Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Insect Science

Recent progress and prospects for advancing arachnid genomics

Jessica E Garb[1], Prashant P Sharma[2] and Nadia A Ayoub[3]

Arachnids exhibit tremendous species richness and adaptations of biomedical, industrial, and agricultural importance. Yet genomic resources for arachnids are limited, with the first few spider and scorpion genomes becoming accessible in the last four years. We review key insights from these genome projects, and recommend additional genomes for sequencing, emphasizing taxa of greatest value to the scientific community. We suggest greater sampling of spiders whose genomes are understudied but hold important protein recipes for silk and venom production. We further recommend arachnid genomes to address significant evolutionary topics, including the phenotypic impact of genome duplications. A barrier to high-quality arachnid genomes are assemblies based solely on short-read data, which may be overcome by long-range sequencing and other emerging methods.

Addresses
[1] Department of Biological Sciences, 198 Riverside Street, Olsen Hall 414, University of Massachusetts Lowell, Lowell, MA 01854, USA
[2] Department of Integrative Biology, 352 Birge Hall, 430 Lincoln Drive, University of Wisconsin-Madison, WI 53706, USA
[3] Department of Biology, 204 West Washington Street, Howe Hall, Washington and Lee University, Lexington, VA 24450, USA

Corresponding author: Garb, Jessica E (Jessica_Garb@uml.edu)

Current Opinion in Insect Science 2018, **25**:51–57

This review comes from a themed issue on **Insect genomics**

Edited by **Christopher Childers**, **Anna Childers** and **Stephen Richards**

https://doi.org/10.1016/j.cois.2017.11.005

Introduction

Arachnids are an arthropod class containing over 130 000 described species in 12 extant orders including Araneae (spiders), Scorpiones (scorpions), Acariformes (mites), Parasitiformes (ticks), Opiliones (harvestmen) and Thelyphonida (vinegaroons) (Figure 1a; [1,2]). Despite their diversity and key phylogenetic position, the first arachnid genome became available as recently as 2008 and genomes have only become accessible for non-acarine (mite and tick) arachnids in the last four years [3–5,6••,7••,8].

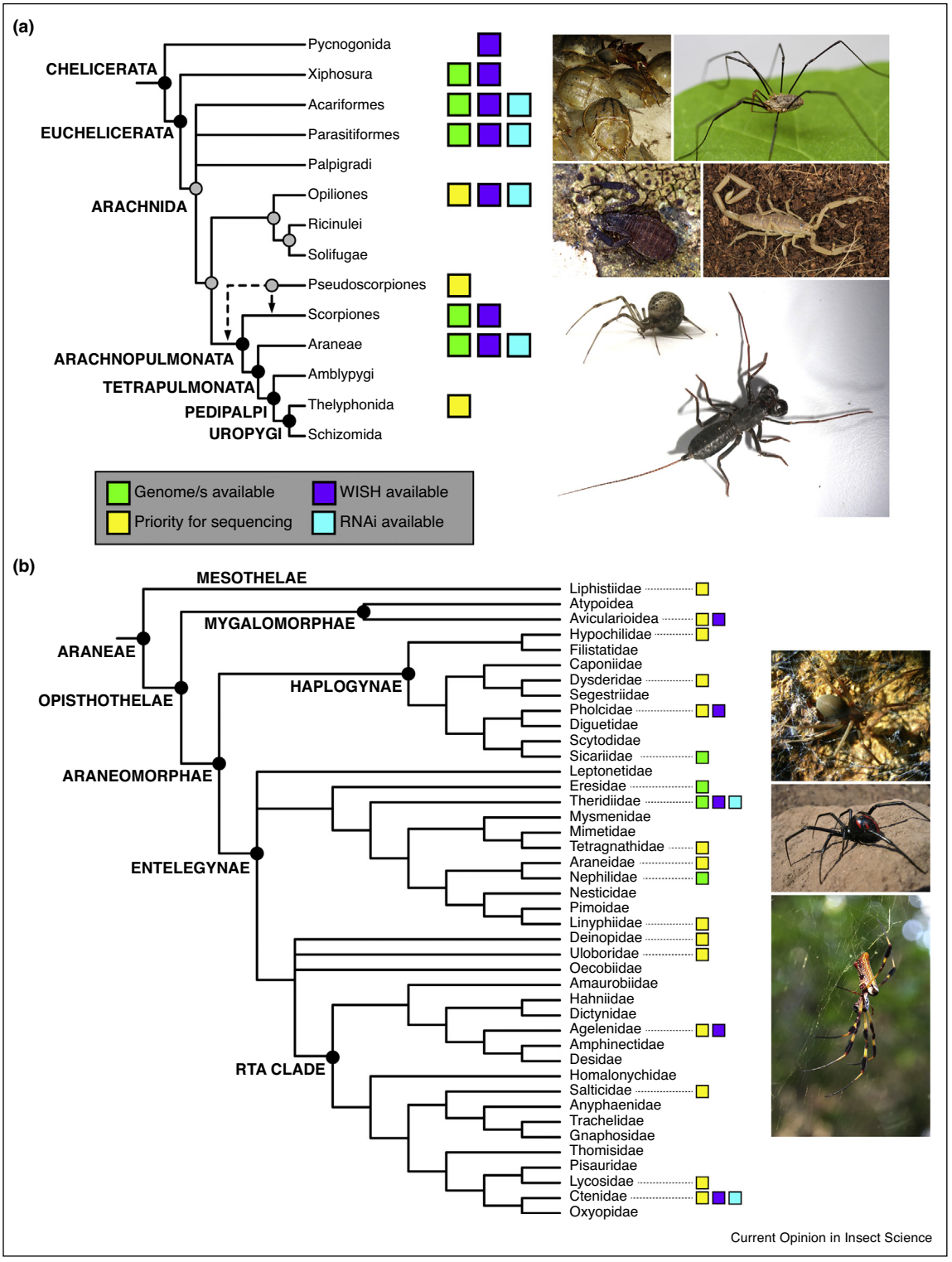
The earliest arachnid genome sequencing focused on members of the Acari that are important plant and animal pests, such as *Ixodes scapularis* (black legged tick, transmitter of Lyme disease), *Varroa destructor* (honeybee mite), *Tetranychus urticae* (red spider mite), *Rhipicephalus microplus* (southern cattle tick), and *Galendromus occidentalis* (western predatory mite) [9–11]. More recently, two scorpion and five spider genomes were deposited in NCBI (Table 1; [3–5,6••,7••]), several of which were produced as part of the i5k, an initiative to sequence 5000 medically and agriculturally important arthropods [4]. In this review we focus on significant findings uncovered by new spider and scorpion genomes and make recommendations for additional genomes urgently needed to address research questions of broadest scientific interest.

Spider silk biology

A spider-specific trait of particular economic interest is silk production due to the impressive mechanical properties and biomimetic potential of these fibers. Although silk production has evolved multiple times in arthropods, it has reached greatest sophistication in spiders, which can make up to seven distinct types of silk fibers and glues [12]. These include dragline silks with toughness that exceeds Kevlar, prey capture threads that can reversibly extend 300% [12], and environmentally responsive silk glues [13]. Spider silk fibers are primarily composed of different members of a spider-specific family of structural proteins (spidroins) that dictate their divergent material properties. Genomes, in concert with transcriptomic and proteomic data, are an important resource to comprehensively characterize spidroins as well as other proteins composing silks. High quality sequences are required to produce silk-like synthetic fibers using genetic engineering [14,15].

A particular challenge to assembling spidroin gene sequences is that they encode long, highly repetitive proteins. Of the five available spider genomes, only one (the social velvet spider) successfully assembled multiple full-length spidroins using short-read sequences alone [5]. This was attributed to the low heterozygosity of this highly inbred species. A promising solution for assembling full-length silk genes from an outbred species was to use genomic information from spidroin gene fragments for long-distance PCR of full-length genes [6••]. These PCR products were subsequently sequenced completely with Pacific Biosciences (SMRT) methods [6••]. This approach remarkably recovered 20 full-length

Figure 1



Established and targeted genomic resources across arachnids. Colored icons indicate availability of genomes (green), whole mount in situ hybridization techniques for RNA expression assays (purple), and functional tools (blue). High-value targets for genomic sequencing are indicated with yellow icons. **(a)** Simplified phylogeny of Chelicerata from Sharma *et al.* [2] showing well resolved nodes (black circles) and nodes recovered only by slowly evolving genes (gray circles). Photographs clockwise from top left: the horseshoe crab *Limulus polyphemus* (photo: P. Funch); the harvestman *Phalangium opilio* (photo: arthropod i5k wiki); the scorpion *Centruroides sculpturatus* (photo: G. Giribet); the vinegaroon *Mastigoproctus giganteus* (photo: G. Giribet); the cobweb spider *Parasteatoda tepidariorum* (photo: E.V.W. Setton); and the pseudoscorpion

Table 1

Summary statistics for available arachnid genomes deposited in NCBI. Three high-quality genomes of the 15 acarine genomes are shown with all other available arachnid genomes as of September 2017

Species	Length (Mb)	Coverage	Contig N50 (bp)	Scaffold N50 (bp)	Scaffolds	Genes[a]	NCBI Accession [Reference]
Stegodyphus mimosarum (Social velvet spider)	2738	70×	40 146	480 636	68 653	27 252	GCA_000611955.2
Ixodes scapularis (Deer tick)	1765	6×	2942	76 228	369 492	24 785	GCA_000208615.1
Tetranychus urticae (Red spider mite)	91	8.05×	212 780	2 993 488	641	13 382	GCA_000239435.1 [9]
Galendromus occidentalis (Western predatory mite)	152	17.7×	200 706	896 831	2211	12 289	GCA_000255335.1 [10]
Parasteatoda tepidariorum (Common house spider)	1445	137.7×	10 147	4 055 356	16 533	27 990	GCA_000365465.2 [7••]
Loxosceles reclusa (Brown recluse spider)	3262	86.0×	1834	63 237	143 665	20 617	GCA_001188405.1
Latrodectus hesperus (Western black widow)	1137	48.0×	2223	13 889	151 814	17 364	GCA_000697925.1
Nephila clavipes (Golden orb-weaver)	2439	98.5×	7993	62 959	180 236	14 025	GCA_002102615.1 [6••]
Centruroides sculpturatus (Bark scorpion)	926.4	181.1×	5173	342 549	10 457	30 465	GCA_000671375.1 [7••]
Mesobuthus martensii (Chinese scorpion)	925.5	200.0×	45 228	223.6 kb	14 798	32 016	GCA_000484575.1 [3]

[a] Gene counts for *L. reclusa*, *L. hesperus* and *C. sculpturatus* taken from Models.gff3 files corresponding to each species from https://i5k.nal.usda.gov/content/data-downloads.

spidroin genes from the golden orb-weaver *Nephila clavipes*, ranging in length from 1218 to17 817 bp of protein-encoding sequence.

Venom production in scorpions and spiders

Much scientific attention has focused on scorpion and spider venoms due to their utility for drug discovery, as a tool for biomedical research, and for treating hazardous bites [16,17]. Venoms are produced by nearly all scorpions and spiders to capture prey. Because these venoms are primarily composed of toxic proteins, sequence analyses integrating genomes, venom gland transcriptomes, and venom proteomes provide an unparalleled system for toxin characterization [18]. Further, venoms are fascinating from an evolutionary perspective because of their high molecular diversity [16,18]. Genomes present an essential resource to discern the degree to which venom molecular diversity is generated from multiple paralogs of gene families, polymorphic alleles, and/or alternative transcripts of loci.

The first genome-based analysis of arachnid venom in the Chinese scorpion *Mesobuthus martensii* [3] revealed that its venom neurotoxins are encoded by some of this species' largest gene families (≥116 genes) and uncovered many previously unknown neurotoxins. This also revealed scorpion genes for endogenous K^+ channels that have evolved resistance to their own venom toxins, providing a mechanism for venom self-immunity. Venomic analysis of the first spider genome from *Stegodyphus mimosarum*, similarly revealed a large number of venom genes, including 51 knottin-like toxin genes, 26 of which were confirmed as producing distinct toxins found in this species' venom [5]. The knottin-like (cystein knot) toxins typically bind to and disrupt neuronal ion channels, making them primary targets for drug discovery [19]. Surprisingly, spidroin proteins were also found in *S. mimosarum* venom. However, the expression of silk proteins in venom glands may be ubiquitous for spiders, since analysis of the golden orb-weaver (*N. clavipes*) genome found high expression of a flagelliform spidroin in venom glands [6••].

Genome-based analyses of venom proteins have also been carried out in the common house spider *Parasteatoda tepidariorum* [20•]. *P. tepidariorum* is confamilial with black widows *(Latrodectus)*, notorious for their highly toxic venom which is dominated by a family of neurotoxins (latrotoxins) previously only reported from *Latrodectus* and its sister genus *Steatoda*. Gendreau *et al.* [20•] found the house spider genome encodes a minimum of 47 latrotoxin genes, many of which are tandemly arrayed in close proximity. That these latrotoxins are also highly divergent from those in black widows illustrates the dynamic nature of venom gene evolution. Gendreau *et al.* [20•] also showed high sequence similarity between house spider latrotrotoxins and proteins from bacterial endosymbionts of arthropods, suggesting lateral transfer of latrotoxin genes.

Whole genome duplication

Gene duplication is arguably the most important contributor to evolving new functions. Whole genome duplications in vertebrates have been cited as instrumental in the diversification and evolution of complexity in this group [21,22]. In contrast, a whole genome duplication in horseshoe crabs is not correlated with morphological or phylogenetic diversity [23,24]. One of the first indicators that spiders experienced pervasive gene duplications came

(Figure 1 Legend Continued) *Synsphyronus apimelus* (photo: G. Giribet). **(b)** Simplified phylogeny of spiders from Garrison *et al.* [48•] showing major lineages (for detailed resolution of all derived lineages, the reader is referred to the Spider Tree of Life by Wheeler *et al.* [45•]). Note that the genome of *Acanthoscurria geniculata* [5] is not included here due to incompleteness. Photographs from top: the brown recluse *Loxosceles reclusa*; the black widow *Latrodectus hesperus*; and the golden orb weaver *Nephila clavipes* (photos: J. Ballesteros Chávez).

from analysis of a few *Hox* genes, which had duplicated and diverged in expression patterns in the wandering spider, *Cupiennius salei* [25]. A similar pattern was found for two of the *Wnt* ligand families in the house spider, *P. tepidariorum* [26]. Transcriptomic and whole-genome analysis of the scorpions, *Centruroides sculpturatus* and *M. martensii*, respectively, later revealed duplication of almost the entire *Hox* gene complement [27–29]. Analyses of multiple spider transcriptomes, and the genomes of a spider, a scorpion, and a tick indicated that gene duplications were pervasive in spiders, extending beyond these developmental patterning gene families and might have provided fodder for the diversification of silk production in spiders [30]. Duplications of microRNA genes also abounded in the house spider [31•]. However, the scale and timing of the gene duplication event was still ambiguous.

Excepting the inbred velvet spider, genomes from spiders and scorpions based on Illumina short-reads alone have resulted in comparatively poor assemblies (i.e. short scaffolds; Table 1) possibly because of their generation from outbred individuals. High quality sequencing of the house spider genome that included both short read sequencing and long-range 'Chicago' scaffolding [32] allowed the first synteny analysis of an arachnid genome. It revealed large tracts of syntenic segments on separate scaffolds (likely separate chromosomes), indicative of a whole genome duplication [7••]. In the same paper, the scorpion, *C. sculpturatus* genome was sequenced and analyzed, but the lack of long-range scaffolding precluded syntenic analysis. Nevertheless, the gene complement of this species was more complete than the previous genome of a scorpion (*Mesobuthus martensii*, [3]). Molecular evolutionary analyses of these two genomes with additional arachnid transcriptomes and genomes suggested that the whole genome duplication occurred in the ancestor of spiders and scorpions. Further analyses demonstrated that more distantly related arachnids, such as ticks and mites, showed no signature of a whole genome duplication and that the spider-scorpion duplication was independent of the one(s) in horseshoe crabs [7••]. A demonstration that every *Hox* gene paralog had diverged in expression in the house spider, further suggests that the whole genome duplication was important for morphological diversification of spiders and scorpions [7••].

Arachnid evo–devo and phylogeny

The field of developmental genetics is one of the proving grounds for validation of genomic sequencing and translational application of sequence data to functional tools. For *P. tepidariorum*, its newly sequenced genome represents a major step forward in its establishment as an emerging model organism. Existing resources for this species already include a detailed developmental staging system [33], capacity for high-quality whole mount in situ hybridization [34], and techniques for parental and embryonic RNA interference (e.g. [35,36]). The experimental tractability of this system was most recently exemplified by the works of Pechmann [37] and Pechmann *et al.* [38••], who performed single-cell injections at the 16-cell stage to induce ectopic gene expression via a capped mRNA construct. At the time of this writing, CRISPR-Cas9 mediated genome editing has not been achieved in *P. tepidariorum*. To date, no published experimental work has demonstrated the feasibility of injecting a 1-cell stage spider embryo [37]. Breaking this impasse is the next major challenge for researchers of spider developmental biology.

Other arachnid models for the study of development include the mites *Tetranychus urticae* [9] and *Archegozetes longisetosus* [39], the harvestman *Phalangium opilio* [40,41], and the scorpion *C. sculpturatus* [28]. However, manipulative tools like gene silencing and/or genomic resources are limited in some of these species. Due to the long gestation period and live bearing condition of scorpions, RNA interference is unlikely to be achieved in *C. sculpturatus*. Inversely, embryonic RNA interference was achieved early in the spider *C. salei*, but this former mainstay of developmental biology was superseded by the more tractable species *P. tepidariorum*, and now lacks a genome.

Beyond development, the advent of genomes is necessary for an improved understanding of arachnid relationships, which have proven recalcitrant to resolution in spite of interrogation with genome scale datasets [2,42]. Due to the incidence of four long-branch orders in the arachnid tree of life (mites, ticks, pseudoscorpions, and palpigrades), arachnid monophyly is difficult, albeit possible, to obtain in phylogenomic analyses ([2]; Figure 1). New genomes from the unrepresented arachnid orders are anticipated to provide both improvements in the inference of gene orthology, as well as a potential source of rare genomic changes (e.g. shared transposon insertions, duplicated Hox clusters) that may aid in phylogenetic resolution of the orders [7••].

Non-spider arachnid genomes needed

To date, sequenced genomes are available only for the orders Acariformes, Araneae, Parasitiformes, and Scorpiones. Of the eight arachnid orders that lack genomes, we propose the following three groups as high priority targets for sequencing.

The harvestman *Phalangium opilio* is an opportune representative of the order Opiliones. This species is synanthropic and broadly distributed throughout the Northern Hemisphere, as well as New Zealand. It is in active use as a laboratory model system for the study of development; techniques available in this species include gene and protein expression assays, and embryonic RNA

interference [43]. It is also a known generalist predator of such major agricultural pests as the soy aphid and the corn earworm.

A species of pseudoscorpion in the clade Iocheirata should also be prioritized, as this group of pseudoscorpions has evolved venom and silk separately from spiders and scorpions. The venom of these pseudoscorpions is produced from the pedipalp, whereas their silk glands are associated with the chelicerae. A natural exemplar species is the harlequin beetle-riding species *Cordylochernes scorpioides*, whose biology has long been under study [44].

Finally, a non-spider tetrapulmonate should be prioritized toward understanding the genetic basis for the evolution of the book lung. While lacking silk or venom glands, the order Uropygi is of biological interest due to the group's pygidial glands, which produce a noxious spray for repelling predators that is rich in ketones, caprylic acid, and acetic acid. The vinegaroon *Mastigoproctus giganteus* (order Uropygi) is a natural choice in this regard due to its large size, hardiness, and ability to be cultured in captivity.

Spider genomes needed

The available spider genomes were sequenced from representatives of four of the 112 presently described families (Ref. [45•]). Future genome sequencing projects should target representatives of all major spider clades (Figure 1b), to reveal key innovations in silk and venom production, provide resources for functional genetic studies, and improve spider phylogeny. Genome sequencing any members of the large RTA-clade, including scientifically important families such as jumping spiders (Salticidae), wolf spiders (Lycosidae), funnel-web spiders (Agelenidae), and wandering spiders (Ctenidae), is urgently needed (Figure 1b). Other targets should include a species of Deinopidae (net-casting spider) and Uloboridae (hackled orb-weavers) to resolve long-standing questions regarding silk evolution, as well as additional genomes from araneoid orb-weavers such as a member of Tetragnathidae and Araneidae. Genomes are also needed from representatives of the Haplogynae clade beyond the brown recluse in Sicariidae (e.g. from Hypochilidae, Dysderidae and Pholcidae); and mygalomorphs from Theraphosidae (tarantulas) and Hexathelidae (targeting the Sydney funnel-web spider due to its medically important venom). Finally, Liphistiidae, sister to all other spider families, should be prioritized due to its key phylogenetic position and retention of ancestral traits.

The selection of species to target should be balanced by practical considerations such as genome size and potential to obtain sufficient DNA from few closely related individuals to reduce allelic diversity and facilitate assembly. Additional funding should also be directed to improving the more fragmented spider genomes, specifically *L. hesperus* (Western black widow) and *Loxosceles reclusa* (brown recluse), with long-range SMRT sequencing, given their scientifically important venom.

A second complementary approach should target genomes from a few closely related species to discern rapid genomic changes that shape adaptions in silk and venom use. We suggest targeting multiple species from the genus *Latrodectus*, containing black widows, because several species have available transcriptomes [30], the preliminary genome for *L. hesperus* [4], and because the silk and venom of these species has been extensively studied due to the ease of collecting and lab-rearing *Latrodectus* [46•,47].

Conclusions

Genomic resources for arachnids are particularly sparse given their substantial phenotypic diversity and scientific importance. Advances in sequencing technologies have recently yielded the first spider and scorpion genomes, revealing exciting findings including whole genome duplications, mechanisms of venom self-resistance, and a much wider diversity of silk and venom genes. The varying quality of these genome assemblies emphasizes the need to incorporate long-range sequencing methods (e.g. SMRT and Chicago) when planning to sequence outbred species. The additional genomes we recommend for sequencing will transform arachnology into a genome-enabled discipline, providing a comparative framework essential to pinpoint genes underlying the many fascinating adaptations of arachnids.

Conflict of interest

The authors have no actual or potential conflict of interest including any financial, personal or other relationships with other people or organizations within three years of beginning the submitted work that could inappropriately influence, or be perceived to influence, their work.

Acknowledgements

We thank Stephen Richards for the invitation to submit this review article. Photographs in Figure 1 were provided by Emily Setton and Jesus Ballesteros Chávez (Sharma Lab, University of Wisconsin-Madison); Peter Funch (Aarhus University); and Gonzalo Giribet (Harvard University). Research originating from our labs was funded by the National Science Foundation (IOS-0951086 to NAA; IOS-1552610 to PPS), National Institutes of Health (2 R15 GM097714-02 to JEG; F32 GM78875-1A to NAA), and Washington and Lee University through Lenfest Summer Fellowships to NAA.

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Coddington JA, Giribet G, Harvey MS, Prendini L, Walter DE: **Arachnida**. In *Assembling the Tree of Life*. Edited by Cracraft J, Donoghue M. Oxford University Press; 2004:296-318.

2. Sharma PP, Kaluziak ST, Pérez-Porro AR, González VL, Hormiga G, Wheeler WC, Giribet G: **Phylogenomic interrogation of Arachnida reveals systemic conflicts in phylogenetic signal**. *Mol Biol Evol* 2014, **31**:2963-2984.

3. Cao Z, Yu Y, Wu Y, Hao P, Di Z, He Y, Chen Z, Yang W, Shen Z, He X *et al.*: **The genome of *Mesobuthus martensii* reveals a unique adaptation model of arthropods**. *Nat Commun* 2013, **4**:2602.

4. i5K Consortium: **The i5K initiative: advancing arthropod genomics for knowledge, human health, agriculture, and the environment**. *J Hered* 2013, **104**:595-600.

5. Sanggaard KW, Bechsgaard JS, Fang X, Duan J, Dyrlund TF, Gupta V, Jiang X, Cheng L, Fan D, Feng Y *et al.*: **Spider genomes provide insight into composition and evolution of venom and silk**. *Nat Commun* 2014, **5**.

6. •• Babb PL, Lahens NF, Correa-Garhwal SM, Nicholson DN, Kim EJ, Hogenesch JB, Kuntner M, Higgins L, Hayashi CY, Agnarsson I *et al.*: **The *Nephila clavipes* genome highlights the diversity of spider silk genes and their complex expression**. *Nat Genet* 2017, **49**:895-903.

Nephila clavipes is considered a workhorse for understanding spider silk synthesis and this paper represents the second published spider genome. Although short-read sequencing resulted in a fragmented genome, targeted long-range PCR and SMRT sequencing recovered a remarkable spider silk complement of 28 full or nearly full length genes.

7. •• Schwager EE, Sharma PP, Clarke T, Leite DJ, Wierschin T, Pechmann M, Akiyama-Oda Y, Esposito L, Bechsgaard J, Bilde T *et al.*: **The house spider genome reveals an ancient whole-genome duplication during arachnid evolution**. *BMC Biol* 2017, **15**:62.

The house spider is an emerging model system for chelicerate development and its genome represents one of the most complete for an arachnid. Especially notable were the ability to identify large syntenic tracts indicative of a whole genome duplication, dating the whole genome duplication to the ancestor of spiders and scorpions, and demonstrating the consequences of duplication for expression patterns of the developmental patterning Hox genes.

8. Brewer MS, Cotoras DD, Croucher PJP, Gillespie RG: **New sequencing technologies, the development of genomics tools, and their applications in evolutionary arachnology**. *J Arachnol* 2014, **42**:1-15.

9. Grbic M, Leeuwen T, Clark RM, Rombauts S, Rouze P, Grbic V, Osborne EJ, Dermauw W, Ngoc PC, Ortego F: **The genome of *Tetranychus urticae* reveals herbivorous pest adaptations**. *Nature* 2011, **479**.

10. Hoy MA, Waterhouse RM, Wu K, Estep AS, Ioannidis P, Palmer WJ, Pomerantz AF, Simão FA, Thomas J, Jiggins FM *et al.*: **Genome sequencing of the phytoseiid predatory mite *Metaseiulus occidentalis* reveals completely atomized Hox genes and superdynamic intron evolution**. *Genome Biol Evol* 2016, **8**:1762-1775.

11. Gulia-Nuss M, Nuss AB, Meyer JM, Sonenshine DE, Roe RM, Waterhouse RM, Sattelle DB, Fuente J, Ribeiro JM, Megy K: **Genomic insights into the *Ixodes scapularis* tick vector of Lyme disease**. *Nat Commun* 2016, **7**.

12. Blackledge TA, Hayashi CY: **Silken toolkits: biomechanics of silk fibers spun by the orb web spider *Argiope argentata* (Fabricius 1775)**. *J Exp Biol* 2006, **209**:2452-2461.

13. Jain D, Zhang C, Cool LR, Blackledge TA, Wesdemiotis C, Miyoshi T, Dhinojwala A: **Composition and function of spider glues maintained during the evolution of cobwebs**. *Biomacromolecules* 2015, **16**:3373-3380.

14. Andersson M, Jia Q, Abella A, Lee X-Y, Landreh M, Purhonen P, Hebert H, Tenje M, Robinson CV, Meng Q *et al.*: **Biomimetic spinning of artificial spider silk from a chimeric minispidroin**. *Nat Chem Biol* 2017, **13**:262-264.

15. Krishnaji ST, Kaplan DL: **Bioengineered chimeric spider silk-uranium binding proteins**. *Macromol Biosci* 2013, **13**:256-264.

16. King GF, Hardy MC: **Spider-venom peptides: structure, pharmacology, and potential for control of insect pests**. *Annu Rev Entomol* 2013, **58**:475-496.

17. Dardevet L, Rani D, Abd El Aziz T, Bazin I, Sabatier J-M, Fadl M, Brambilla E, De Waard M: **Chlorotoxin: a helpful natural scorpion peptide to diagnose glioma and fight tumor invasion**. *Toxins* 2015, **7**:1079-1101.

18. Casewell NR, Wüster W, Vonk FJ, Harrison RA, Fry BG: **Complex cocktails: the evolutionary novelty of venoms**. *Trends Ecol Evol* 2013, **28**:219-229.

19. Saez NJ, Senff S, Jensen JE, Er SY, Herzig V, Rash LD, King GF: **Spider-venom peptides as therapeutics**. *Toxins* 2010, **2**:2851-2871.

20. • Gendreau KL, Haney RA, Schwager EE, Wierschin T, Stanke M, Richards S, Garb JE: **House spider genome uncovers evolutionary shifts in the diversity and expression of black widow venom proteins associated with extreme toxicity**. *BMC Genomics* 2017, **18**:178.

Discovered an unexpectedly large diversity of latrotoxin (black widow venom neurotoxin) genes in the common house spider genome, as well as potential evidence of lateral gene transfer with bacteria.

21. Putnam NH, Butts T, Ferrier DE, Furlong RF, Hellsten U, Kawashima T, Robinson-Rechavi M, Shoguchi E, Terry A, Yu JK: **The amphioxus genome and the evolution of the chordate karyotype**. *Nature* 2008, **453**.

22. Hakes L, Pinney JW, Lovell SC, Oliver SG, Robertson DL: **All duplicates are not equal: the difference between small-scale and genome duplication**. *Genome Biol* 2007, **8**.

23. Nossa CW, Havlak P, Yue JX, Lv J, Vincent KY, Brockmann HJ, Putnam NH: **Joint assembly and genetic mapping of the Atlantic horseshoe crab genome reveals ancient whole genome duplication**. *GigaScience* 2014, **3**.

24. Kenny NJ, Chan KW, Nong W, Qu Z, Maeso I, Yip HY, Chan TF, Kwan HS, Holland PW, Chu KH: **Ancestral whole-genome duplication in the marine chelicerate horseshoe crabs**. *Heredity* 2016, **116**.

25. Schwager EE, Schoppmeier M, Pechmann M, Damen WG: **Duplicated Hox genes in the spider *Cupiennius salei***. *Front Zool* 2007, **4**.

26. Janssen R, Gouar M, Pechmann M, Poulin F, Bolognesi R, Schwager EE, Hopfen C, Colbourne JK, Budd GE, Brown SJ: **Conservation, loss, and redeployment of Wnt ligands in protostomes: implications for understanding the evolution of segment formation**. *BMC Evol Biol* 2010, **10**.

27. Sharma PP, Santiago MA, Gonzalez-Santillan E, Monod L, Wheeler WC: **Evidence of duplicated Hox genes in the most recent common ancestor of extant scorpions**. *Evol Dev* 2015, **17**.

28. Sharma PP, Schwager EE, Extavour CG, Wheeler WC: **Hox gene duplications correlate with posterior heteronomy in scorpions**. *Proc R Soc Lond B Biol Sci* 2014, **281**:20140661.

29. Di Z, Yu Y, Wu Y, Hao P, He Y, Zhao H, Li Y, Zhao G, Li X, Li W: **Genome-wide analysis of homeobox genes from *Mesobuthus martensii* reveals Hox gene duplication in scorpions**. *Insect Biochem Mol Biol* 2015, **61**.

30. Clarke TH, Garb JE, Hayashi CY, Arensburger P, Ayoub NA: **Spider transcriptomes identify ancient large-scale gene duplication event potentially important in silk gland evolution**. *Genome Biol Evol* 2015, **7**.

31. • Leite DJ, Ninova M, Hilbrant M, Arif S, Griffiths-Jones S, Ronshaugen M, McGregor AP: **Pervasive microRNA duplication in Chelicerates: insights from the embryonic microRNA repertoire of the spider *Parasteatoda tepidariorum***. *Genome Biol Evol* 2016, **8**.

Comprehensive characterization of micro RNAs (miRNA) expressed in embryos of the house spider identified 101 conserved miRNA families and 47 new families. Comparative analysis with published chelicerate genomes further demonstrated that most of the new families were unique to the house spider, that miRNA families conserved with spiders and scorpions were often duplicated, and that the duplicated miRNAs had diverged in arm usage indicative of shifts in function.

32. Putnam NH, O'Connell BL, Stites JC, Rice BJ, Blanchette M, Calef R, Troll CJ, Fields A, Hartley PD, Sugnet CW: **Chromosome-**

scale shotgun assembly using an in vitro method for long-range linkage. *Genome Res* 2016, **26**.

33. Mittmann B, Wolff C: **Embryonic development and staging of the cobweb spider *Parasteatoda tepidariorum* C. L. Koch, 1841 (syn.: *Achaearanea tepidariorum*; Araneomorphae; Theridiidae)**. *Dev Genes Evol* 2012, **222**.

34. Akiyama-Oda Y, Oda H: **Multi-color FISH facilitates analysis of cell-type diversification and developmental gene regulation in the *Parasteatoda* spider embryo**. *Dev Growth Differ* 2016, **58**:215-224.

35. Schwager EE, Meng Y, Extavour CG: ***vasa* and *piwi* are required for mitotic integrity in early embryogenesis in the spider *Parasteatoda tepidariorum***. *Dev Biol* 2015, **402**:276-290.

36. Turetzek N, Pechmann M, Schomburg C, Schneider J, Prpic NM: **Neofunctionalization of a duplicate *dachshund* gene underlies the evolution of a novel leg segment in arachnids**. *Mol Biol Evol* 2016, **33**.

37. Pechmann M: **Formation of the germ-disc in spider embryos by a condensation-like mechanism**. *Front Zool* 2016, **13**:35.

38. •• Pechmann M, Benton MA, Kenny NJ, Posnien N, Roth S: **A novel role for *Ets4* in axis specification and cell migration in the spider *Parasteatoda tepidariorum***. *eLife* 2017, **6**.
Using single-cell injections of capped mRNAs at 16-cell stages, the authors demonstrate that *Ets4* is necessary and sufficient for the integrity and migration of the cumulus, the spider's equivalent of an embryonic organizer. This work draws upon genomic, transcriptomic, and functional tools in *P. tepidariorum*, toward demonstrating the full capabilities of this emerging model system.

39. Barnett AA, Thomas RH: **Posterior Hox gene reduction in an arthropod: *Ultrabithorax* and *Abdominal-B* are expressed in a single segment in the mite *Archegozetes longisetosus***. *EvoDevo* 2013, **4**:23.

40. Sharma PP, Schwager EE, Extavour CG, Giribet G: **Hox gene expression in the harvestman *Phalangium opilio* reveals divergent patterning of the chelicerate opisthosoma**. *Evol Dev* 2012, **14**.

41. Sharma PP, Schwager EE, Giribet G, Jockusch EL, Extavour CG: ***Distal-less* and *dachshund* pattern both plesiomorphic and apomorphic structures in chelicerates: RNA interference in the harvestman *Phalangium opilio* (Opiliones)**. *Evol Dev* 2013, **15**:228-242.

42. Regier JC, Shultz JW, Zwick A, Hussey A, Ball B, Wetzer R, Martin JW, Cunningham CW: **Arthropod relationships revealed by phylogenomic analysis of nuclear protein-coding sequences**. *Nature* 2010, **463**.

43. Sharma PP, Tarazona OA, Lopez DH, Schwager EE, Cohn MJ, Wheeler WC, Extavour CG: **A conserved genetic mechanism specifies deutocerebral appendage identity in insects and arachnids**. *Proc Biol Sci* 2015, **282**:20150698.

44. Zeh JA, Bonilla MM, Su EJ, Padua MV, Anderson RV, Zeh DW: **Constant diurnal temperature regime alters the impact of simulated climate warming on a tropical pseudoscorpion**. *Sci Rep* 2014, **4**.

45. • Wheeler WC, Coddington JA, Crowley LM, Dimitrov D, Goloboff PA, Griswold CE, Hormiga G, Prendini L, Ramírez MJ, Sierwald P *et al.*: **The spider tree of life: phylogeny of Araneae based on target-gene analyses from an extensive taxon sampling**. *Cladistics* 2016 http://dx.doi.org/10.1111/cla.12182.
The most taxon-rich resolution of the spider tree of life to date, this analysis— the fruit of nearly 15 years of work— sampled six Sanger sequenced loci from 932 specimens, representing 115 of the 116 spider families (with four new family designations proposed).

46. • Clarke TH, Garb JE, Haney RA, Chaw RC, Hayashi CY, Ayoub NA: **Evolutionary shifts in gene expression decoupled from gene duplication across functionally distinct spider silk glands**. *Sci Rep* 2017, **7**:8393.
RNA-sequencing of three cobweb weaving spider species' silk glands and three other tissues provided the first comprehensive analyses of gene expression patterns in each of the seven functionally differentiated silk gland types. Reconstruction of ancestral expression states demonstrated that evolution of tissue-specific gene expression can arise through multiple pathways, with duplication followed by a shift in gene expression playing a much more minor role than previously thought.

47. Haney RA, Clarke TH, Gadgil R, Fitzpatrick R, Hayashi CY, Ayoub NA, Garb JE: **Effects of gene duplication, positive selection, and shifts in gene expression on the evolution of the venom gland transcriptome in widow spiders**. *Genome Biol Evol* 2016, **8**:228-242.

48. • Garrison NL, Rodriguez J, Agnarsson I, Coddington JA, Griswold CE, Hamilton CA, Hedin M, Kocot KM, Ledford JM, Bond JE: **Spider phylogenomics: untangling the Spider Tree of Life**. *PeerJ* 2016, **4**:e1719.
Analysis of nearly 3400 gene regions improved resolution of the basal tree of spiders and demonstrated accelerated rates of evolution in the asymmetrically diverse RTA-clade, and separately in the Mygalomorphae. This work also suggested an older evolutionary origin of the orb web than previously inferred.