for this analysis were derived from *O. nubilalis*, which might require longer regions to form stable hybridization signals with the heterologous *O. furnacalis* probes.

Thus, we first screened the *O. nubilalis* BAC library, followed by screening of a newly constructed fosmid library from *O. furnacalis* pupae from which we obtained 75 264 colonies (average insert size, 33.1 kb; total estimate size, 2491 Mb). Consequently, we isolated 10 BACs (Table 1) from the *O. nubilalis* BAC library, and then isolated 14 fosmids (Table 1) by screening the *O. furnacalis* fosmid library with markers that had failed to isolate positive *O. nubilalis* BACs.

### FISH and genetic analysis

We performed FISH analysis against *O. nubilalis* chromosomes using the 24 BAC and fosmid probes described above. In parallel with the FISH analysis, we performed genetic analysis of the candidate genes using 24 $BF_1$ progeny between *O. nubilalis* and *O. scapulalis* to validate and complement the FISH results, because we could not isolate BACs or fosmids for six genes (Table 1).

We obtained clear signals from five BACs and two fosmids containing *Ostrinia* orthologues of *B. mori* genes located on chromosome 11. Two BACs, 02G09 and 32C13, which had been assigned to previously identified *O. nubilalis* chromosome 29, and signals from two BACs (08O23, 32O01) and two fosmids (13G13, 49B19), were mapped onto the same chromosome (Figure 3a). Thus, we designated this newly identified chromosome as *O. nubilalis* chromosome 11. In addition, we obtained polymorphic PCR products from nine orthologues that were mapped onto *O. nubilalis* chromosomes 11 or 29, as expected (Table 1, Supplementary Table S5). Taking all these data together, the boundary between ancestral chromosomes 11 and 29 on *B. mori* chromosome 11 was narrowed to the region between gene models BGIBMGA011986 and BGIBMGA011911 (Table 1, Figure 4a).

Unexpectedly, BAC 02B23, which contained an *O. nubilalis* orthologue (LC002983) of a *B. mori* full-length cDNA, AK383118, was localized on another chromosome, excluding it as belonging to *O. nubilalis* chromosome 11 or 29. Genetic analysis revealed that this *O. nubilalis* orthologue was located on *O. nubilalis* chromosome 23 (Table 1, Supplementary Table S5). However, there remained the possibility that AK383118 was incorrectly assembled into Bm_scaf59. Thus, we performed linkage mapping of AK383118 in *B. mori* and confirmed its location on *B. mori* chromosome 11 (data not shown).

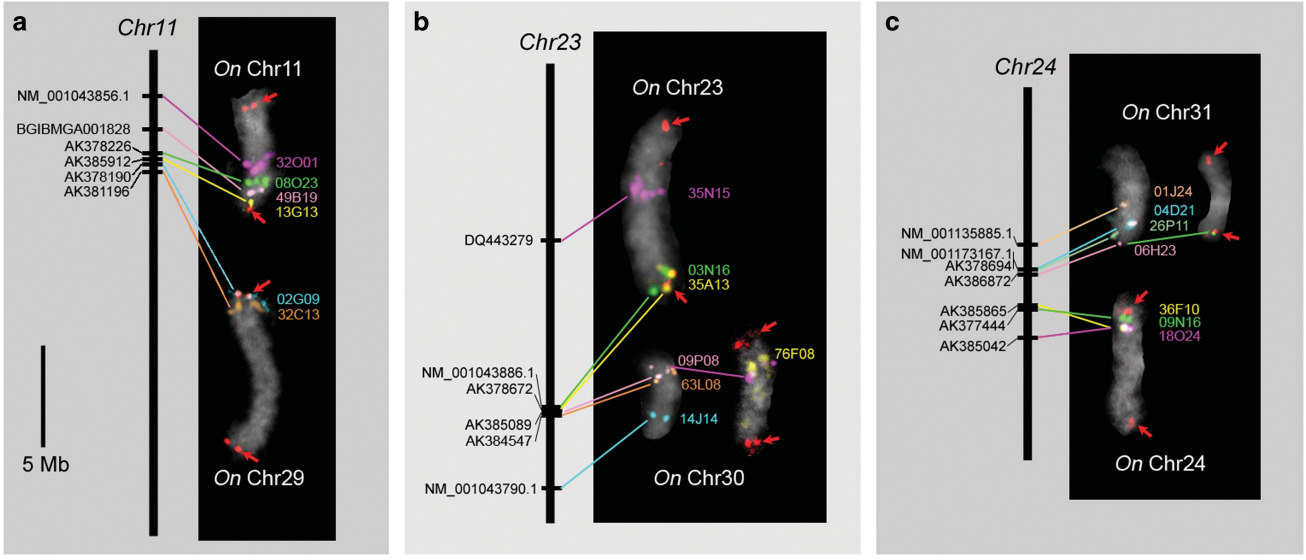


**Figure 3** FISH detection of *O. nubilalis* chromosomes 11 and 29 (**a**), 23 and 30 (**b**) and 24 and 31 (**c**). Signals on the *O. nubilalis* chromosome (white numbers) are pseudocolored and probe names are shown to the right of the FISH images. See Table 1 for details. Arrows point to (TTAGG)n signals pseudocolored red. Note that left images of chromosomes 30 and 31 were not hybridized with (TTAGG)n probes. Left-side bars represent orthologous *B. mori* chromosomes (italic black numbers). Locations of *B. mori* orthologues are taken from Kaikobase. FISH images are from different individuals and/or preparations. Hence, *O. nubilalis* bivalent lengths vary depending on the pachytene stage.

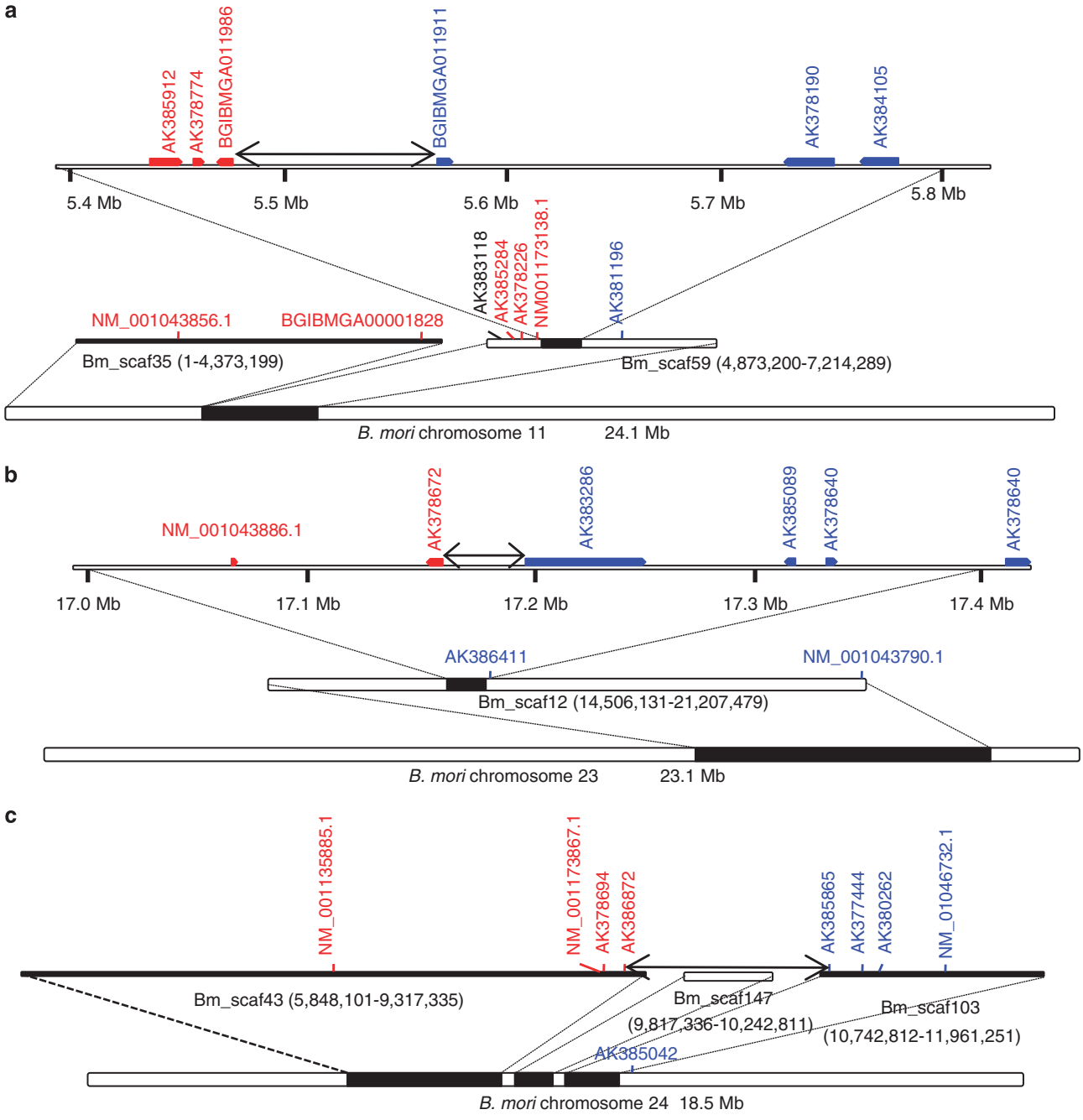


**Figure 4** Candidate regions of ancient chromosomal fusions in *B. mori* chromosomes 11 (**a**), 23 (**b**) and 24 (**c**). Candidate regions are shown as two-headed arrows. Vertical letters indicate *B. mori* genes for which orthologues are located on *O. nubilalis* chromosomes 11 (red letters, **a**), 23 (red letters, **b**), 24 (red letters, **c**), 29 (blue letters, **a**), 30 (blue letters, **b**) and 31 (blue letters, **c**). Letters in parentheses indicate locations of *B. mori* scaffolds in Kaikobase.

This was the only example for which we found evidence suggesting an ancient translocation.

Similarly, FISH analysis identified missing *O. nubilalis* chromosome 30 (Figure 3b) on *B. mori* chromosome 23, where the boundary between ancestral chromosomes 23 and 30 was estimated to be localized between full-length cDNAs AK378672 and AK383286 (Table 1, Figure 4b). Finally, *O. nubilalis* orthologues of *B. mori* genes on Bm_scaf43 and Bm_scaf103 were assigned to *O. nubilalis* chromosomes 31 and 24, respectively (Figures 3c and 4c).

Intervals between *B. mori* scaffolds in Kaikobase are not experimentally confirmed but are designated arbitrarily as 500 kb (http://sgp.dna.affrc.go.jp/KAIKObase/). Hence, we suspected that the distance between Bm_scaf43 and Bm_scaf103 was much shorter than the estimated length of 1.5 Mb from the genome database. Conversely, assigning unmapped contigs located in this gap might enlarge the actual distance. Thus, we isolated *B. mori* BACs, 17J12F and 19D5C, containing the most distal full-length cDNAs of the scaffolds, AK386872 and AK385865, respectively (Table 1). FISH probing with

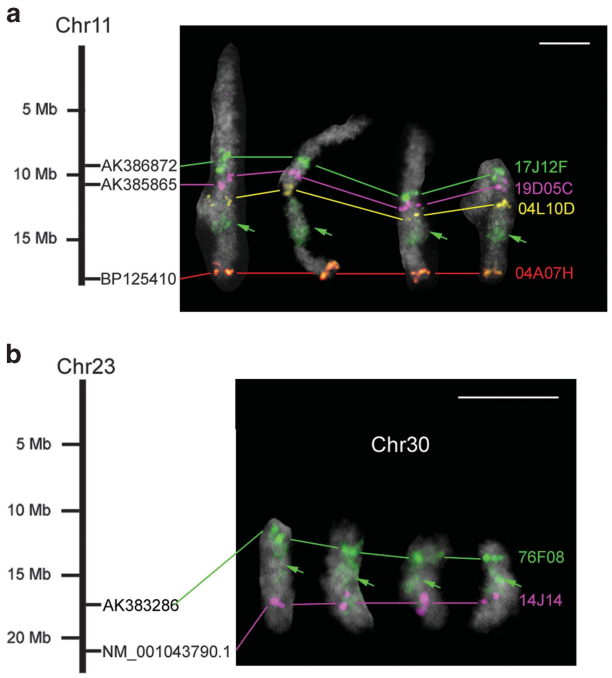


**Figure 5** (**a**) A sequence gap between Bm_scaf43 and Bm_scaf 103 of *B. mori* chromosome 11 revealed by BAC-FISH. Signals from 17J12F (green), 19D05C (magenta) 04L10D (yellow) and 04A07H (red) represent AK386872, AK385865, NM_001043588.1 and BP125410, respectively. Note that the green signals from 17J12F faintly stain the heterochromatic region (refer to Figure 2 in Yoshido *et al.* 2005) of *B. mori* chromosome 11 (green arrows). (**b**) Ectopic signals from an *O. furnacalis* fosmid (76F08) located near the end of chromosome 30 appear in *O. nubilalis* chromosome 30 (Chr30 shown in white letters). Two fosmids, 76F08 and 14J14, carrying representative single genes, LC002996 and LC003001, are orthologues of *B. mori* AK383286 and NM_001043790.1, respectively (see Table 1). Black bars represent *B. mori* chromosomes 11 and 23 (Chr11 and Chr23). Scale bars, 5 μm.

these clones suggested that the interval between Bm_scaf43 and Bm_scaf103 was longer than 1 Mb, although the estimate varied depending on the state of chromosome condensation (Figure 5).

Intriguingly, the 17J12F probe also painted a segment of chromosome 24 (Figure 5a) which was positionally consistent with a previously described 4',6-diamidino-2-phenylindole-positive and presumably heterochromatic segment (Yoshido *et al.*, 2005). Similar ectopic painting was also observed for *Ostrinia* BACs and fosmids putatively located near the ends of chromosomes (Figure 5b). A possible explanation is that these clones contain subtelomeric heterochromatin and tend to label heterochromatin scattered around chromosomes. If so, 17J12F may contain subtelomeric heterochromatin that originated from an ancient chromosomal end.

### Comparison with the *M. cinxia* genome data

Judging from the results described above, we estimated the positions of boundaries between ancient chromosomes in *B. mori* chromosomes 11, 23 and 24 as 5 474 195–5 565 779, 17 155 881–17 195 618 and 9 237 493–10 769 257, respectively. In the proceeding work comparing *B. mori* and *M. cinxia* genomes, these boundaries were estimated to be 5 410 160–5 798 000 in chromosome 11, 17 180 240–17 247 200 in chromosome 23, and 9 953 070–10 742 570 in chromosome 24 (Ahola *et al.*, 2014). Our estimates based on FISH and genetic analysis of *O. nubilalis* conserved genes is in good accordance with those based on *M. cinxia* genome sequences in spite of the distant phylogenetic relationships between the two species. Taken together, these data provide strong evidence to support the hypothesis that species having 31 chromosomes per haploid genome basically retain the ancestral karyotype of Lepidoptera.

### Search for telomere-associated sequences

The telomeres of *B. mori* are composed of pentanucleotide telomeric repeats, (TTAGG)n, and several families of long interspersed elements (Okazaki *et al.*, 1993, 1995). The TTAGG telomeric repeats are conserved among all arthropod groups, although lacking in spiders and some insect orders (Okazaki *et al.*, 1993; Sahara *et al.*, 1999; Vítková *et al.*, 2005). Thus, we searched five *B. mori* scaffolds located on the boundaries of putative chromosomal fusions for TTAGG repeats. Many of the repeat-like sequences we found were short and incomplete (Supplementary Table S6), but a complete 28-mer repeat was located adjacent to a 117 bp-gap on Bm_scaf59 (chr. 11: position 5 933 800–5 933 773), raising the possibility that this repeat might be longer (Supplementary Table S6).

We also searched for long interspersed elements, especially the R1 clade which includes telomeric repeat-specific TRAS family and SART1 retrotransposons (Kubo *et al.*, 2001; Osanai-Futahashi *et al.*, 2008). The proportion of long interspersed elements at the boundary region in *B. mori* chromosome 11 was much higher than the estimated average of the whole genome, 13.8% (Osanai-Futahashi *et al.*, 2008). In addition, we found retrotransposon-derived sequences similar to TRAS3 and SART1, as well as clear peaks in the proportion of R1 clade non-LTRs in this region (Figure 6). Intriguingly, a TTAGG repeat was neighboring the TRAS3-like sequence (Figure 6), which might be another vestigial feature of former subtelomeric heterochromatin.

TTAGG repeats, TRAS3-like sequences and accumulation of the R1 clade long interspersed elements were detected not in the candidate region but in the surrounding region of the boundary in *B. mori* chromosome 23 (Figure 6). The region between the signals of telomeric repeats and 76F08 on *O. nubilalis* chromosome 23 was relatively long compared with other chromosomal ends (Figures 3b and 5b), which might mean that internal deletion and translocation of the boundary region occurred after the fusion event generating *B. mori* chromosome 23. Taken together, we speculate that the fusion event generating *B. mori* chromosome 23 occurred before those generating chromosome 11.

In contrast, neither telomeric repeat-specific retrotransposons nor TTAGG repeats were found within or around the boundary in *B. mori* chromosome 24 (Figure 6). As described above, the maximum estimate of the boundary between ancient chromosomes is much longer in *B. mori* chromosome 24 (1532 kb) compared with the estimates for chromosomes 11 (91.6 kb) and 23 (39.7 kb) (Figure 4). As many long gaps in the sequence assembly still remain, including two inter-scaffold ones in this region, we were unable to exclude the possibility that sequences harboring telomeric features are located in these gaps.

Further analysis is needed to understand sequential changes that have occurred in the former chromosomal ends after fusion events. Significant stability of lepidopteran chromosomes enabled us to distinguish which portion fused chromosomes were derived from. Diverged clades with a reduced and variable chromosome number are adequate for this purpose because many boundaries at different temporal stages are available and the chronological order of fusion events can be determined to some extent by detection of shared fused chromosomes among closely or distantly related species. For instance, we previously compared subspecies of a wild silkmoth, *Samia cynthia* (2n = 25–28), with *B. mori*, and identified more than 15 boundaries between ancestral chromosomes excluding 24 and 30 (Yoshido *et al.*, 2011).

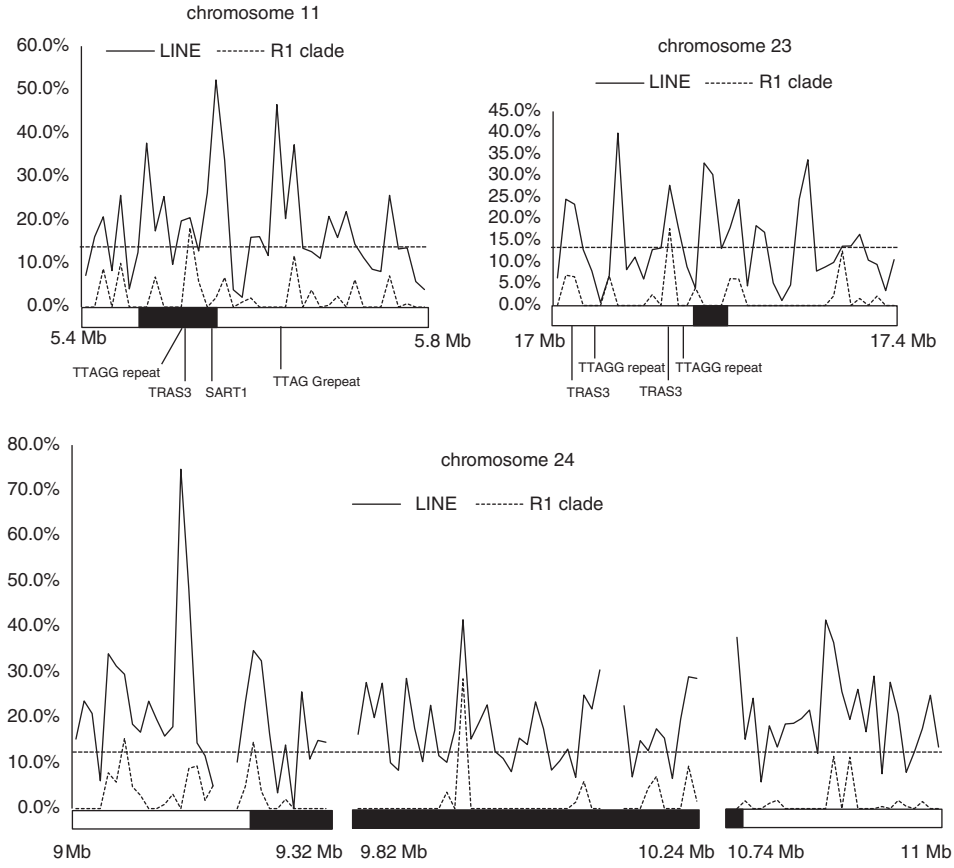


**Figure 6** Distribution of long interspersed elements (LINEs) and telomeric repeats around the candidate regions of ancient chromosomal fusions. Genomic sequences with gaps were divided into 10-kb windows, and the ratio of the length of all (solid lines) or R1-clade (dotted lines) LINEs in total genomic sequence without gaps was calculated for each window. Filled squares indicate candidate regions. Horizontal dotted lines indicate 13.8%, average of LINE content in the *B. mori* genome (Osanai-Futahashi *et al.*, 2008). Vertical lines indicate insertion sites of a TTAGG repeat and truncated copies of TTAGG repeat-specific retrotransposons, TRAS3 and SART1.

Current knowledge about shared gene order in Lepidoptera is limited to the clade Ditrysia, which constitutes 98% of lepidopteran species (Regier *et al.*, 2013), and little is known about karyotype evolution in more primitive groups. Additionally, Trichoptera is not only most closely related with Lepidoptera among extant insect orders (Wiegmann *et al.*, 2009), but also shares unique features such as a sex determination system with female heterogamety and holocentric chromosomes. Analysis of basal non-ditrysian species and Trichoptera is critical to reveal the origin of the $n=31$ karyotype and assess the stability of lepidopteran chromosomes.

## CONCLUSION

To establish new markers for a FISH-based chromosome map of *O. nubilalis* by efficient screening of *Ostrinia* orthologues of *B. mori* genes located on targeted chromosomal regions lacking markers, we determined and assembled 511 Mb genome sequence and constructed a fosmid library of *O. furnacalis*. Consequently, we constructed a cytogenetic map of *O. nubilalis* consisting of 122 BAC and eight fosmid probes covering all 31 chromosomes. This map can be applied to other *Ostrinia* moths and related species, and will be useful for the detection of local chromosomal rearrangements that have occurred uniquely in particular species where genetic recombination is suppressed. The targeted selection of BAC and fosmid clones reported here is especially effective for rapid determination of extensive genomic sequences or genome-wide characterization of multiple-copy genes in organisms lacking genome reference sequences.

We estimated the boundaries between ancient chromosomes on *B. mori* chromosomes 11, 23 and 24 by mapping *O. nubilalis* orthologues, which agreed well with a previous estimation based on sequence comparison between *B. mori* and *M. cinxia*. These findings strongly suggest that lepidopteran species having 31 chromosomes per haploid genome basically retain the ancestral karyotype of Lepidoptera.

## DATA ARCHIVING

Sequence data listed in Table 1 have been submitted to the DDBJ/EMBL/GenBank databases under the accession numbers LC002981–LC003010. Sequence scaffolds and singleton contigs used in this study are available from the Dryad Digital Repository: http://dx.doi.org/10.5061/dryad.q2k15.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## ACKNOWLEDGEMENTS

We are grateful to Dr M. R. Goldsmith for critical reading of the manuscript. We thank Aya Kawai and Noriko Fujikawa-Kojima for their technical assistance in FISH analysis. A part of this study is financially supported by the Program for Promotion of Basic Research Activities for Innovative Biosciences

(PROBRAIN) and grants 22380040, 23248008, 23380030, 25292204 of Japan Society for the Promotion of Science (JSPS).

Ahola V, Lehtonen R, Somervuo P, Salmela L, Koskinen P, Rastas P *et al.* (2014). The Glanville fritillary genome retains an ancient karyotype and reveals selective chromosomal fusions in Lepidoptera. *Nat Commun* **5**: 4737.

Alföldi J, Di Palma F, Grabherr M, Williams C, Kong L, Russell EP *et al.* (2011). The genome of the green anole lizard and a comparative analysis with birds and mammals. *Nature* **477**: 587–591.

Baxter SW, Davey JW, Johnston JS, Shelton AM, Heckel DG, Jiggins CD *et al.* (2011). Linkage mapping and comparative genomics using next-generation RAD sequencing of a non-model organism. *PLoS ONE* **6**: e19315.

Beldade P, Saenko SV, Pul N, Long AD (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.

Bolger AM, Lohse M, Usadel B (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics* **30**: 2114–2120.

Dopman EB, Bogdanowicz SM, Harrison RG (2004). Genetic mapping of sexual isolation between E and Z pheromone strains of the European corn borer (*Ostrinia nubilalis*). *Genetics* **167**: 301–309.

Dopman EB, Pérez L, Bogdanowicz SM, Harrison RG (2005). Consequences of reproductive barriers for genealogical discordance in the European corn borer. *Proc Natl Acad Sci USA* **102**: 14706–14711.

Heliconius Genome Consortium (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* **487**: 94–98.

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

Kemkemer C, Kohn M, Cooper DN, Froenicke L, Högel J, Hameister H *et al.* (2009). Gene synteny comparisons between different vertebrates provide new insights into breakage and fusion events during mammalian karyotype evolution. *BMC Evol Biol* **9**: 84.

Kim CG, Hoshizaki S, Huang YP, Tatsuki S, Ishikawa Y (1999). Usefulness of mitochondrial COII gene sequences in examining phylogenetic relationships in the Asian corn borer, *Ostrinia furnacalis*, and allied species (Lepidoptera: Pyralidae). *Appl Entomol Zool* **34**: 405–412.

Kubo Y, Okazaki S, Anzai T, Fujiwara H (2001). Structural and phylogenetic analysis of TRAS, telomeric repeat-specific retrotransposon family in lepidopteran insects. *Mol Biol Evol* **18**: 848–857.

Lassance JM (2010). Journey in the *Ostrinia* world: From pest to model in chemical ecology. *J Chem Ecol* **36**: 1155–1169.

Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J *et al.* (2012). SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *Gigascience* **1**: 18.

Murray MG, Thompson WF (1980). Rapid isolation of high molecular weight plant DNA. *Nucl Acids Res* **8**: 4321–4326.

Nene V, Wortman JR, Lawson D, Haas B, Kodira C, Tu ZJ *et al.* (2007). Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science* **316**: 1718–1723.

Okazaki S, Ishikawa H, Fujiwara H (1995). Structural analysis of TRAS1, a novel family of telomeric repeat-associated retroposons in the silkworm Bombyx mori. *Mol Cell Biol* **15**: 4545–4552.

Okazaki S, Tsuchida K, Maekawa H, Ishikawa H, Fujiwara H (1993). Ideintfication of a pentanucleotide telomeric sequence, (TTAGG)*n*, in the silkworm, *Bombyx mori* and other insects. *Mol Cell Biol* **13**: 1424–1432.

Osanai-Futahashi M, Suetsugu Y, Mita K, Fujiwara H (2008). Genome-wide screening and characterization of transposable elements and their distribution analysis in the silkworm Bombyx mori . *Insect Biochem Mol Biol* **38**: 1046–1057.

Pokorná M, Giovannotti M, Kratochvíl L, Caputo V, Olmo E, Ferguson-Smith MA *et al.* (2012). Conservation of chromosomes syntenic with avian autosomes in squamate reptiles revealed by comparative chromosome painting. *Chromosoma* **121**: 409–418.

Pringle EG, Baxter SW, Webster CL, Papanicolaou A, Lee SF, Jiggins CD (2007). Synteny and chromosome evolution in the Lepidoptera: Evidence from mapping in *Heliconius melpomene*. *Genetics* **177**: 417–426.

Regier JC, Mitter C, Zwick A, Bazinet AL, Cummings MP, Kawahara AY *et al.* (2013). A large-scale, higher-level, molecular phylogenetic study of the insect order Lepidoptera (moths and butterflies). *PLoS ONE* **8**: e58568.

Robinson R (1971). *Lepidoptera Genetics*. Pergamon: Oxford.

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, Shibata F, Fujikawa N, Okabe T, Tanaka-Okuyama M *et al.* (2013). FISH identification of *Helicoverpa armigera* and *Mamestra brassicae* chromosomes by BAC and fosmid probes. *Insect Biochem Mol Biol* **43**: 644–653.

Shibata F, Sahara K, Naito Y, Yasukochi Y (2009). Reprobing of multicolour FISH in preparations of lepidopteran chromosomes. *Zool Sci* **26**: 187–190.

Shimomura M, Minami H, Suetsugu Y, Ohyanagi H, Satoh C, Antonio B *et al.* (2009). KAIKObase: an integrated silkworm genome database and data mining tool. *BMC Genomics* **10**: 486.

Smit AFA, Hubley R, Green P (2013–2015). RepeatMasker Open-4.0.5 <http://www.repeatmasker.org>.

Streiff R, Courtois B, Meusnier S, Bourguet D (2014). Genetic mapping of two components of reproductive isolation between two sibling species of moths, *Ostrinia nubilalis* and *O. scapulalis*. *Heredity* **112**: 370–381.

Van't Hof AE, Nguyen P, Dalíková M, Edmonds N, Marec F, Saccheri IJ (2013). Linkage map of the peppered moth, *Biston betularia* (Lepidoptera, Geometridae): a model of industrial melanism. *Heredity* **110**: 283–295.

Vítková M, Král J, Traut W, Zrzavý J, Marec F (2005). The evolutionary origin of insect telomeric repeats. $(TTAGG)_n$. *Chromosome Res* **13**: 145–156.

Wadsworth CB, Li X, Dopman EB (2015). A recombination suppressor contributes to ecological speciation in *OSTRINIA* moths. *Heredity* **114**: 593–600.

Wiegmann BM, Trautwein MD, Kim JW, Cassel BK, Bertone MA, Winterton SL *et al.* (2009). Single-copy nuclear genes resolve the phylogeny of the holometabolous insects. *BMC Biol* **7**: 34.

Wu C, Asakawa S, Shimizu N, Kawasaki S, Yasukochi Y (1999). Construction and characterization of bacterial artificial chromosome libraries from the silkworm Bombyx mori. *Mol Gen Genet* **261**: 698–706.

Yasukochi Y, Ashakumary L, 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, Miura N, Nakano R, Sahara K, Ishikawa Y (2011a). Sex-linked pheromone receptor genes of the European corn borer, *Ostrinia nubilalis*, are in tandem arrays. *PLoS ONE* **6**: e18843.

Yasukochi Y, Tanaka-Okuyama M, Kamimura M, Nakano R, Naito Y, Ishikawa Y *et al.* (2011b). Isolation of BAC clones containing conserved genes from libraries of three distantly related moths: a useful resource for comparative genomics of Lepidoptera. *J Biomed Biotechnol* **2011**: 165894.

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 (2005). The *Bombyx mori* karyotype and the assignment of linkage groups. *Genetics* **170**: 675–685.

Yoshido A, Sahara K, Yasukochi Y (2014). Silk Moths (Lepidoptera). In: Sharakhov IV (ed). *Protocols for Cytogenetic Mapping of Arthropod Genomes*. CRC Press: Boca Ranton, FL, USA, pp 219–256.

Yoshido A, Yasukochi Y, Sahara K (2011). *Samia cynthia* versus *Bombyx mori*: comparative gene mapping between a species with a low-number karyotype and the model species of Lepidoptera. *Insect Biochem Mol Biol* **41**: 370–377.

Supplementary Information accompanies this paper on Heredity website (http://www.nature.com/hdy)