

# Chromosome Mapping of Dragline Silk Genes in the Genomes of Widow Spiders (Araneae, Theridiidae)

Yonghui Zhao\*, Nadia A. Ayoub<sup>‡</sup>, Cheryl Y. Hayashi

Department of Biology, University of California Riverside, Riverside, California, United States of America

## Abstract

With its incredible strength and toughness, spider dragline silk is widely lauded for its impressive material properties. Dragline silk is composed of two structural proteins, MaSp1 and MaSp2, which are encoded by members of the spidroin gene family. While previous studies have characterized the genes that encode the constituent proteins of spider silks, nothing is known about the physical location of these genes. We determined karyotypes and sex chromosome organization for the widow spiders, *Latrodectus hesperus* and *L. geometricus* (Araneae, Theridiidae). We then used fluorescence *in situ* hybridization to map the genomic locations of the genes for the silk proteins that compose the remarkable spider dragline. These genes included three loci for the MaSp1 protein and the single locus for the MaSp2 protein. In addition, we mapped a *MaSp1* pseudogene. All the *MaSp1* gene copies and pseudogene localized to a single chromosomal region while *MaSp2* was located on a different chromosome of *L. hesperus*. Using probes derived from *L. hesperus*, we comparatively mapped all three *MaSp1* loci to a single region of a *L. geometricus* chromosome. As with *L. hesperus*, *MaSp2* was found on a separate *L. geometricus* chromosome, thus again unlinked to the *MaSp1* loci. These results indicate orthology of the corresponding chromosomal regions in the two widow genomes. Moreover, the occurrence of multiple *MaSp1* loci in a conserved gene cluster across species suggests that *MaSp1* proliferated by tandem duplication in a common ancestor of *L. geometricus* and *L. hesperus*. Unequal crossover events during recombination could have given rise to the gene copies and could also maintain sequence similarity among gene copies over time. Further comparative mapping with taxa of increasing divergence from *Latrodectus* will pinpoint when the *MaSp1* duplication events occurred and the phylogenetic distribution of silk gene linkage patterns.

**Citation:** Zhao Y, Ayoub NA, Hayashi CY (2010) Chromosome Mapping of Dragline Silk Genes in the Genomes of Widow Spiders (Araneae, Theridiidae). PLoS ONE 5(9): e12804. doi:10.1371/journal.pone.0012804

**Editor:** Sergios-Orestis Kolokotronis, American Museum of Natural History, United States of America

**Received:** June 17, 2010; **Accepted:** August 19, 2010; **Published:** September 21, 2010

**Copyright:** © 2010 Zhao et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** The research was funded by the Army Research Office (W911NF-06-1-0455 to CH) and National Science Foundation (IOS-0951086 to NA, IOS-0951061 to CH). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** The authors have declared that no competing interests exist.

\* E-mail: yonghuiz@ucr.edu

<sup>‡</sup> Current address: Department of Biology, Washington and Lee University, Lexington, Virginia, United States of America

## Introduction

With over 41,000 described species, spiders are one of the most diverse metazoan orders [1]. Numerous studies have been published on spider ecology, physiology, systematics, venom toxicology, silk biology, *etc.* [2]. Despite the prevalence of spiders and research on them, their genetic characterization remains extremely limited. To date, there are no sequenced spider genomes. Rather, genomic analyses are mostly limited to measures of genome size and cytogenetic studies [3], [4]. The cytogenetic work has been restricted to determinations of karyotype, diploid chromosome number, C-banding patterns, nucleolus organizer regions, and sex chromosomes [5][6]–[7]. Although karyotypes of more than 500 species of spiders have been reported [4], no genes have been localized onto chromosomes.

Western black widow (*Latrodectus hesperus*) and brown widow (*L. geometricus*) spiders are in the cob-web weaver family, Theridiidae (Araneae). Previous studies on these two species have largely focused on systematics, venom pharmacology, and silk characterizations, such as in [8][9][10][11][12]–[13]. In terms of cytogenetics, there is one report on the diploid chromosome number of female brown widows [14], but no descriptions of the karyotypes of brown widow males or either sex of the Western black widow.

Spider dragline silk is renowned for its high tensile strength and toughness, with Western black widow draglines among the strongest on record [15], [16]. The two primary constituents of orb-weaver dragline silk are the proteins, major ampullate spidroins 1 and 2 (MaSp1 and MaSp2) [17]. Complementary DNA (cDNA) encoding MaSp1 and MaSp2 of the golden orb-weaver, *Nephila clavipes*, were the first published spider silk gene sequences [18], [19]. These breakthroughs helped lead to the characterization of MaSp1 and MaSp2 from other species. For brown widows, MaSp1 and MaSp2 were described from conceptual translations of cDNA and genomic DNA clones [10], [20]. MaSp1 and MaSp2 of Western black widows were also initially characterized by cDNAs [21], [22].

Silk protein transcripts are typically very long and highly repetitive [23]. Given these attributes and the technical limitations of cDNA synthesis and cloning, nearly all spider silk genes are known only from partial-length cDNAs or gene fragments. Through construction of a large-insert genomic DNA library for the Western black widow, the first full-length *MaSp1* and *MaSp2* genes were deciphered, along with substantial amounts of flanking regions [11]. More recently, it was reported that *MaSp1* is a multi-copy gene in several spider species [24][25]–[26]. For widow spiders, three *MaSp1* loci were diagnosed in the *L. hesperus* and *L.*

*geometricus* genomes, and a *MaSp1* pseudogene was also found in *L. hesperus*. A series of PCR amplifications using DNA extractions from individual spiders were performed to distinguish between gene copies (separate loci) and allelic variation [25]. There was no evidence for more than one *MaSp2* gene in *L. hesperus*.

The spatial relationships of the *MaSp1* multiple loci and *MaSp2* locus in *Latrodectus* genomes are unknown. Indeed, to our knowledge, no silk gene has ever been physically located within any spider genome. To map the locations of the major ampullate spidroin loci, we use fluorescence *in situ* hybridization (FISH). FISH is typically performed with probes and targets derived from the same species. However, it can also be used for comparative mapping of genomic fragments onto the chromosomes of closely related species, without additional cloning or DNA sequencing, as shown in [27], [28].

The construction of a large-insert genomic DNA library for *L. hesperus* provides a rich resource for FISH probes that can be used to map genes and trace genome evolution in widow spiders [11]. To enable such studies, here we present the karyotypes and chromosomal C-banding for *L. hesperus* and *L. geometricus*. Next, we map five fosmid clones of *L. hesperus* to specific chromosomes of *L. hesperus* with FISH. Each clone harbors one of three distinct *MaSp1* gene copies, a *MaSp1* pseudogene, or *MaSp2*. Furthermore, the three *MaSp1* gene copies and *Masp2* are heterologously mapped onto the karyotype of *L. geometricus*. Through these efforts, we aim to test scenarios of how the spidroin gene family evolved. If the silk genes are dispersed among chromosomes, that may indicate the birth of new paralogs via retroposition, among other possibilities. However, if the silk genes are localized to a single region in the

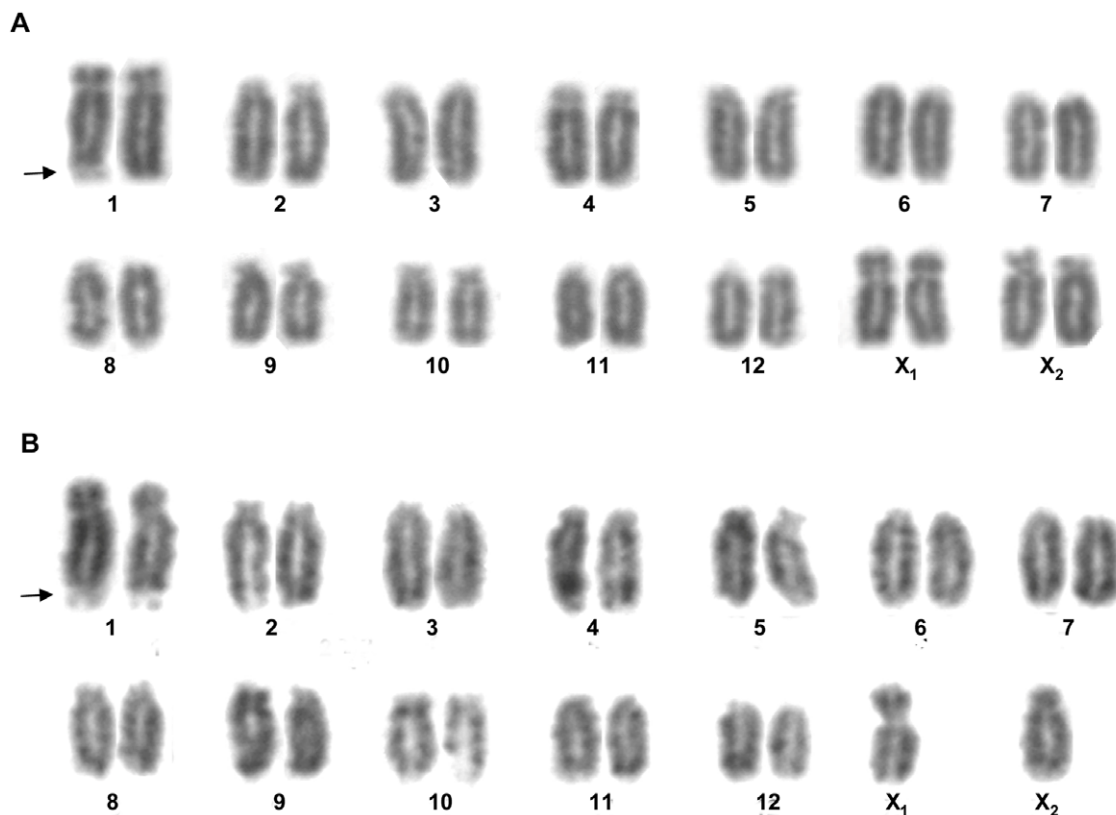
genome, that pattern would be consistent with gene duplication from unequal crossing over events [29].

## Results and Discussion

### Karyotypes of *L. hesperus* and *L. geometricus*

We obtained over 100 well-defined metaphase spreads from developing *L. hesperus* (LH) eggs (20 eggs analyzed, 5–6 spreads per egg). From these visualizations, we determined that Western black widows have diploid chromosome numbers of 26 or 28 (Figures 1, S1A, S1B). These counts are identical to the 26/28 chromosome sets reported for both *L. indistinctus* and *L. curacaviensis* [14], [30]. Aside from a submetacentric chromosome that is present singly in the 26 chromosome sets and paired in the 28 chromosome sets, all *L. hesperus* chromosomes are telo/acrocentric.

The autosome pairs in the *L. hesperus* karyotypes are numbered according to descending size (Figure 1). The largest autosome pair of *L. hesperus* (LH-1) has a secondary constriction and a satellite chromosome at the end of the long arm (Figures 1, S1A, S1B). One or two additional, small, telocentric chromosomes are infrequently found in the metaphase spreads (Figure S2). These small elements could be supernumerary chromosomes (e.g., B-chromosomes), which are known from eukaryotes such as grasshoppers, wasps and plants [31]. Because supernumerary chromosomes are often not stably transmitted, this could explain the variable number (zero, one, or two) of these small elements among *L. hesperus* karyotypes. The significance of the satellite and supernumerary chromosomal elements, and their polymorphic distribution among *L. hesperus* individuals, is an area for future inquiry.



**Figure 1. Karyotypes of *L. hesperus*.** (A) Metaphase spread with 28 chromosomes (female). (B) Metaphase spread with 26 chromosomes (male). Autosomes are numbered and sex chromosomes designated as "X<sub>1</sub>" or "X<sub>2</sub>." Arrows point to the terminal secondary constriction of the first autosome pair. Original metaphase spreads shown in Figures S1A and S1B. doi:10.1371/journal.pone.0012804.g001

For *L. geometricus* (LG), ~50 high-quality metaphase spreads were derived from 20 developing eggs. There were primarily two types of cells, those with a diploid chromosome number of 17 and those with 18 (Figures 2, S1C, S1D). All chromosomes were telo/acrocentric and no supernumerary chromosomes were observed. As with the *L. hesperus* karyotype, *L. geometricus* autosome pairs were numbered by descending size (Figure 2).

### Sex chromosome organization

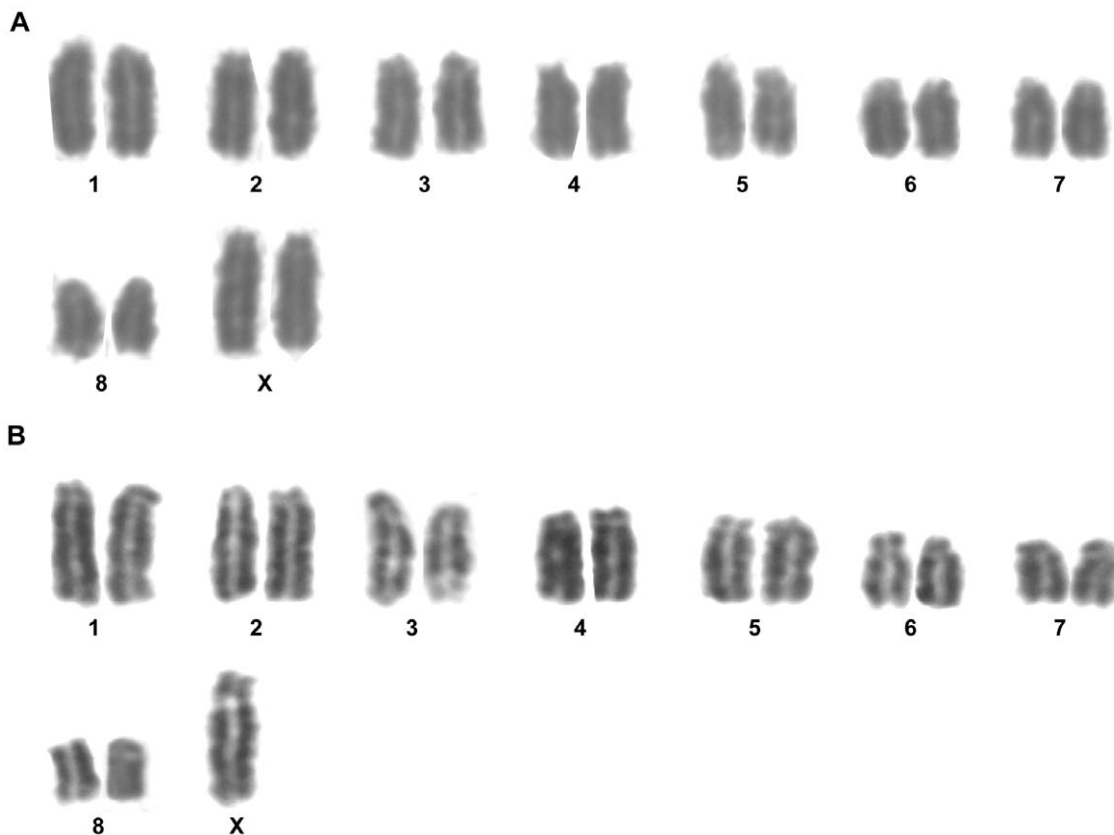
In spiders, males and females of a species usually differ in the number of sex chromosomes they possess [6], [32]. Females typically have more copies of the sex chromosomes. The most common sex chromosome organization of spiders in the Entelegynae, which includes theridiids, is an  $X_1X_20$  system [7]. In this scheme, there are two pairs of X chromosomes in females ( $X_1X_1X_2X_2$ ), while males have two unpaired X chromosomes ( $X_1X_2$ ). Thus, the diploid chromosome number of female spiders is commonly higher than that of males. Various other X chromosome systems (e.g.,  $X0$ ,  $X_1X_2X_30$ ,  $X_1X_2X_3X_40$ ), as well as some XY systems (e.g.,  $X_1X_2X_3Y$ ), have also been reported in spiders [33]–[35].

For both *L. hesperus* and *L. geometricus*, developing eggs, not individuals of known sex, were used to obtain the metaphase spreads. Thus additional information is needed to associate each karyotype with its respective sex. Considering the *L. hesperus* chromosome morphology and diploid chromosome numbers of 26 and 28, we deduce that the sex chromosome organization of the Western black widow is  $X_1X_2$  in males ( $2n = 26$ ) and  $X_1X_1X_2X_2$  in females ( $2n = 28$ ). The  $X_1$  sex chromosome is identifiable by its

unique, submetacentric appearance and its presence as either single or paired in the two types of metaphase chromosome spreads that we observed. The  $X_2$  sex chromosome is distinguishable by its short arm (p), which is longer than those of similar-sized acrocentric autosomes (Figure 1).

*L. geometricus* metaphase spreads have either 17 or 18 diploid chromosomes (Figure 2). The count of 18 chromosomes is consistent with the karyotype for a female *L. geometricus* that was previously described [14]. Hence, the sex chromosome organization of *L. geometricus* is  $X0$  (an unpaired X) in males ( $2n = 17$ ) and  $XX$  in females ( $2n = 18$ ). The brown widow X chromosome is distinguishable as the largest chromosome in the genome (Figure 2).

Further investigation is needed to determine the relationship between the two X chromosomes of *L. hesperus* and the single X chromosome of *L. geometricus*. Possible evolutionary scenarios include the origin of the *L. geometricus* X from fusion of two ancestral sex chromosomes, or the derivation of the two *L. hesperus* X chromosomes from fission of a single ancestral sex chromosome. Insight can be gained by considering the karyotypes of closely related species. The subfamily Latrodectinae is composed of three genera, *Latrodectus*, *Steatoda*, and *Crustulina* [36]. Karyotypes with 22 and 24 chromosomes have been published for *Steatoda bipunctata*, *S. grossa*, *S. triangulosa*, and *Crustulina guttata*, indicating that these species all have two sex chromosomes [30], [37]. Considering that *L. hesperus*, *L. indistinctus*, and *L. curacaviensis* also have two sex chromosomes, this is most likely the ancestral condition in latrodectine spiders [14], [30]. If so, then the single X chromosome in *L. geometricus* arose through chromosomal fusion.



**Figure 2. Karyotypes of *L. geometricus*.** (A) Metaphase spread with 18 chromosomes (female). (B) Metaphase spread with 17 chromosomes (male). Autosomes are numbered and sex chromosomes designated by "X." Original metaphase spreads shown in Figures S1C and S1D. doi:10.1371/journal.pone.0012804.g002

When molecular markers are developed for the *Latrodectus* X chromosomes, this hypothesis can be tested.

### C-banding of *L. hesperus* and *L. geometricus*

C-bands are regions of chromosomes that are formed by constitutive heterochromatin. They stain darkly due to the dense packing of chromatin. For many organisms, C-bands can serve as chromosomal markers because the distribution of heterochromatin varies among chromosomes within a genome and between the genomes of different species. In *L. hesperus*, we found that most of the C-bands are located in the telomeric and centromeric areas of each chromosome (Figure 3A). The chromosomes have similar banding patterns, except that the sex chromosomes have extra C-bands on their long arms, toward the centromeres (arrows in Figure 3A). In *L. geometricus*, the majority of C-bands are in the centromeric region of chromosomes, and LG-4 is distinctive in having more heterochromatin than any of the other chromosomes (Figure 3B).

The C-banding of *L. hesperus* is extensive, indicating that the ratio of constitutive heterochromatin to euchromatin is high in the *L. hesperus* genome. By contrast, there is much less C-banding, and hence relatively lower amounts of constitutive heterochromatin on *L. geometricus* chromosomes (Figure 3). The greater proportion of C-banding in *L. hesperus* compared to *L. geometricus* cannot be easily explained by genome size since *L. hesperus* has the smaller genome (1.29 pg for *L. hesperus* vs. 1.54 pg for *L. geometricus*) [38]. It is currently unknown whether there is functional significance to the variation in constitutive heterochromatin in *Latrodectus*, as has been reported in other eukaryotes [39].

### Mapping of *MaSp1* and *MaSp2* onto *Latrodectus* chromosomes by FISH

Clones from an *L. hesperus* fosmid library were successfully hybridized to chromosomes of *L. hesperus* and *L. geometricus*. The fosmid clones contained inserts of ~40 kb of genomic DNA. Four fosmid clones, each of which contained one of three loci of *MaSp1* or the *MaSp1* pseudogene, were mapped to the same general region, the long arm (q) of LH-1 (LH-1q; Figures 4A–4C, S3). Further analysis showed that these clones are located in the same specific region, as determined by measuring the Fractional Length from the p terminus (FLpter) of their hybridization signals (Table 1). A fifth fosmid clone that contained *MaSp2* was mapped onto LH-8q (Figure 4D).

The three *L. hesperus* fosmid clones with *MaSp1* gene copies were successfully mapped onto *L. geometricus* chromosomes. All three

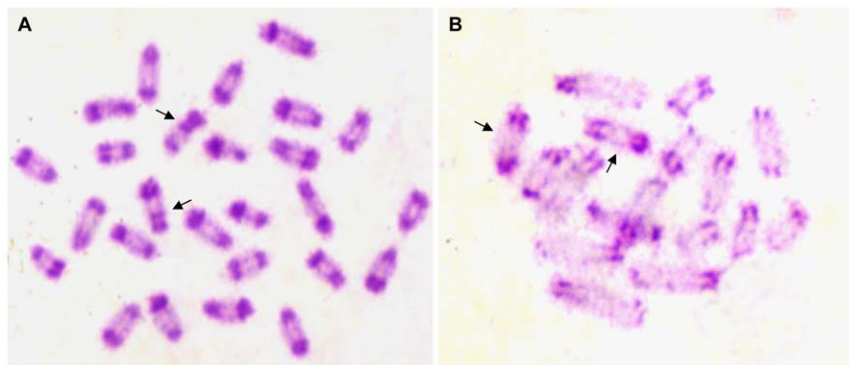
gene copies hybridized to the long arm of the fourth largest autosome pair of *L. geometricus* (LG-4q; Figure 5A–5C). FLpter values indicate that the three loci are not just on the same arm of the same chromosome, but they are in the identical chromosomal region (Table 1). The *L. hesperus* fosmid clone with *MaSp2* was comparatively mapped to the long arm of the second-largest autosome pair of *L. geometricus* (LG-2q; Figure 5D).

The relative sizes of the *MaSp1* and *MaSp2* containing chromosomes vary between *Latrodectus* species. In *L. hesperus*, the *MaSp1* loci are on LH-1, the largest chromosome, while *MaSp2* is on LH-8, a considerably smaller chromosome. By contrast, in *L. geometricus*, *MaSp2* is on a larger chromosome (LG-2) than the *MaSp1* loci (LG-4). Furthermore, as indicated by FLpter values, *L. hesperus* *MaSp2* is near the middle of the chromosome long arm, while in *L. geometricus*, the gene is close to the telomere (Table 1). Thus, the FLpter values and morphometry of the *MaSp1* and *MaSp2* bearing chromosomes are signposts of the genomic rearrangements that have occurred since *L. geometricus* and *L. hesperus* diverged from their most recent common ancestor.

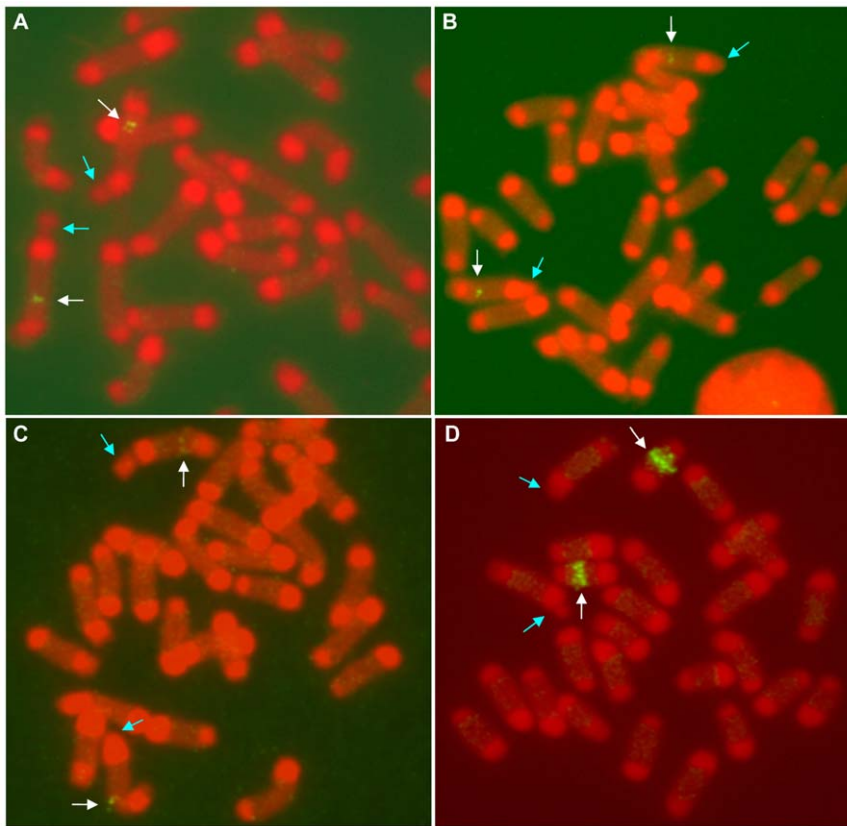
### Evolution of *Latrodectus* silk genes and genomic regions

Based on analyses of mitochondrial cytochrome c oxidase subunit 1, the genus *Latrodectus* is divided into two monophyletic groups, the *geometricus* clade and the *mactans* clade [8]. The *geometricus* clade is named after its member, *L. geometricus*, and the *mactans* clade includes *L. hesperus*. Thus, *L. geometricus* and *L. hesperus* span the maximum phylogenetic divergence in *Latrodectus*. Differences in number of autosome pairs, sex chromosome systems, and pattern of constitutive heterochromatin are further evidence of the marked genomic changes between these species (Figures 1–3). Yet, by comparative mapping of *MaSp1* and *MaSp2* genes with FISH, we are able to predict orthology of genomic regions on LH-1q with LG-4q, and LH-8q with LG-2q (Table 1).

The mapping of multiple *MaSp1* genes to single chromosomal regions in *L. hesperus* and *L. geometricus* (Figures 4 and 5) implies that in widow genomes, the *MaSp1* loci are in a gene cluster. Based on the resolution of FISH with metaphase chromosomes, we deduce that this gene cluster is within a 2–3 Mb region [40]. Neighboring of multicopy genes has also been observed in other organisms, such as *DAZ* fertility factor genes on the human Y chromosome, patatin genes in potato, and histone genes in grasshopper [41][42]–[43]. The *MaSp1* gene cluster contains a minimum of three copies and it is possible that there are more copies that have yet to be characterized. However, if there are additional copies, there is no evidence that they are located outside of the gene cluster. Given that the full-length *MaSp1* gene copy (locus 1) is a



**Figure 3. C-banding of chromosomes from male *Latrodectus*.** (A) *L. hesperus* with arrows pointing to sex chromosomes X<sub>1</sub> and X<sub>2</sub>, and (B) *L. geometricus* with arrows pointing to LG-4. doi:10.1371/journal.pone.0012804.g003



**Figure 4. FISH mapping of *L. hesperus* chromosomes with probes.** (A) *MaSp1* locus 1, (B) *MaSp1* locus 2, (C) *MaSp1* locus 3, and (D) *MaSp2*. White arrows point to hybridization signals, blue arrows to the secondary constriction of LH-1. Further details about probes and signal locations are in Table 1.

doi:10.1371/journal.pone.0012804.g004

~10 kb single exon gene, cross-hybridization could occur among gene copies due to the substantial amounts of sequence conservation observed among the three known loci [11], [25]. Yet, FISH did not reveal signal outside of the gene cluster region, which would have indicated dispersion of *MaSp1* loci in *L. hesperus* or *L. geometricus* (Figures 4A–4C, 5A–5C).

The presence of *MaSp1* gene clusters in both *Latrodectus* species, suggests that *MaSp1* gene duplications predate their speciation. If the multiple *MaSp1* loci in *Nephila* (Nephilidae) [26] and *Euprostheno*s (Pisauridae) [24] are also similarly localized in their respective genomes, then the *MaSp1* gene cluster could be extremely ancient. Based on the estimated divergence time of these spider lineages, the

gene cluster could have persisted for over 200 million years [44]. Alternatively, *MaSp1* gene clusters could have multiple origins. To distinguish between the scenarios of single vs. convergent evolution of gene clusters, the genomic distribution of *MaSp1* genes and adjacent chromosomal regions need to be mapped across spider families. Regardless of the history of the *MaSp1* multiple loci, there appears to be selection for the maintenance of the gene copies. Possible functional significance could be to increase gene dosage or as part of tissue-specific, developmental-stage, or other condition-specific differential gene expression [25].

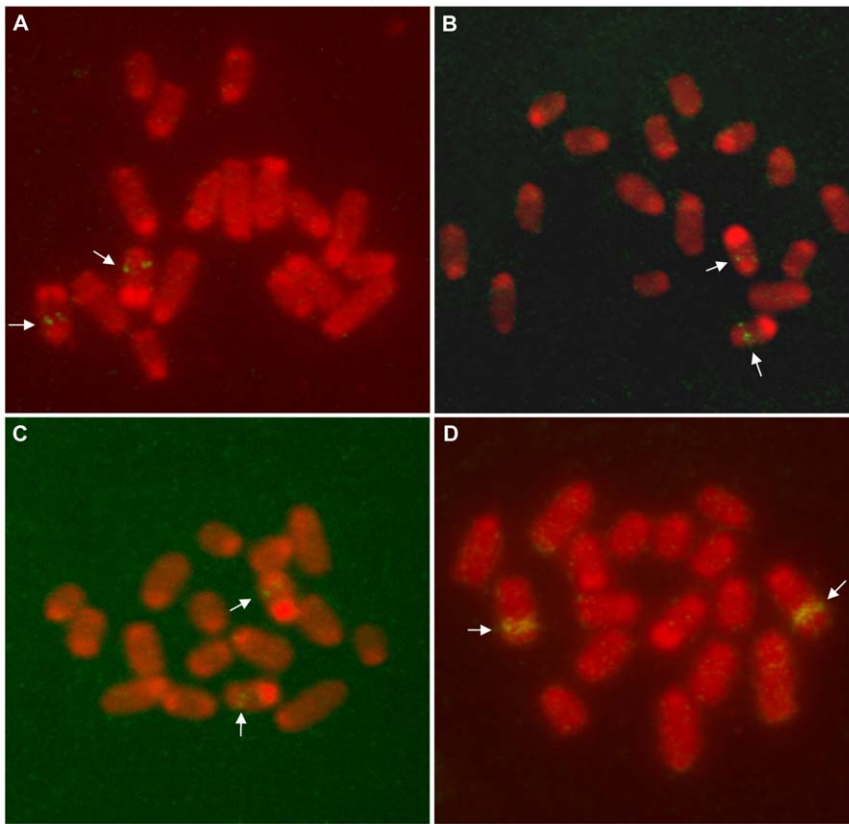
The localization of *Latrodectus MaSp1* loci to a single chromosomal region is also consistent with the spawning of gene copies by

**Table 1.** Genomic location (chromosome number-arm) and FLpter values of *MaSp1* gene copies, *MaSp1* pseudogene, or *MaSp2* in *L. hesperus* (LH) and *L. geometricus* (LG).

Fosmid probe	Genbank Accession	LH location	LH FLpter	LG location	LG FLpter
<i>MaSp1</i> locus 1	EF595246	LH-1q	0.43±0.03	LG-4q	0.36±0.03
<i>MaSp1</i> locus 2	EU177653	LH-1q	0.44±0.03	LG-4q	0.34±0.03
<i>MaSp1</i> locus 3	EU177650	LH-1q	0.45±0.03	LG-4q	0.35±0.03
<i>MaSp1</i> pseudogene	EU177647	LH-1q	0.46±0.03	n/a	n/a
<i>MaSp2</i>	EF595245	LH-8q	0.44±0.05	LG-2q	0.72±0.05

Chromosomes numbered as in Figure 1. Each FLpter value is based on 8–15 metaphase spreads. In calculating the FLpter values on LH-1q, the relative lengths of secondary constrictions were not included because they were not visible in all metaphase spreads.

doi:10.1371/journal.pone.0012804.t001



**Figure 5. FISH comparative mapping of *L. geometricus* chromosomes with probes.** (A) *MaSp1* locus 1, (B) *MaSp1* locus 2, (C) *MaSp1* locus 3, and (D) *MaSp2*. Arrows point to hybridization signals. Further details about probes and signal locations are in Table 1. doi:10.1371/journal.pone.0012804.g005

tandem duplication. Segmental duplication via unequal recombination of sister chromatids is one mechanism that can generate new paralogs [29], [45]. Fibroin genes in particular are especially prone to unequal recombination because they are largely composed of repetitive sequences [46]. For example, the aciniform spidroin gene (*AcSp1*) from the banded garden spider, *Argiope trifasciata*, is built from a tandem array of 600 bp repeat units, some of which share an extraordinary 100% pairwise nucleotide identity [47]. Remarkable sequence conservation has also been noted in other members of the spidroin gene family, including *MaSp1* and *MaSp2* from widow spiders [11], [22], [48], [49]. Specifically, in the *L. hesperus* *MaSp1* locus 1, there are 20 tandem-arrayed iterations of a repeat that average a scant 2.5% pairwise sequence divergence over 348 nucleotides [11]. The unusual repetitive design of silk genes provides ample opportunities for misalignment of chromosomes during meiosis and unequal crossing over. The genetic shuffling caused by these events results in duplication or deletion not only of intragenic repeat units (e.g., allelic length variants; [50]), but also of entire gene copies (e.g., tandem duplicates). The genomic localization of multiple *MaSp1* loci in *Latrodectus* is consistent with the persistence of tandem duplicates over evolutionary time.

Tandem duplication of *MaSp1* results in tandem arrayed gene copies. This physical arrangement is a higher level of repetitiveness beyond the ensemble repeats and subrepeats previously described for spider silks. Sequence conservation among tandem repeat units within a gene can occur by concerted evolution that homogenizes the repeats throughout an array. Similarly, physical proximity on a chromosome can explain much of the sequence conservation

noted for *Latrodectus* *MaSp1* loci [25] through concerted evolution of genes, such as observed for arrays of rRNA genes [51].

The location of *MaSp2* on a different chromosome from the *MaSp1* loci in *Latrodectus* species is also informative about the evolutionary dynamics of the spidroin gene family (Table 1). Phylogenetic analyses of *MaSp1* and *MaSp2* indicate that they form a clade to the exclusion of other spidroin gene family members [17], [20], [22]. Furthermore, *Latrodectus* *MaSp2* appears to have a *MaSp1* ancestor, meaning that *MaSp2* is a divergent *MaSp1* paralog [25]. This inference is in agreement with the demonstration from FISH mapping that *MaSp2* is unlinked to *MaSp1*. Accordingly, *MaSp2* likely originated from a *MaSp1* gene copy either via retroposition of a *MaSp1* transcript (hence, *MaSp2*'s disparate location), or translocation to a different chromosome. Once unlinked from the other *MaSp1* paralogs, the repetitive region of *MaSp2* dramatically diverged in sequence from *MaSp1*. Analyses to detect past recombination events, however, reveal that there has been recombination between the amino- and carboxyl-termini encoding regions of *Latrodectus* *MaSp1* loci and *MaSp2*. Yet, this recombination rate is less frequent than the homogenization among *MaSp1* loci [25].

In summary, the karyotyping of the brown widow and Western black widow revealed their autosome and sex chromosome organization (Figures 1 and 2). Characterization of these karyotypes enabled the first comparative mapping of spider genomes. FISH was used to locate the positions of multiple *MaSp1* gene copies and the *MaSp2* gene in *Latrodectus* species (Figures 4, 5, S3). Given the substantial phylogenetic divergence, differences in genomic organization, and near absence of genomic

sequence data for these species, our results attest to the power of FISH for gaining insight into the molecular evolution of understudied organisms.

The comparative genomic hybridization results shed light on the evolutionary dynamics of the *MaSp1* and *MaSp2* genes that underlie the structural proteins of one of nature's most renowned materials, the spider dragline. *MaSp1* loci, including a pseudogene, are clustered (Figures 4, 5, S3; Table 1), consistent with the origin of paralogs through segmental duplication events and homogenization via concerted evolution [29], [45]. *MaSp2*, which is closely related to *MaSp1* [25], is unlinked from the *MaSp1* array. This physical separation is consistent with the substantial differences that have accumulated between *MaSp1* and *MaSp2*. Indeed, *MaSp2* has distinctive attributes, such as prevalent proline residues, that contribute extensibility and toughness to spider draglines [23]. Thus, the mapping of silk genes on chromosomes has implications for understanding the genomic environments in which the members of the spider silk gene family evolved to produce a diversity of high-performance biomaterials.

## Materials and Methods

### Preparation of chromosome spreads

Mature female *Latrodectus hesperus* (Western black widow) and *Latrodectus geometricus* (brown widow) were captured in Riverside and San Diego, California, respectively. Females were maintained in the lab and observed daily for the production of egg cases. Eggs were allowed to develop for 8–10 days and then were collected for chromosomes. For karyotyping, one egg was used per chromosome preparation. For FISH, multiple eggs were combined per preparation. Standard cytogenetic techniques were used to prepare and stain chromosome spreads [37], [52].

### Preparation of FISH probes

We cultured five fosmid clones that were isolated from a previously constructed *L. hesperus* genomic library [11]. The five clones harbored one of three *MaSp1* loci, a *MaSp1* pseudogene, or a *MaSp2* gene. Fosmid DNA was labeled with digoxigenin (Roche Applied Science) then co-precipitated with a 50-fold excess of salmon sperm DNA and sonicated *L. hesperus* genomic DNA. Immediately prior to hybridization, the precipitate was dissolved in hybridization solution (10% dextran sulfate sodium salt, 25% formamide, 2×SSC), denatured at 75°C, and put on ice.

### Fluorescence *in situ* hybridization

FISH was performed following previously described methods [53]. Briefly, slides of chromosome preparations were incubated at

60°C for 2 h, denatured at 72°C for 3 min in 70% formamide/2×SSC, and dehydrated at –20°C in an ethanol series. The denatured FISH probes were then added to the slides and held in place by cover glass. All five probes described above were used individually with *L. hesperus* chromosome preparations. Four of the probes (the functional genes) were used for heterologous FISH of *L. geometricus* chromosomes. The pseudogene containing probe was not used because it had only ~500 bp of sequence identifiable as *L. hesperus MaSp1* and an orthologous region was not known in *L. geometricus*. Hybridizations were performed for 17 h at 37°C. After washing, hybridization signals were detected with mouse monoclonal anti-digoxigenin FITC antibodies that were amplified with rabbit anti-mouse IgG FITC and goat anti-rabbit IgG FITC antibodies (Sigma-Aldrich). Chromosomes were counterstained with propidium iodide. Images were captured with a Zeiss epifluorescence microscope equipped with a CCD camera.

## Supporting Information

**Figure S1** Metaphase spreads used to derive the karyotypes shown in Figures 1 and 2. *L. hesperus* with (A) 28 chromosomes or (B) 26 chromosomes. Arrows point to second constriction of LH-1. *L. geometricus* with (C) 18 chromosomes, or (D) 17 chromosomes. Found at: doi:10.1371/journal.pone.0012804.s001 (1.24 MB TIF)

**Figure S2** Metaphase spread from *L. hesperus* with arrows pointing to two supernumerary chromosomes. Found at: doi:10.1371/journal.pone.0012804.s002 (0.76 MB TIF)

**Figure S3** FISH mapping of the *MaSp1* pseudogene probe onto *L. hesperus* chromosomes. White arrows point to hybridization signals, blue arrows to the second constriction of LH-1. Further details about probe and signal location are in Table 1. Found at: doi:10.1371/journal.pone.0012804.s003 (0.88 MB TIF)

## Acknowledgments

A. Lukaszewski generously helped with the microscopy. M.L. Casem, J. Garb, M. Hedin, and J. Starrett assisted with the collection of spiders. M. Collin, J. Král, A. Lukaszewski, and J. Starrett improved the manuscript.

## Author Contributions

Conceived and designed the experiments: YZ NA CH. Performed the experiments: YZ NA. Analyzed the data: YZ NA CH. Contributed reagents/materials/analysis tools: CH. Wrote the paper: YZ NA CH.

## References

- Platnick NI (2010) The world spider catalog, Version 10.5. Am Mus Nat Hist <http://research.amnh.org/iz/spiders/catalog/>.
- Foelix RF (1996) Biology of Spiders. Oxford: Oxford University Press. 336 p.
- Gregory TR (2006) Animal genome size database. <http://www.genomesize.com/>.
- Řezáč M, Král J, Musilová J, Pekar S (2006) Unusual karyotype diversity in the European spiders of the genus *Ahytus* (Araneae: Atypidae). Hereditas 143: 123–129.
- Chen S (1999) Cytological studies on six species of spiders from Taiwan (Araneae: Theridiidae, Pschridae, Uloboridae, Oxyopidae, and Ctenidae). Zool Stud 38: 423–434.
- Araújo D, Cella MD, Brescovit DA (2005) Cytogenetic analysis of the neotropical spider *Nephilengys cruentata* (Araneomorphae, Tetragnathidae): standard staining norms, C-bands and base-specific fluorochromes. Braz J Biol 65: 193–202.
- Král J, Musilová J, Štáhlavský F, Řezáč M, Akan Z, et al. (2006) Evolution of the karyotype and sex chromosome systems in basal clades of araneomorph spiders (Araneae: Araneomorphae). Chromosome Res 14: 859–880.
- Garb JE, Gonzalez A, Gillespie RG (2004) The black widow spider genus *Latrodectus* (Araneae: Theridiidae): phylogeny, biogeography, and invasion history. Mol Phylogenet Evol 31: 1127–1142.
- Blackledge TA, Summers AP, Hayashi CY (2005) Gumfooted lines in black widow cobwebs and the mechanical properties of spider capture silk. Zoology 108: 41–46.
- Motriuk-Smith D, Smith A, Hayashi CY, Lewis RV (2005) Analysis of the conserved N-terminal domains in major ampullate spider silk proteins. Biomacromolecules 6: 3152–3159.
- Ayoub NA, Garb JE, Tinghitella RM, Collin MA, Hayashi CY (2007) Blueprint for a high-performance biomaterial: full-length spider dragline silk genes. PLoS One 2: e514.
- Blasingame E, Tuton-Blasingame T, Larkin L, Falick AM, Zhao L, et al. (2009) Pyriform spidroin 1, a novel member of the silk gene family that anchors dragline silk fibers in attachment discs of the black widow spider, *Latrodectus hesperus*. J Biol Chem 284: 29097–29108.
- Guerrero B, Finol HJ, Reyes-Lugo M, Salazar AM, Sánchez EE, et al. (2010) Activities against hemostatic proteins and adrenal gland ultrastructural changes caused by the brown widow spider *Latrodectus geometricus* (Araneae: Theridiidae) venom. Comp Biochem Physiol C Toxicol Pharmacol 151: 113–121.
- Martindale C (1980) Studies on *Latrodectus*. P Entomol Congr Entomol Soc South Af 3: 45–46.

15. Gosline JM, Demont ME, Denny MW (1986) The structure and properties of spider silk. *Endeavour* 10: 37–43.
16. Swanson BO, Blackledge TA, Summers AP, Hayashi CY (2006) Spider dragline silk: correlated and mosaic evolution in high-performance biological materials. *Evolution* 60: 2539–2551.
17. Garb JE, Dimauro T, Vo V, Hayashi CY (2006) Silk genes support the single origin of orb webs. *Science* 312: 1762.
18. Xu M, Lewis RV (1990) Structure of a protein superfiber: spider dragline silk. *Proc Natl Acad Sci USA* 87: 7120–7124.
19. Hinman MB, Lewis RV (1992) Isolation of a clone encoding a second dragline silk fibroin. *Nephila clavipes* dragline silk is a two-protein fiber. *J Biol Chem* 267: 19320–19324.
20. Gatesy J, Hayashi C, Motriuk D, Woods J, Lewis R (2001) Extreme diversity, conservation, and convergence of spider silk fibroin sequences. *Science* 291: 2603–2605.
21. Lawrence BA, Vierra CA, Moore AM (2004) Molecular and mechanical properties of major ampullate silk of the black widow spider, *Latrodectus hesperus*. *Biomacromolecules* 5: 689–695.
22. Garb JE, Hayashi CY (2005) Modular evolution of egg case silk genes across orb-weaving spider superfamilies. *Proc Natl Acad Sci USA* 102: 11379–11384.
23. Hayashi C, Shipley N, Lewis R (1999) Hypotheses that correlate the sequence, structure, and mechanical properties of spider silks. *Int J Biol Macromol* 24: 271–275.
24. Rising A, Johansson J, Larson G, Bongcam-Rudloff E, Engstrom W, et al. (2007) Major ampullate spidroins from *Euprosthenops australis*: multiplicity at protein, mRNA and gene levels. *Insect Mol Biol* 16: 551–561.
25. Ayoub NA, Hayashi CY (2008) Multiple recombining loci encode MaSp1, the primary constituent of dragline silk, in widow spiders (*Latrodectus*: Theridiidae). *Mol Biol Evol* 25: 277–286.
26. Gaines WT, Marcotte WJ (2008) Identification and characterization of multiple Spidroin 1 genes encoding major ampullate silk proteins in *Nephila clavipes*. *Insect Mol Biol* 17: 465–474.
27. Kukekova AV, Vorobieva NV, Beklemisheva VR, Johnson JL, Temnykh SV, et al. (2009) Chromosomal mapping of canine-derived BAC clones to the red fox and American mink genomes. *J Hered* 100(Suppl 1): S42–53.
28. Yasukochi Y, Tanaka-Okuyama M, Shibata F, Yoshida A, Marec F, 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.
29. Zhang J (2003) Evolution by gene duplication: an update. *Trends Ecol Evol* 18: 292–298.
30. Araújo D, Maia UM, Brescovit AD (2010) The first cytogenetic characterization of the poisonous black widow spider *Latrodectus* gr. *curacaviensis* from Brazil, with chromosomal review of the family Theridiidae (Arachnida, Araneae). *Micron* 41: 165–168.
31. Camacho JP, Sharbel TF, Beukeboom LW (2000) B-chromosome evolution. *Philos Trans R Soc Lond B Biol Sci* 355: 163–178.
32. Král J (2007) Evolution of multiple sex chromosomes in the spider genus *Malthonica* (Araneae: Agelenidae) indicates unique structure of the spider sex chromosome systems. *Chromosome Res* 15: 863–879.
33. Maddison WP (1982) XXXY sex chromosomes in males of the jumping spider genus *Pellenes* (Araneae: Salticidae). *Chromosoma* 85: 23–37.
34. Datta SN, Chatterjee K (1983) Chromosome number and sex determining system in fifty-two species of spiders from North-East India. *Chromosome Information Service* 35: 6–9.
35. Rowell DM (1985) Complex sex-linked fusion heterozygosity in the Australian huntsman spider *Delena cancerides* (Araneae: Sparassidae). *Chromosoma* 93: 169–176.
36. Arnedo MA, Coddington J, Agnarsson I, Gillespie RG (2004) From a comb to a tree: phylogenetic relationships of the comb-footed spiders (Araneae, Theridiidae) inferred from nuclear and mitochondrial genes. *Mol Phylogenet Evol* 31: 225–245.
37. Tugmon C, Brown J, Horner N (1990) Karyotypes of seventeen USA spiders (Araneae, Araneidae, Gnaphosidae, Loxoscelidae, Lycosidae, Oxyopidae, Philodromidae, Salticidae and Theridiidae). *J Arachnol* 18: 41–48.
38. Gregory TR, Shorthouse DP (2003) Genome sizes of spiders. *J Hered* 94: 285–290.
39. Dimitri P, Caizzi R, Giordano E, Carmela Accardo M, Lattanzi G, et al. (2009) Constitutive heterochromatin: a surprising variety of expressed sequences. *Chromosoma* 118: 419–435.
40. Laan M, Kallioniemi O, Hellsten E, Alitalo K, Peltonen L, et al. (1995) Mechanically stretched chromosomes as targets for high-resolution FISH mapping. *Genome Res* 5: 13–20.
41. Repping S, de Vries J, van Daalen S, Korver C, Leschot N, et al. (2003) The use of spermHALO-FISH to determine DAZ gene copy number. *Mol Hum Reprod* 9: 183–188.
42. Ganai MW, Bonierbale MW, Roeder MS, Park WD, Tanksley SD (1991) Genetic and physical mapping of the patatin genes in potato and tomato. *Mol Gen Genet* 225: 501–509.
43. Cabrero J, Lopez-Leon MD, Teruel M, Camacho JP (2009) Chromosome mapping of H3 and H4 histone gene clusters in 35 species of acridid grasshoppers. *Chromosome Res* 17: 397–404.
44. Ayoub NA, Hayashi CY (2009) Spiders (Araneae). in SBHedges, SKumar, eds. *Timetree of Life* Oxford University Press. pp 255–259.
45. Hurler M (2004) Gene duplication: the genomic trade in spare parts. *PLoS Biol* 2: e206.
46. Mita K, Ichimura S, James TC (1994) Highly repetitive structure and its organization of the silk fibroin gene. *J Mol Evol* 38: 583–592.
47. Hayashi CY, Blackledge TA, Lewis RV (2004) Molecular and mechanical characterization of aciniform silk: uniformity of iterated sequence modules in a novel member of the spider silk fibroin gene family. *Mol Biol Evol* 21: 1950–1959.
48. Hayashi CY, Lewis RV (2000) Molecular architecture and evolution of a modular spider silk protein gene. *Science* 287: 1477–1479.
49. Garb J, DiMauro T, Lewis R, Hayashi C (2007) Expansion and intragenic homogenization of spider silk genes since the Triassic: evidence from Mygalomorphae (tarantulas and their kin) spidroins. *Mol Biol Evol* 24: 2454–2464.
50. Chinali A, Vater W, Rudakoff B, Sponner A, Unger E, et al. (2010) Containment of extended length polymorphisms in silk proteins. *J Mol Evol* 70: 325–338.
51. Ganley AR, Kobayashi T (2007) Highly efficient concerted evolution in the ribosomal DNA repeats: total rDNA repeat variation revealed by whole-genome shotgun sequence data. *Genome Res* 17: 184–191.
52. Peng Y, Hu C, Zhao J, Chen J (1998) Karyotype analysis of *Pirata piraticus* (Araneae: Lycosidae). *Acta Arachnol Sin* 7: 142–145.
53. Coppiters W, Zijlstra C, Van de Weghe A, Bosma AA, Peelman L, et al. (1994) A porcine minisatellite located on chromosome 14q29. *Mamm Genome* 5: 591–593.