

1 **A SoxB gene acts as an anterior gap gene and**
2 **regulates posterior segment addition in the spider**
3 ***Parasteatoda tepidariorum***

4

5 Christian L. B. Paese¹, Anna Schoenauer¹, Daniel J. Leite¹, Steven Russell²,
6 Alistair P. McGregor^{1*}.

7

8 1. Laboratory of Evolutionary Developmental Biology, Department of
9 Biological and Medical Sciences, Oxford Brookes University, Gipsy Lane,
10 Oxford, OX3 0BP, UK.

11

12 2. Department of Genetics, University of Cambridge, Downing Street,
13 Cambridge, CB2 3EH, UK.

14

15 *To whom correspondence should be addressed: amcgregor@brookes.ac.uk
16 (A.P.M.)

17

18 Running Head: Sox gene regulation of spider segmentation

19

20 Key words. Sox genes, *Parasteatoda tepidariorum*, spider, evolution,
21 development.

22 **Summary**

23 The Sox gene family encode a set of highly conserved HMG domain
 24 transcription factors that regulate many key processes during metazoan
 25 embryogenesis. In insects, the SoxB gene *Dichaete* is the only Sox gene
 26 known to be involved in embryonic segmentation. To determine if similar
 27 mechanisms are used in other arthropods, we investigated the role of Sox
 28 genes during segmentation in the spider *Parasteatoda tepidariorum*. While
 29 *Dichaete* does not appear to be involved in spider segmentation, RNAi
 30 knockdown of the closely related *Sox21b-1* gene results in a gap like
 31 phenotype in the developing prosoma and also perturbs the sequential
 32 addition of opisthosomal segments. We show that this is in part due to a role
 33 for *Sox21b-1* in regulating the expression of *Wnt8* and influencing Delta-Notch
 34 signalling during the formation of the segment addition zone. Thus, we have
 35 found that two different mechanisms for segmentation in a non-mandibulate
 36 arthropod are regulated by a Group B Sox gene. Our work provides new
 37 insights into the function of an important and conserved gene family across
 38 arthropods, and the evolution of the regulation of segmentation in these
 39 animals.

40

41

42 Introduction

43 Arthropods are the most speciose and widespread of the animal phyla, and it
 44 is thought that their diversification and success is at least in part explained by
 45 their segmented body plan (1). In terms of development, insects utilise either
 46 derived long germ embryogenesis, where all body segments are made more
 47 or less simultaneously, or short/intermediate germ embryogenesis, where a
 48 few anterior segments are specified and posterior segments are added
 49 sequentially from a growth or segment addition zone (SAZ) (2, 3). It is thought
 50 that segmentation in the ancestral arthropod resembled the short germ mode
 51 seen in most insects (2, 4). Understanding the regulation of segmentation
 52 more widely across the arthropods is important for understanding both the
 53 development and evolution of these highly successful animals.

54 We have a detailed and growing understanding of the regulation of
 55 segmentation in various insects, especially the long germ dipteran *Drosophila*
 56 *melanogaster* and the short germ beetle *Tribolium castaneum*. However,
 57 studies of other arthropods including the myriapods *Strigamia maritima* and
 58 *Glomeris marginata*, and chelicerates, such as the spiders *Cupiennius salei*
 59 and *Parasteatoda tepidariorum*, have provided important mechanistic and
 60 evolutionary insights into arthropod segmentation (2, 5-9). Previous studies
 61 have shown that different genetic mechanisms are used to generate
 62 segments along the anterior-posterior axis of spider embryos. In the anterior
 63 tagma, the prosoma or cephalothorax, the cheliceral and pedipalpal segments
 64 are generated by dynamic waves of *hedgehog* (*hh*) and *orthodenticle* (*otd*)
 65 expression (10, 11). The leg bearing segments are specified by gap gene like
 66 functions of *hunchback* (*hb*) and *distal-less* (*dll*) (12, 13). In contrast, the

67 segments of the posterior tagma, the opisthosoma or abdomen, are
 68 generated sequentially from a SAZ. This process is regulated by dynamic
 69 interactions between Delta-Notch and Wnt8 signalling to regulate *caudal*
 70 (*cad*), which in turn is required for oscillatory expression of pair-rule gene
 71 orthologues including *even-skipped* (*eve*), and *runt* (*run*) (4, 14, 15).
 72 Interestingly, these pair-rule gene orthologues are not involved in the
 73 production of the prosomal segments (15). Therefore, the genetic regulation
 74 of segmentation along the anterior-posterior axis in the spider exhibits
 75 similarities and differences to segmentation in both long germ and short germ
 76 insects.

77 The Group B Sox family gene *Dichaete* is required for correct
 78 embryonic segmentation in the long germ *D. melanogaster*, where it regulates
 79 pair-rule gene expression (16,17). Interestingly, it was recently discovered
 80 that a *Dichaete* orthologue is also involved in segmentation in the short germ
 81 *T. castaneum* (18). This similarity is consistent with work inferring that these
 82 modes of segmentation are more similar than previously thought and provides
 83 insights into how the long germ mode evolved (18, 19, 20). However, it
 84 appears that despite these similarities, *Dichaete* can play different roles in *D.*
 85 *melanogaster* and *T. castaneum* consistent with the generation of segments
 86 simultaneously via a gap gene mechanism in the former and sequentially from
 87 a posterior SAZ in the latter (18).

88 We recently described the discovery of 14 Sox genes in the genome of
 89 the spider *P. tepidariorum* (21) and that several of the spider Sox genes are
 90 represented by multiple copies likely produced during the whole genome
 91 duplication (WGD) in the lineage leading to this arachnid (21). Interestingly,

92 while *Dichaete* is not expressed in a pattern consistent with a role in
 93 segmentation (21), we found that the closely related SoxB gene, *Sox21b-1*, is
 94 expressed in both the prosoma and opisthosoma before and during
 95 segmentation (21). Here we report that in *P. tepidariorum*, *Sox21b-1*,
 96 regulates both prosomal and opisthosomal segmentation. In the prosoma,
 97 *Sox21b-1* has a gap-like gene function and is required for the generation of
 98 the four leg bearing segments. In addition, *Sox21b-1* appears to act upstream
 99 of both *Delta* and *Wnt8* to regulate the formation of the SAZ, and knockdown
 100 of *Sox21b-1* results in truncated embryos missing all opisthosomal segments.
 101 Therefore, while prosomal and opisthosomal segments are generated by
 102 different mechanisms in the spider, our analysis shows that *Sox21b-1* is
 103 required for segmentation in both regions of the developing spider embryo.

104

105

106 RESULTS

107 ***Sox21b-1* is maternally deposited and is subsequently expressed in the** 108 **germ disc and germ band of spider embryos**

109 We previously identified and assayed the expression of the complement of
110 Sox genes in the genome of the spider *P. tepidariorum* (21). Our phylogenetic
111 analysis indicates that *P. tepidariorum Sox21b-1* and its paralog *Sox21b-2* are
112 members of the Sox group B, closely related to the *Drosophila Dichaete* and
113 *Sox21b* genes (Fig. S1). In insects (22, 23), *Dichaete*, *Sox21a* and *Sox21b*
114 are clustered in the genome, however, both *Sox21b* paralogs are dispersed in
115 the spider genome (21). This suggests that *Sox21b-1* and *Sox21b-2* possibly
116 arose from the WGD event in the ancestor of arachnoplumonates (24), rather
117 than by a more recent tandem duplication (20) (Fig. S1).

118 In light of its interesting expression pattern we elected to analyse
119 *Sox21b-1* further. Pre-vitellogenic *P. tepidariorum* oocytes contain a Balbiani's
120 body (25), where maternally deposited factors are enclosed, and we found
121 that *Sox21b-1* is abundant in this region, indicating that it is maternally
122 contributed (Fig. 1A). However, after fertilization *Sox21b-1* is not expressed
123 again until early stage 5, when weak expression is detected throughout the
124 germ disc, with stronger expression in more central cells (Fig. 1B). At late
125 stage 5, expression becomes more restricted to the centre of the germ disc
126 (Fig. 1C). During stages 5 and 6, the cumulus migrates to the rim of the germ
127 disc, opening the dorsal field and giving rise to an axially symmetric germ
128 band (Fig. 1D) (see 26). In early stage 6 embryos, *Sox21b-1* is observed in
129 the middle of the presumptive prosoma in a broad stripe (Fig. 1D), which
130 develops further during stage 7 in the region where the leg bearing segments

131 will form (Fig. 1E). This expression pattern resembles the previously
132 described expression of the gap gene *hb* (13).

133 During these and subsequent stages, dynamic expression of *Sox21b-1*
134 is observed in the SAZ and the most anterior region of the germ band that will
135 give rise to the head segments (Fig. 1F-H). Later in development, the
136 expression of *Sox21b-1* resembles that of *SoxNeuro* (*SoxN*), another Group B
137 Sox gene (21). This expression is similar to that of both *SoxN* and *Dichaete* in
138 *D. melanogaster*, which are expressed in neuroblasts of the neuroectoderm
139 and then differentiating neurons in the ventral nerve cord (22) (Fig. 1G, H).
140 Expression of the related group B Sox genes, *Dichaete* and *Sox21b-2* are not
141 detected in *P. tepidariorum* during embryonic development (20). The
142 expression of *Sox21b-1* in the embryo suggests that it is involved in both
143 anterior and posterior segmentation in this spider, as well as later during
144 nervous system development.

145

146 ***Sox21b-1* regulates prosomal and opisthosomal segmentation**

147 To assay the function of *Sox21b-1* during embryogenesis we knocked down
148 the expression of the gene using a parental RNAi approach (27). We
149 observed three phenotypic classes, which were consistent between both of
150 the non-overlapping *Sox21b-1* fragments we used for RNAi (Fig. 2; Fig. S2;
151 Table S2).

152 Class I embryos developed a presumptive head region (Fig. 2A-C), as
153 well as normal cheliceral, pedipalpal and first leg bearing (L1) segments (Fig.
154 2C). The identity of these segments was confirmed by expression of *labial*
155 (*lab*) in the pedipalps and L1, and *Deformed-A* (*Dfd-A*) in L1 (Fig S3A-B).

156 However, the other three leg bearing segments, L2 - L4, as well as all of the
157 opisthosomal segments were missing in Class I embryos. These embryos
158 exhibited a truncated germ band, terminating in disorganised tissue in the
159 region of the SAZ (Fig. 2C). In the case of Class II phenotypes, embryos only
160 differentiated the head region and the cheliceral and pedipalpal segments
161 (Fig. 2D; Fig S3A-B): all leg bearing segments of the prosoma and
162 opisthosomal segments produced from the SAZ were missing (Fig. 2D). In
163 Class III embryos, the germ band did not form properly from the germ disc
164 (Fig. 2E) and we therefore looked earlier in development to understand how
165 this severe phenotype arose. We observed that the formation of the primary
166 thickening occurs normally at stage 4 (27, 28, 29), but subsequently the
167 cumulus, the group of mesenchymal cells that arise as the primary thickening
168 at the centre of the germ disc, fails to migrate properly to the rim of the germ
169 disc during stage 5 (Fig. S4). Since migration of the cumulus is required for
170 the transition from germ disc to germ band, this observation at least in part
171 explains the subsequent Class III phenotype. In some embryos, we did
172 observe the opening of the dorsal field in stage 6 embryos: therefore, we
173 suggest these embryos later develop Class I and II phenotypes (Fig. S4B-C).

174 We next examined the effect of *Sox21b-1* depletion on cell death and
175 proliferation at stages 5 and 9 in knockdown and control embryos using
176 antibodies against Caspase-3 and phosphorylated Histone 3 (PHH3) (Fig.
177 S5). At the germ disc stage there is no detectable cell death in control
178 embryos (n = 10), but we observed some small clusters of apoptotic cells in
179 the *Sox21b-1* knockdown embryos (n = 10) (Fig. S5A-B). At stage 9, a few
180 cells expressed Caspase-3 in the posterior-most part of the SAZ (Fig. S5C),

181 but we did not observe cell death in this region of *Sox21b-1* knockdown
182 embryos (Fig. S5D). However, we did observe pronounced cell death in the
183 anterior extraembryonic layer of the same embryos, (n = 10) (Fig. S5D).

184 Expression of PHH3 at stages 5 and 9, indicated that *Sox21b-1*
185 knockdown embryos show decreased cell proliferation compared to controls
186 (n= 10 for each) (Fig. S5E-H). Interestingly the cells were also clearly larger in
187 *Sox21b-1* knockdown embryos compared to controls, which may reflect
188 perturbed cell proliferation (Fig. S5E-H). Thus, our functional analysis shows
189 that *Sox21b-1* regulates cell proliferation and the transition from radial to axial
190 symmetry. Moreover, *Sox21b-1* is involved in two different segmentation
191 mechanisms in spiders: it has a gap gene like function in the prosoma, as well
192 as a requirement for the formation of the SAZ and subsequent production of
193 opisthosomal segments.

194

195 **Effects of *Sox21b-1* knockdown on the germ disc and mesoderm**

196 In *P. tepidariorum*, *decapentaplegic* (*dpp*) and *Ets4* are required for cumulus
197 formation (29, 30). To investigate if *Sox21b-1* is involved in the formation of
198 this cell cluster we assayed the expression of *dpp* and *ets4* in *Sox21b-1* RNAi
199 knockdown embryos. However, both genes were expressed normally and
200 cumulus formation was unaffected (Fig. 3 E, F).

201 The rim of the spider germ disc develops into the head structures and
202 is regulated in part by *hh*, while the mesodermal and endodermal layers of the
203 head are specified by the mesendodermal gene *forkhead* (*fkh*) (10, 27). To
204 investigate if anterior expression of *Sox21b-1* (Fig. 1B) is involved in the
205 formation of the head rudiment and differentiation of the mesodermal and

206 endodermal layers in particular, we assayed the expression of *hh* and *fkh* in
207 class I and II *Sox21b-1* knockdown embryos.

208 *hh* is expressed at the rim of the germ disc in the ectoderm (Fig. 3 D)
209 (31) and remains unaffected by *Sox21b-1* knockdown (Fig. 3H). *fkh* is also
210 expressed in cells around the rim, as well as in the centre of the germ disc in
211 mesendodermal cells (Fig. 3C). In *Sox21b-1* knockdown embryos both of
212 these *fkh* expression domains are lost (Fig. 3G) and it therefore appears that
213 *Sox21b-1* is required for specification of mesendodermal cells in the germ
214 disc of spider embryos. Indeed in the germ disc at stage 5, when *fkh*
215 expression commences, we observed invaginating cells forming a second
216 layer (Fig. S3G), however, in *Sox21b-1* knockdown embryos we observed a
217 lower number of invaginating cells, which exhibit bigger nuclei compared to
218 the controls (Fig. S3H).

219 In both spiders and flies, the *twist* (*twi*) gene is involved in mesoderm
220 specification (32) and we therefore examined the expression of this gene after
221 *Sox21b-1* knockdown to further evaluate if the loss of *fkh* affects the formation
222 of the internal layers. *twi* is expressed in the visceral mesoderm of the limb
223 buds from L1 to L4, in the opisthosomal segments O1 to O4 and in an anterior
224 mesodermal patch in the central part of the developing head (32) (Fig. S3E).
225 While the head expression persists in *Sox21b-1* class I embryos, expression
226 in all the limb and opisthosomal segments is lower or absent (Fig. S3F). In
227 orthogonal projections the anterior-most region of the embryo three layers of
228 cells can be identified in control embryos (Fig. S3I). However, in *Sox21b-1*
229 knockdown embryos the formation of these layers is perturbed (Fig. S3J).

230 These data suggest that the ectodermal segmentation in the prosomal region
231 occurs even upon a reduction of the internal layers of the embryo.

232

233 **Effects of *Sox21b-1* knockdown on segmentation**

234 In *P. tepidariorum*, formation of the SAZ and production of posterior segments
235 requires the Wnt8 and Delta-Notch signalling pathways (14, 15). Interactions
236 between these pathways regulate *hairy* (*h*) and, via *cad*, the expression of
237 pair-rule gene orthologues including *eve* (14, 15). To better understand the
238 loss of segments we observe in *Sox21b-1* knockdown embryos we analysed
239 the expression of *Dl*, *Wnt8*, *h* and *cad* in these embryos compared to controls.

240 *Dl* is expressed at stages 5 and 6 in the forming SAZ, in the region of
241 the L4 primordia, and in the presumptive head (33) (Fig. 4A). Subsequently at
242 stage 9, *Dl* expression is visible in clusters of differentiating neuronal cells and
243 oscillates in the SAZ, an expression pattern associated with the sequential
244 addition of new segments (Fig. 4B). In *Sox21b-1* knockdown embryos, *Dl*
245 expression is not detected at stage 5 (Fig. 4C) and is absent in the posterior
246 at stage 9 (Fig. 4D). However, expression in the anterior neuroectoderm
247 seems normal up to the pedipalpal segment, although neurogenesis is
248 apparently perturbed in the presumptive L1 segment (Fig. 4D). This suggests
249 that the ectoderm up to the L1 segment differentiates normally, but the
250 development of the SAZ and posterior segment addition controlled by *Dl* is
251 lost upon *Sox21b-1* knockdown.

252 *Wnt8* is initially expressed at stage 5 in the centre and at the rim of the
253 germ disc (Fig. 4E). At stage 9, striped expression of *Wnt8* is seen from the
254 head to the posterior segments and in the posterior cells of the SAZ (Fig. 4G).

Knockdown of *Sox21b-1* results in the loss of *Wnt8* expression in late stage 5 embryos (Fig. 4F). At stage 9, *Wnt8* expression is observed in the cheliceral, pedipalpal and first walking limb segments of *Sox21b-1* knockdown embryos, but no expression is detected in the remaining posterior cells (Fig. 4H). Consistent with the loss of *DI* and *Wnt8*, *cad* expression is also lost in stage 5 and stage 9 *Sox21b-1* knockdown embryos (Fig. 4I-L). These observations indicate that *Sox21b-1* acts upstream of *Wnt8* and Delta-Notch signalling to regulate the formation of the SAZ and the subsequent production of posterior segments. In support of this regulatory relationship we find that *Sox21b-1* expression is still detected in the posterior regions of the truncated embryos produced by RNAi knockdown of either *DI* or *Wnt8* (Fig. 5).

The spider orthologue of *h* is expressed in the presumptive L2-L4 segments and dynamically in the SAZ (14) (Fig. 4M, O). In late stage 5 *Sox21b-1* knockdown embryos, the expression of *h* is lost throughout the entire germ disc (Fig. 4N). In addition, in Class I phenotype embryos at stage 9, the expression of *h* is completely absent in the tissue posterior to the pedipalpal segment (Fig. 4P). Therefore the loss of *h* expression is consistent with the loss of leg bearing segments in the anterior gap-like phenotype that results from knockdown of *Sox21b-1* as well as loss of segments produced by the SAZ.

To look at the effect of *Sox21b-1* knockdown on segmentation in more detail we examined the expression of *engrailed* (*en*) and *hh*. At stage 9 *en* is expressed segmentally from the cheliceral to the O3 segment in control embryos (Fig. 4Q). However, in *Sox21b-1* knockdown embryos, expression of *en* was only observed in the cheliceral, pedipalpal and L1 segments,

280 consistent with the loss of all the more posterior segments (Fig. 4S). *hh* has a
 281 similar expression pattern to *en* at stage 9, except it exhibits an anterior
 282 splitting wave in the cheliceral segment and is also expressed earlier in
 283 opisthosomal segments and in the SAZ (Fig. 4R). Upon *Sox21b-1*
 284 knockdown, *hh* is only detected in shortened stripes in the cheliceral and
 285 pedipalpal segments (Fig. 4T).

286 Taken together, our analysis of *P. tepidariorum* embryos where
 287 *Sox21b-1* is depleted by parental RNAi reveals an important role for this
 288 Group B Sox gene in both gap-like segmentation of the prosoma, as well as
 289 posterior segment formation from the SAZ. These experiments further
 290 emphasise the critical role this class of transcription factors play in arthropod
 291 segmentation.

292

293 Discussion

294 *A SoxB gene is required for two different mechanisms of spider segmentation*

295 The Sox (Sry-Related High-Mobility Group box) gene family encodes
296 transcription factors that regulate many important processes underlying the
297 embryonic development of metazoans (34-37). One such gene, *Dichaete*, is
298 expressed in a gap gene pattern and is involved in regulating the canonical
299 segmentation cascade in *D. melanogaster* (16, 17). Recently, the analysis of
300 the expression of *Dichaete* in the flour beetle *T. castaneum* strongly suggests
301 a role in short germ segmentation (18), further supported by knockdown of the
302 *Dichaete* orthologue in *Bombyx mori*, which resulted in the loss of posterior
303 segmentation (38).

304 Here we show that, while *Dichaete* is not involved in spider
305 segmentation (21), the closely related SoxB gene, *Sox21b-1*, regulates
306 formation of both prosomal and opisthosomal segments. In the prosoma
307 *Sox21b-1* has a gap gene role and is required for the specification of L1-L4
308 segments (Fig. 6), resembling the roles of *hb* and *Dll* in prosomal
309 segmentation in this spider (12, 13) and, at least superficially, gap gene
310 function in *Drosophila*.

311 In *Drosophila* the gap genes regulate pair rule gene expression, and
312 while our results indicate that *Sox21b-1* is required for the expression of *h* and
313 the generation of leg bearing prosomal segments (Fig 4E; Fig. 6), in contrast
314 to insects, in spiders this does not involve the orthologues of *eve* and *runt*
315 because they are not expressed in the developing prosomal segments (15,
316 39).

317 In the posterior, *Sox21b-1* knockdown perturbs SAZ formation and
 318 consequently results in truncated embryos missing all opisthosomal
 319 segments. Therefore, *Sox21b-1* regulates development of the SAZ, and our
 320 observations indicate this is at least in part through roles in organising the
 321 germ layers and specification of mesendodermal cells during stages 5 and 6.
 322 This is supported by the loss of *fkf* expression upon *Sox21b-1* knockdown,
 323 which is required for mesoderm and endoderm formation in both spiders and
 324 insects (10, 40, 41). Moreover, the subsequent dynamic expression of
 325 *Sox21b-1* in the SAZ after stage 6 is suggestive of a role in segment addition.

326 Our work on *Sox21b-1* provides an important new insight into the gene
 327 regulatory network (GRN) underlying the formation of the SAZ and the
 328 sequential addition of segments from this tissue. We show that *Sox21b-1* acts
 329 upstream of Wnt8 and Delta-Notch signalling in this GRN and is required for
 330 the activation of these important signalling pathways during posterior
 331 development (Fig. 6). Further work is needed to determine if Group B Sox
 332 genes, such as *Dichaete* and *Sox21b-1*, occupy a similar position in the
 333 regulatory hierarchy for posterior segmentation in other arthropods. This could
 334 provide important new insights into the evolution of the regulation of
 335 segmentation in arthropods since a Wnt-Delta-Notch-Cad regulatory cassette
 336 was probably used ancestrally in arthropods to regulate posterior
 337 development (4, 8, 9, 14, 15, 42). Interestingly, SoxB genes also cooperate
 338 with Wnt and Delta-Notch signalling in various aspects of vertebrate
 339 development including the patterning of neural progenitors and maintenance
 340 of the stem state in the neuroepithelium (35, 43, 44, 45).

341

342 *Sox21b-1* exhibits highly pleiotropic phenotypes during early spider
343 embryogenesis

344 Our study shows that *Sox21b-1* is not only involved in segmentation but is
345 maternally supplied and regulates cell division in the early germ disc, as well
346 as the transition from radial to axial symmetry during germ band formation.
347 Further experiments with *Sox21b-1* are required to fully elucidate the
348 mechanisms by which it affects these early functions. Furthermore, while
349 spider head development is less affected than trunk segmentation by
350 knockdown of *Sox21b-1*, it is clear from our experiment that *Sox21b-1*
351 regulates cell fate in this region. Interestingly, Sox2 is involved with the neuro-
352 mesodermal fate choice in mice and *Dichaete* has a role in embryonic brain
353 development in *Drosophila* (45, 46): consequently, SoxB genes may play an
354 ancestral role in the patterning of the head ectoderm and mesoderm in
355 metazoans (45, 46).

356

357 *The evolution of Sox21b-1*

358 The evolution and diversification of Group B Sox genes in insects is not fully
359 resolved due to difficulties in clearly assigning orthologues based on the
360 highly conserved HMG domain sequence (17, 35, 47). However, despite
361 these ambiguities it is clear that the *Dichaete* and *Sox21b* class genes in all
362 arthropods examined to date are closely related and likely arose from a
363 duplication in the common ancestor of this phylum (see 47 for discussion).
364 Note that in insects *Dichaete*, *Sox21a* and *Sox21b* are clustered (29),
365 however, while *Dichaete* and *Sox21a* are also clustered in *P. tepidariorum*,
366 the *Sox21b* paralogs are dispersed in the genome of this spider (21). We

367 believe it is highly significant that two very closely related SoxB genes are
 368 involved in segmentation in both the spider *P. tepidariorum* and in insects,
 369 pointing to an ancient role for this subfamily of Sox genes in invertebrates.
 370 Given the close similarity between the HMG domains of Sox21b and
 371 Dichaete, it is possible that in some lineages the Dichaete orthologue
 372 assumed the segmentation role, whereas in others it was Sox21b. In spiders,
 373 *Wnt8* is involved in posterior development while in other arthropods this role is
 374 played by *Wnt1/wg* (14), and therefore the evolution of *Sox21b-1* may have
 375 led to the co-option to different genes and developmental systems drift of the
 376 GRN for posterior development.

377 The spider contains an additional related SoxB gene, *Sox21b-2*, that
 378 possibly arose as part of the whole genome duplication event in the ancestor
 379 of arachnoplumonates over 400 million years ago (22). It will be interesting to
 380 examine any segmentation roles in other spiders and arachnids, including
 381 those that did not undergo a genome duplication. Finally, Blast searches of
 382 the Tardigrade *Hypsibius dujardini* genome reveal a single *Dichaete/Sox21b*
 383 class gene and it will be of some interest to characterise the expression
 384 and/or function of this gene in this sister group to the arthropods.

385

386

387 **Materials and Methods**

388 **Spider Culture**

389 *P. tepidariorum* were cultured at 25°C at Oxford Brookes University. The
390 spiders were fed with *D. melanogaster* with vestigial wings and subsequently
391 small crickets (*Gryllus bimaculatus*). Cocoons from mated females were
392 removed and a small number of embryos were immersed in halocarbon oil for
393 staging according to (26).

394

395 **Phylogenetic analysis of *P. tepidariorum* Sox genes**

396 To identify the phylogenetic relationship of *P. tepidariorum* Sox genes the
397 HMG domains of *Anopheles gambiae*, *Mus musculus*, *D. melanogaster*, *P.*
398 *tepidariorum* and *S. mimosarum* Sox genes were aligned with ClustalW (21,
399 48). Phylogenetic analysis was performed in RAxML, with support levels
400 estimated implementing the rapid bootstrap algorithm (1000 replicates) (49),
401 under the PROTGAMMALG model of amino acid substitution, which was
402 identified as best fitting using a custom Perl script from the Exelixis Lab
403 website (https://sco.h-its.org/exelixis/web/software/raxml/hands_on.html).

404

405 **Fixation and gene expression analysis**

406 Embryos ranging from the 1-cell stage to stage 13 were dechorionated and
407 fixed according to (31) with a longer fixation time of 1 hr to facilitate yolk
408 removal for flat-mounting. For immunohistochemistry, methanol steps were
409 omitted. Ovaries from adult females were dissected in 1x PBS and fixed in 4%
410 formaldehyde for 30 min. Probe synthesis and RNA *in situ* hybridisation was
411 carried out with minor modifications to (27), omitting Proteinase K treatment

412 and post-fixation steps. Poly-L-lysine (Sigma-Aldrich) coated coverslips were
413 used for flat-mounting embryos. Nuclei were stained by incubating embryos in
414 1 µg/ml 4-6-diamidino-2-phenylindol (DAPI) in PBS with 0.1% Tween-20 for
415 15 min.

416

417 **Imaging, Live Imaging and Image Analysis**

418 For imaging of flat-mounted embryos after *in situ* hybridisation, an AxioZoom
419 V16 stereomicroscope (Zeiss) equipped with an Axiocam 506-Mono and a
420 colour digital camera were used. Immunostained embryos were imaged with
421 Zeiss LSM 800 or 880 with Airyscan confocal microscopes. For live imaging,
422 embryos were aligned on heptane glue coated coverslips and submersed in a
423 thin layer of halocarbon oil. Bright-field live imaging was performed using an
424 AxioZoom V16 stereomicroscope, while fluorescence live imaging was
425 performed with confocal microscopes. Image stacks were processed in Fiji
426 (50) and Helicon Focus (HeliconSoft). Image brightness and intensity was
427 adjusted in Corel PhotoPaint X5 (CorelDraw) and Fiji.

428

429 **Gene Isolation from cDNA**

430 Fragment of genes were amplified using PCR and cloned into pCR4-TOPO
431 (Invitrogen, Life Technologies). Oligonucleotide sequences are listed in Table
432 S1.

433

434 **Immunohistochemistry**

435 Immunostaining was carried out following (51) with minor modifications:
436 antibodies were not pre-absorbed prior to incubation and the concentration of

437 Triton was increased to 0.1%. The following primary antibodies were used:
 438 mouse anti- α -Tubulin DM1a (Sigma) (1:50), rabbit α cleaved caspase 3 (Cell
 439 Signaling - 9661) (1:200) and rabbit Anti-phospho-Histone H3 (Ser10) (Merck
 440 Millipore - 06-570). For detection the following secondary antibodies were
 441 used: donkey anti-mouse IgG Alexa Fluor 555 (Invitrogen) and goat anti-rabbit
 442 Alexa Fluor 647 (Invitrogen). The counterstaining was carried out by
 443 incubation in 1 μ g/ml 4-6-diamidino-2-phenylindol (DAPI) in PBS + Triton
 444 0,1% for 20 minutes.

445

446 **dsRNA synthesis and Parental RNA interference**

447 Double stranded RNA (dsRNA) for parental RNA interference was
 448 synthesized according to (15) and injected following the standard protocol
 449 from (27). Two non-overlapping fragments of *P. tepidariorum Sox21b-1* were
 450 isolated from the 1134 bp coding sequence of the gene: fragment 1 spanning
 451 549 bp and fragment 2 covering 550 bp. Double stranded RNA for *P.*
 452 *tepidariorum D1* (853 bp), *Wnt8* (714 bp) and the coding sequence of GFP
 453 (720 bp) as used previously (27), were transcribed using the same method.
 454 Synthesis of double stranded RNA was performed using the MegaScript T7
 455 transcription kit (Invitrogen). After purification the dsRNA transcripts were
 456 annealed in a water bath starting at 95°C and slowly cooled down to room
 457 temperature. dsRNA was injected at 2.0 μ g/ μ l in the opisthosoma of adult
 458 females every two days, with a total of five injections (n = 7 for each dsRNA;
 459 n= 2 for GFP controls). The injected spiders were mated after the second
 460 injection and embryos from injected spiders were fixed for gene expression
 461 and phenotypic analyses at three different time points: stage 4 (cumulus

462 formation), stage 5 late (germ disc with migrating cumulus) and stage 9 (head
463 and limbs bud formation).

464

465 **Supplemental Information**

466 Supplemental Information includes five figures and two tables.

467

468 **Declaration of Interest**

469 The authors declare no competing interests.

470

471 **Author Contributions**

472 CLBP, SR and APM designed the project for the paper. CLBP and AS
473 performed most of the experiments. DJL and SR carried out the genomic and
474 bioinformatic analysis. CLBP, SR and APM wrote the manuscript with the help
475 of DJL and AS.

476

477 **Acknowledgements**

478 This research was funded by a CNPq scholarship to CLBP (234586/2014-1),
479 a grant from The Leverhulme Trust (RPG-2016-234) to APM and AS, and in
480 part by a BBSRC grant (BB/N007069/1) to SR.

481

482 **References**

- 483 1. Tautz D (2004) Segmentation. *Dev Cell* 7:301-312.
- 484 2. Peel AD, Chipman AD, & Akam M (2005) Arthropod segmentation:
485 beyond the *Drosophila* paradigm. *Nat Rev Genet* 6:905-916.
- 486 3. Davis GK & Patel NH (2002) Short, long, and beyond: molecular and
487 embryological approaches to insect segmentation. *Annu Rev Entomol*
488 47:669-699.

- 489 4. McGregor AP, Pechmann M, Schwager EE, Damen WG (2009) An
490 ancestral regulatory network for posterior development in arthropods.
491 *Commun Integr Biol* 2:174-176.
- 492 5. Hilbrant M, Damen WG, & McGregor AP (2012) Evolutionary
493 crossroads in developmental biology: the spider *Parasteatoda*
494 *tepidariorum*. *Development* 139:2655-2662.
- 495 6. McGregor AP, et al. (2008) *Cupiennius salei* and *Achaearenea*
496 *tepidariorum*: Spider models for investigating evolution and
497 development. *Bioessays* 30:487-498.
- 498 7. Leite DJ & McGregor AP (2016) Arthropod evolution and development:
499 recent insights from chelicerates and myriapods. *Curr Opin Genet Dev*
500 39:93-100.
- 501 8. Janssen R, Prpic NM, Damen WGM (2004) Gene expression suggests
502 decoupled dorsal and ventral segmentation in the millipede *Glomeris*
503 *marginata* (Myriapoda: Diplopoda). *Dev Biol* 269: 84-104
- 504 9. Brena C, Akam M (2012) The embryonic development of the centipede
505 *Strigamia maritima*. *Dev Biol* 363: 290-307
- 506 10. Kanayama M, et al. (2011) Travelling and splitting of a wave of
507 hedgehog expression involved in spider-head segmentation. *Nat*
508 *Commun* 2:500.
- 509 11. Pechmann M, McGregor AP, Schwager EE, Feitosa NM, Damen WG
510 (2009) Dynamic gene expression is required for anterior regionalization
511 in a spider. *Proc Natl Acad Sci U S A* 106:1468-1472.
- 512 12. Pechmann M, et al. (2011) Novel function of *Distal-less* as a gap gene
513 during spider segmentation. *PLoS Genet* 7:e1002342.
- 514 13. Schwager EE, Pechmann M, Feitosa NM, McGregor AP, Damen WG
515 (2009) *hunchback* functions as a segmentation gene in the spider
516 *Achaearenea tepidariorum*. *Curr Biol* 19:1333-1340.
- 517 14. McGregor AP, et al. (2008) *Wnt8* is required for growth-zone
518 establishment and development of opisthosomal segments in a spider.
519 *Curr Biol* 18:1619-1623.
- 520 15. Schönauer A, et al. (2016) The Wnt and Delta-Notch signalling
521 pathways interact to direct even-skipped expression via caudal during

- 522 segment addition in the spider *Parasteatoda tepidariorum*.
523 *Development* 143:2455-2463.
- 524 16. Nambu PA, JR Nambu (1996) The *Drosophila* fish-hook gene encodes
525 a HMG domain protein essential for segmentation and CNS
526 development. *Development* 122: 3467–3475.
- 527 17. Russell SRH, Sanchez-Soriano N, Wright CR, Ashburner M (1996) The
528 *Dichaete* gene of *Drosophila melanogaster* encodes a SOX-domain
529 protein required for embryonic segmentation. *Development*, 122:3669-
530 3676.
- 531 18. Clark E & Peel AD (2017) Evidence for the temporal regulation of
532 insect segmentation by a conserved set of developmental transcription
533 factors. *bioRxiv*.
- 534 19. Clark E (2017) Dynamic patterning by the *Drosophila* pair-rule network
535 reconciles long-germ and short-germ segmentation. *PLoS Biol*
536 15:e2002439.
- 537 20. Verd B, et al. (2018) A damped oscillator imposes temporal order on
538 posterior gap gene expression in *Drosophila*. *PLoS Biol* 16:e2003174.
- 539 21. Paese CLB, Leite DJ, Schoenauer A, McGregor AP, & Russell S
540 (2017) Duplication and divergence of Sox genes in spiders. *bioRxiv*.
- 541 22. McKimmie C, Woerfel G, Russell S (2005) Conserved genomic
542 organisation of group B Sox genes in insects. *BMC Genetics*, 6:26.21-
543 26.15.
- 544 23. Wilson MJ, Dearden PK (2008) Evolution of the insect Sox genes.
545 *BMC Evol Biol*, 8:120.
- 546 24. Schwager EE, et al. (2017) The house spider genome reveals an
547 ancient whole-genome duplication during arachnid evolution. *BMC Biol*
548 15(1):62.
- 549 25. Jędrzejowska I & Kubrakiewicz IJ (2007) The Balbiani body in the
550 oocytes of a common cellar spider, *Pholcus phalangioides* (Araneae:
551 Pholcidae). *Arthropod Struct Dev*, 36: 317-326.
- 552 26. Mittmann B, Wolff C (2012) Embryonic development and staging of the
553 cobweb spider *Parasteatoda tepidariorum* C. L. Koch, 1841 (syn.:
554 *Achaeearanea tepidariorum*; *Araneomorphae*; *Theridiidae*). *Dev Genes*
555 *Evol*, 222:189-216.

- 556 27. Akiyama-Oda Y, Oda H (2003) Early patterning of the spider embryo: a
557 cluster of mesenchymal cells at the cumulus produces Dpp signals
558 received by germ disc epithelial cells. *Development*, 130:1735-1747.
- 559 28. Pechmann M (2016) Formation of the germ-disc in spider embryos by
560 a condensation-like mechanism. *Frontiers in Zoology* 13:35
- 561 29. Pechmann M, et al. (2017) A novel role for Ets4 in axis specification
562 and cell migration in the spider *Parasteatoda tepidariorum*. *eLife*, 6: 1-
563 20.
- 564 30. Akiyama-Oda Y, Oda H (2006) Axis specification in the spider embryo:
565 dpp is required for radial-to-axial symmetry transformation and sog for
566 ventral patterning. *Development*, 133: 2347-2357.
- 567 31. Akiyama-Oda Y, Oda H (2016) Multi-color FISH facilitates analysis of
568 cell-type diversification and developmental gene regulation in the
569 *Parasteatoda* spider embryo. *Dev Growth Differ*, 58:215-224.
- 570 32. Yamazaki K, Akiyama-Oda Y, Oda H (2005) Expression Patterns of a
571 twist-Related Gene in Embryos of the Spider *Achaearanea*
572 *tepidariorum* Reveal Divergent Aspects of Mesoderm Development in
573 the Fly and Spider. *Zool Sci* 22: 177-185.
- 574 33. Oda H, et al. (2007) Progressive activation of Delta-Notch signaling
575 from around the blastopore is required to set up a functional caudal
576 lobe in the spider *Achaearanea tepidariorum*. *Development*, 134: 2195-
577 2205.
- 578 34. Overton PM, Meadows LA, Urban J and Russell S (2002) Evidence for
579 differential and redundant function of the Sox genes *Dichaete* and
580 *SoxN* during CNS development in *Drosophila*. *Development* 129: 4219-
581 4228.
- 582 35. Wegner, M (1999) From head to toes: the multiple facets of Sox
583 proteins. *Nucleic Acids Research* 27:1409–1420.
- 584 36. Sinclair AH et al (1990) A gene from the human sex-determining region
585 encodes a protein with homology to a conserved DNA-binding motif.
586 *Nature* 346: 240-244.
- 587 37. Lefebvre V (2010) The SoxD transcription factors – Sox5, Sox6, and
588 Sox13 – are key cell fate modulators. *Int J Biochem Cell Biol* 42:429–
589 432.

- 590 38. Nakao H (2018) A Bombyx homolog of ovo is a segmentation gene that
591 acts downstream of Bm-wnt1(Bombyx wnt1 homolog). *Gene*
592 *Expression Patterns* 27: 1-7.
- 593 39. Damen WGM, Weller M, Tautz D (2000) Expression patterns of hairy,
594 even-skipped, and runt in the spider Cupiennius salei imply that these
595 genes were segmentation genes in a basal arthropod. *PNAS*, 97:
596 4515-4519.
- 597 40. Feitosa NM, Pechmann M, Schwager EE, et al (2017) Molecular
598 control of gut formation in the spider Parasteatoda tepidariorum.
599 *genesis* 55: 23-33.
- 600 41. Lan, Q., Cao, M., Kollipara, R. K., Rosa, J. B., Kittler, R. and Jiang, H
601 (2017) FoxA transcription factor Fork head maintains the intestinal
602 stem/progenitor cell identities in Drosophila. *Dev Biol* 433:324-343.
- 603 42. Pueyo JI, Lanfear R, Couso JP (2008) Ancestral Notch-mediated
604 segmentation revealed in the cockroach Periplaneta americana.
605 *Proceedings of the National Academy of Sciences*, 105: 16614-16619.
- 606 43. Holmberg J, et al. (2008) SoxB1 transcription factors and Notch
607 signaling use distinct mechanisms to regulate proneural gene function
608 and neural progenitor differentiation. *Development* 135: 1843-1851.
- 609 44. Kormish JD, Sinner D, Zorn, AM (2010) Interactions between SOX
610 factors and Wnt/ β - catenin signaling in development and disease.
611 *Dev Dyn* 239: 56-68.
- 612 45. Koch F, et al (2017) Antagonistic Activities of Sox2 and Brachyury
613 Control the Fate Choice of Neuro-Mesodermal Progenitors. *Dev Cell*
614 42: 514-526.
- 615 46. Sanchez-Soriano N, Russell S (1998) The Drosophila Sox-domain
616 protein Dichaete is required for the development of the central nervous
617 system midline. *Development* 125:3989-3996.
- 618 47. Zhong L, Wang D, Gan X, Yang T, He S (2011) Parallel Expansions of
619 Sox Transcription Factor Group B Predating the Diversifications of the
620 Arthropods and Jawed Vertebrates. *PLoS ONE*, 6:e16570.
- 621 48. Larkin MA, et al. (2007) Clustal W and Clustal X version 2.0.
622 *Bioinformatics* 23: 2947-2948.

- 623 49. Stamatakis, et al. (2008) A rapid bootstrap algorithm for the RAxML
624 Web servers. *Syst Biol* 57: 758-771.
- 625 50. Schindelin J, et al. (2012) Fiji: an open-source platform for biological-
626 image analysis. *Nature Methods* 9: 676–682.
- 627 51. Schwager EE, Meng Y, Extavour CG (2015) vasa and piwi are required
628 for mitotic integrity in early embryogenesis in the spider *Parasteatoda*
629 *tepidariorum*. *Dev Biol* 402: 276-290.

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

Figures

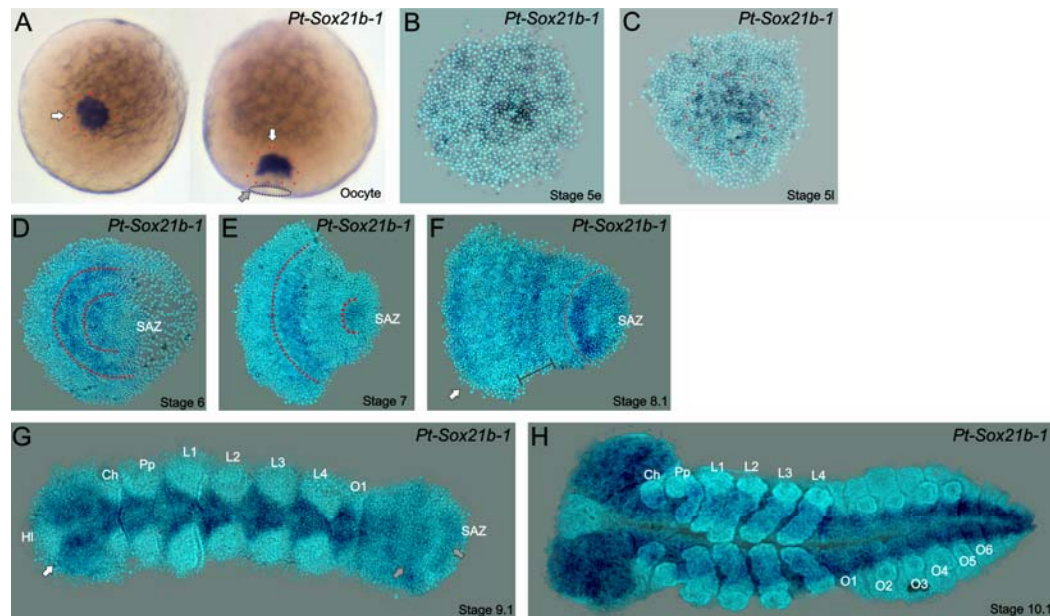
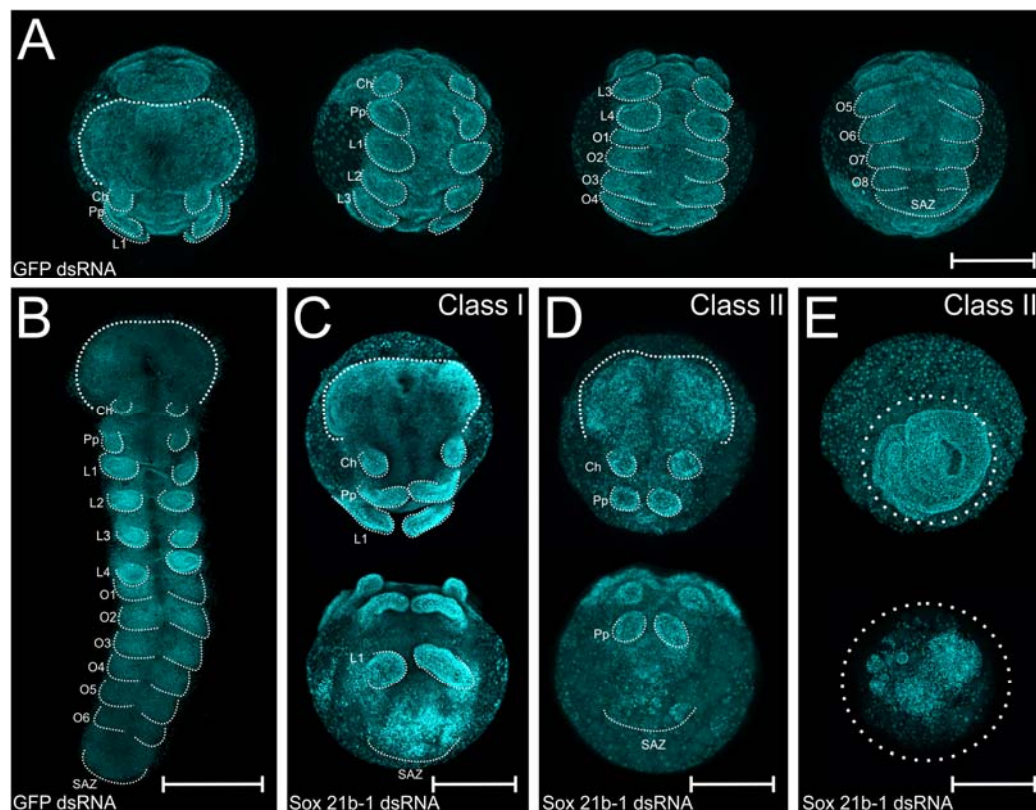


Figure 1. Expression of *Sox21b-1* in *P. tepidariorum* oocytes and embryos

A) Dorsal (left) and lateral (right) views of pre-vitellogenic oocytes showing *Sox21b-1* mRNA in the Balbiani's body (red dashed circle and white arrows). The sperm implantation groove is indicated by a black dashed circle and grey arrow. **B)** At early stage 5, the expression of *Sox21b-1* appears in a salt and pepper pattern in the germ disc. **C)** Expression in the cumulus becomes stronger at late stage 5, with less expression at the periphery of the germ disc (dashed red circle). **D)** At stage 6 *Sox21b-1* is expressed in a broad stripe in the anterior (between the red dashed lines). **E)** At stage 7 there is expression in the region of the presumptive leg bearing segments and in the SAZ (both indicated by red dashed lines). **F)** At stage 8.1 *Sox21b-1* is still expressed in the SAZ and the presumptive leg bearing segments, but nascent expression is observed at the anterior of the germ band (indicated by the white arrows and black brackets). **G)** At stage 9.1, when the limb buds are visible the expression of *Sox21b-1* becomes restricted to the ventral nerve cord (anterior white arrow) and can be observed in the SAZ (posterior grey arrows). **H)** At stage 10.1, *Sox21b-1* expression is restricted to the ventral nerve cord and the head lobes. Ch: Chelicerae; HL: Head lobes; L1 to L4: Prosomal leg

679 bearing segments 1 to 4; O1 to O6: Opisthosomal segments 1 to 6; SAZ:
680 Segment addition zone. Ventral views are shown with anterior to the left,
681 except as described for oocytes.
682

683



684

685 **Figure 2. Embryo phenotypes after *Sox21b-1* parental RNAi knockdown**
686 Whole mount (A) and flat mount (B) control embryos at stage 9 stained with
687 DAPI. Stage 9, Class I (C), Class II (D) and Class III (E) phenotypes from
688 *Sox21b-1* knockdown. In the control embryos (A and B), the head, cheliceral
689 (Ch), pedipalpal (Pp), prosomal walking limbs (L1 to L4), opisthosomal
690 segments (O1 to O6) and a posterior SAZ are all clearly visible as indicated.
691 C) Class I phenotype embryos show a morphologically normal head, pairs of
692 chelicerae, pedipalps and first walking limbs (Ch, Pp, L1), but a disorganised
693 cluster of cells in the posterior where L2-L4, opisthosomal segments and the
694 SAZ should be. D) Class II phenotype embryos consist of fewer cells, but still
695 form a head, chelicerae, pedipalps (Ch, Pp) and a structure resembling the
696 SAZ in the posterior. E) Class III embryos exhibit the most severe phenotype,
697 where, after the germ disc stage, the embryo fails to form an organised germ
698 band. Anterior is to the top, scale bars: 150 µm.

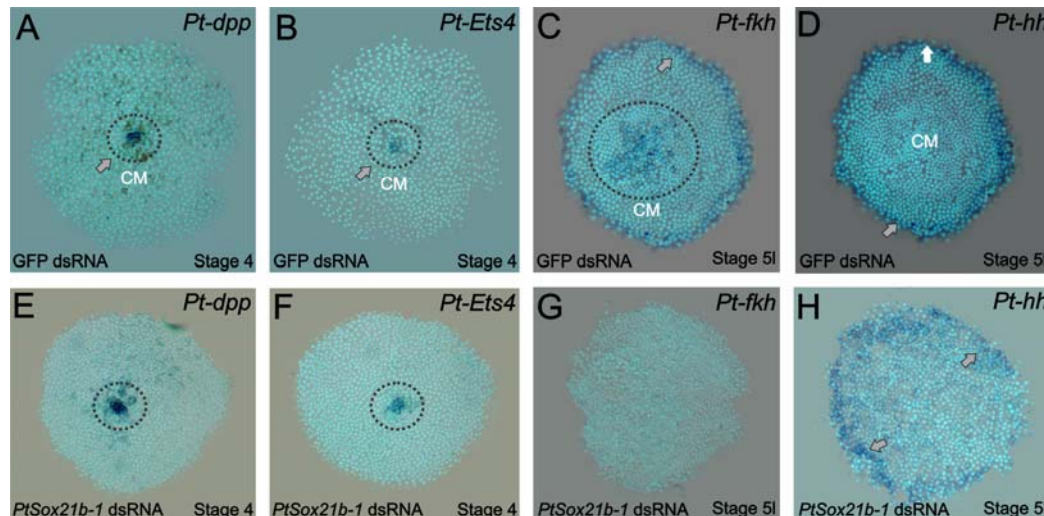


Figure 3. Gene expression in control and *Sox21b-1* knockdowns at the germ disc stage

Pt-dpp (A) and *Pt-Ets4* (B) are expressed in the forming cumulus (CM) in the centre of the germ disc at stage 4 (grey arrow and dotted circle). This expression is unaffected by knockdown of *Sox21b-1* (E and F) (n= 30 for each gene). C) *Pt-fkh* is expressed at the rim and centre of the germ disc at late stage 5 (grey arrow and dotted circle in C), but expression is lost in *Sox21b-1* embryos (n= 30) (G). *Pt-hh* expression at the rim of the germ disc (D) is normal in *Sox21b-1* knockdown embryos (H) (grey arrows).

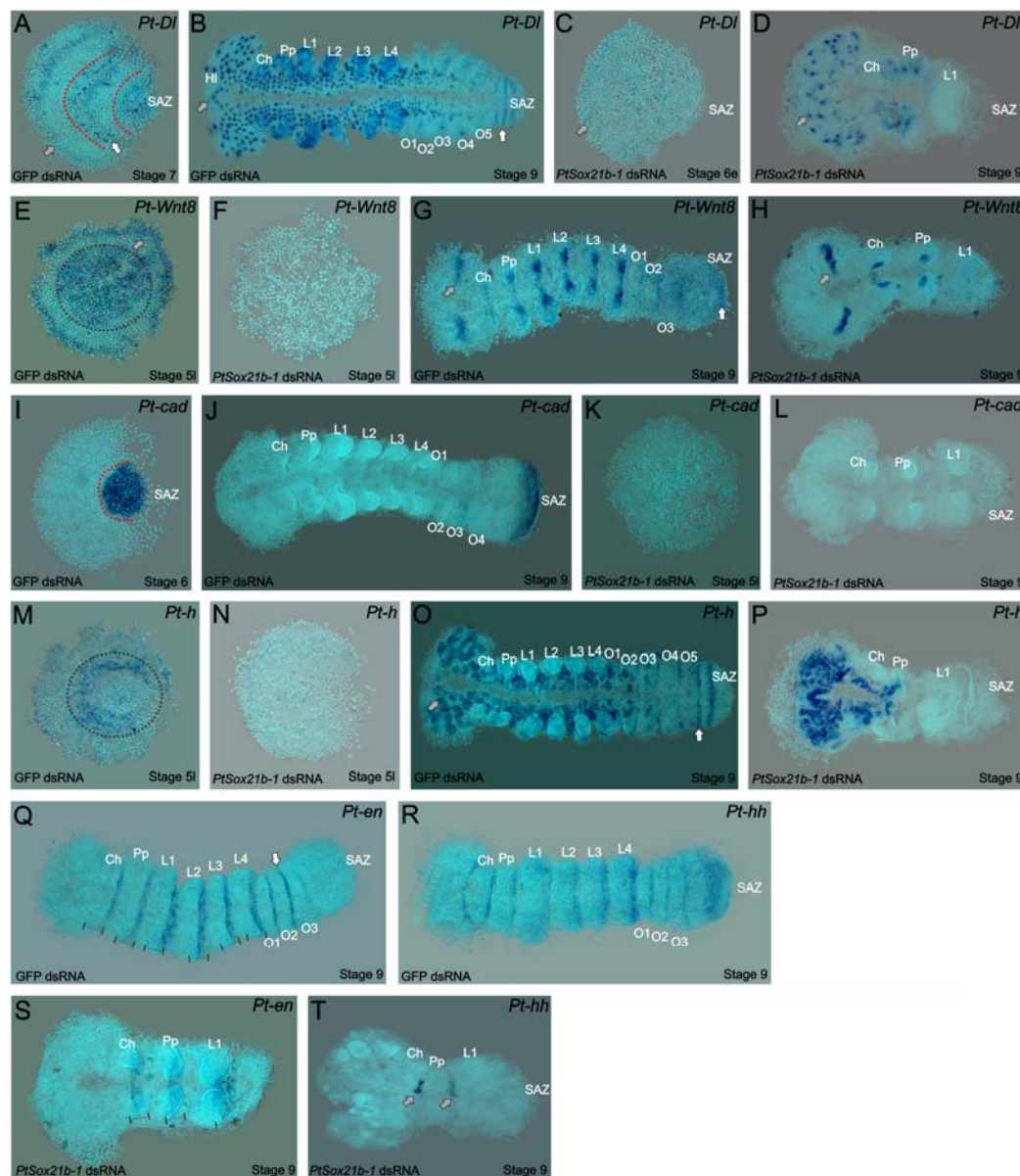


Figure 4. Expression of segmentation genes in *Sox21b-1* pRNAi embryos

A and B) *Pt-Dl* expression at late stage 6/early stage 7 is dynamic in the SAZ and is also observed in the presumptive head region and prosoma of the embryo (red dotted lines and grey arrows). **B)** At stage 9, *Pt-Dl* expression is seen in the SAZ (white arrow) but is restricted to the clusters of proneural differentiation in the anterior region of the embryo (grey arrow in the head lobes). **C)** In *Sox21b-1* knockdown embryos, *Pt-Dl* expression is not detectable in late stage 5/early stage 6 embryos (grey arrow) but can still be observed in the anterior ventral neuroectoderm at stage 9 up to the pedipalpal

segment (n= 17 and n = 14 for stage 5 and 9, respectively) (**D**). *Pt-h* expression at stage 5 in control embryos is seen at the rim and in the centre of the germ disc (black dotted circle in **E**), which is lost in *Sox21b-1* knockdown embryos (**F**). At stage 9, *Pt-h* expression resembles *Pt-Dl*, both in the control and *Sox21b-1* knockdown embryos (**G** and **H**) (n= 15 for both stages). *Pt-Wnt8* expression is similar to *Pt-h* in stage 5 control embryos (black dotted circle in the centre, grey arrow to the rim) and is also lost in *Sox21b-1* knockdown embryos (n= 11) (**I** and **J**). Control embryos at stage 9 show the expression of *Pt-Wnt8* in the medial region of the head (grey arrow), and in distal parts of each segment up to the SAZ (white arrow) (**K**). In *Sox21b-1* knockdown embryos at the same stage, the brain (grey arrow), cheliceral and pedipalpal expression is still present, but the posterior expression is lost (n= 17 for each stage) (**L**). *Pt-cad* is expressed in the SAZ at late stage 5/early stage 6 embryos (**M**), which persists throughout to stage 9 control embryos (**N**). However, *Pt-cad* expression is lost upon *Sox21b-1* knockdown (n= 20 for each stage) (**O** and **P**). *Pt-en* expression is present in the posterior of each segment (black lines in **Q**), and in cheliceral, pedipalpal and L1 segments in *Sox21b-1* knockdown embryos at stage 9 (n= 10) (**S**). *Pt-hh* expression in control embryos at stage 9 is seen in the posterior of each segment and in the SAZ (**R**). When *Sox21b-1* is knocked-down, *Pt-hh* embryos show expression in the middle posterior of the cheliceral and pedipalpal segments (n= 8) (**T**). Ch: Chelicerae; HL: Head Lobes; L1 to L4: Prosomal leg bearing segments; O1 to O5: Opisthosomal segments; SAZ: Segment Addition Zone. Anterior is to the left in stage 9 embryos.

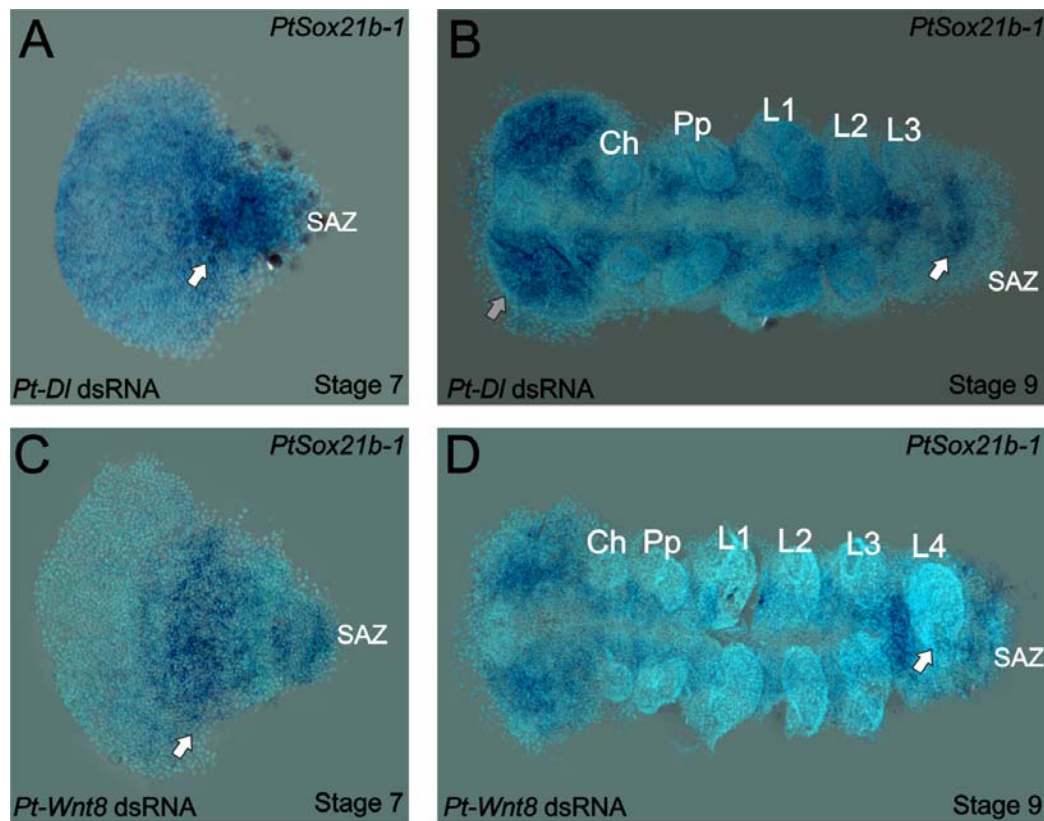


Figure 5. Expression of *Sox21b-1* in *Df* and *Wnt8* pRNAi embryos

Ventral view of stage 7 and 9 knockdown embryos for *Pt-Df* (A and B) and *Pt-Wnt8* (C and D). In knockdown embryos for both *Pt-Df* and *Pt-Wnt8*, *Sox21b-1* is still expressed at the germ band stage (A and C), in a dynamic pattern in the remaining SAZ cells, and in the forming segments in the presumptive prosoma of the embryo (white arrows). In stage 9 *Pt-Wnt8* knockdown embryos, *Sox21b-1* remains highly expressed in the ventral nerve cord (D). *Pt-Df* knockdown embryos lack the posterior L4 segment (white arrow), but brain formation appears normal (grey arrow) (B). *Pt-Wnt8* embryos show a fusion of the L4 limb buds, and *Sox21b-1* is still expressed in the remaining SAZ cells (D). Anterior is to the left in all panels.

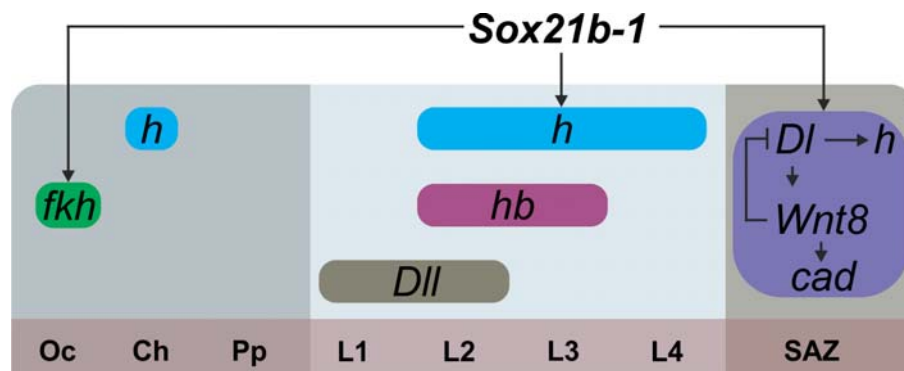
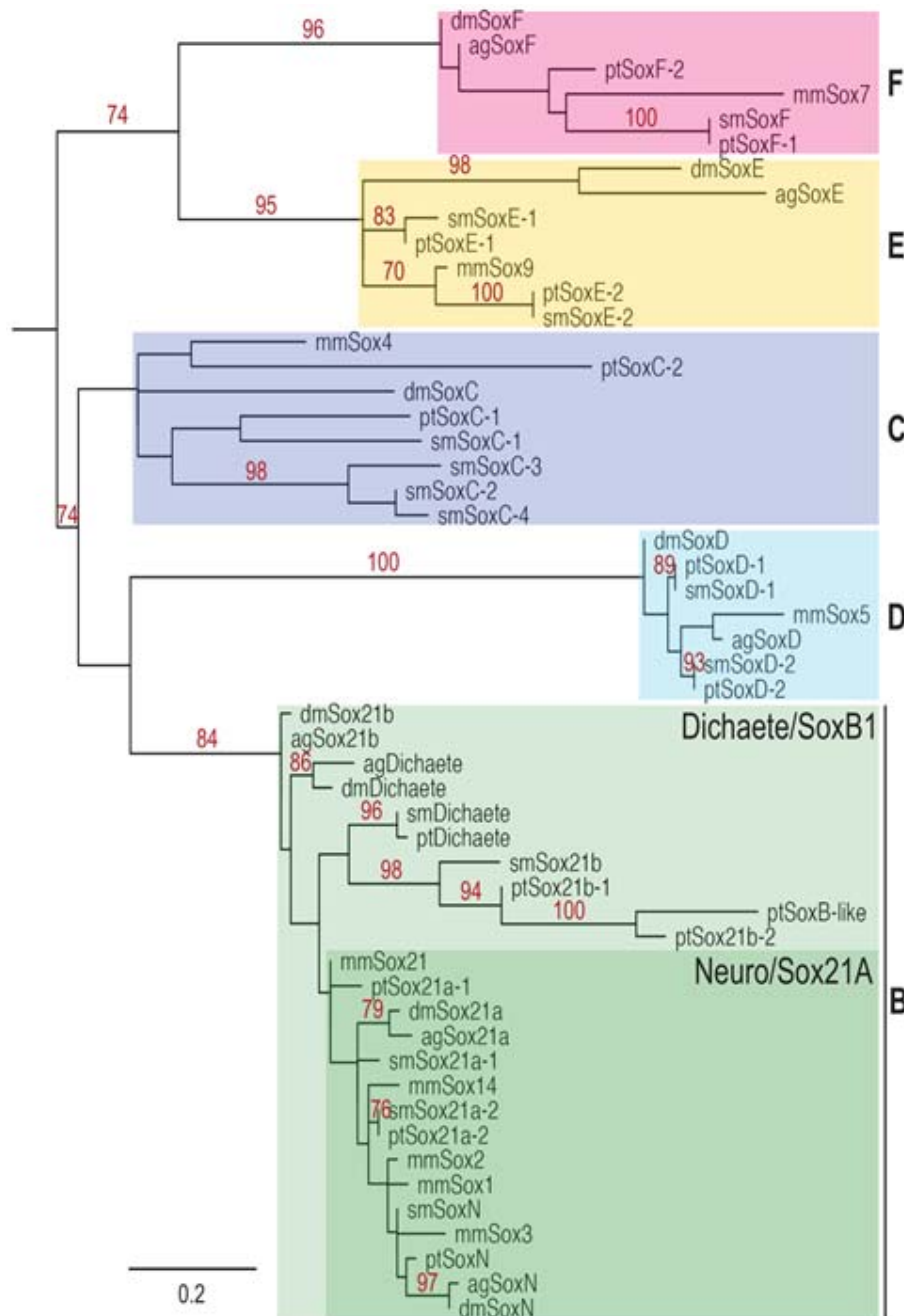


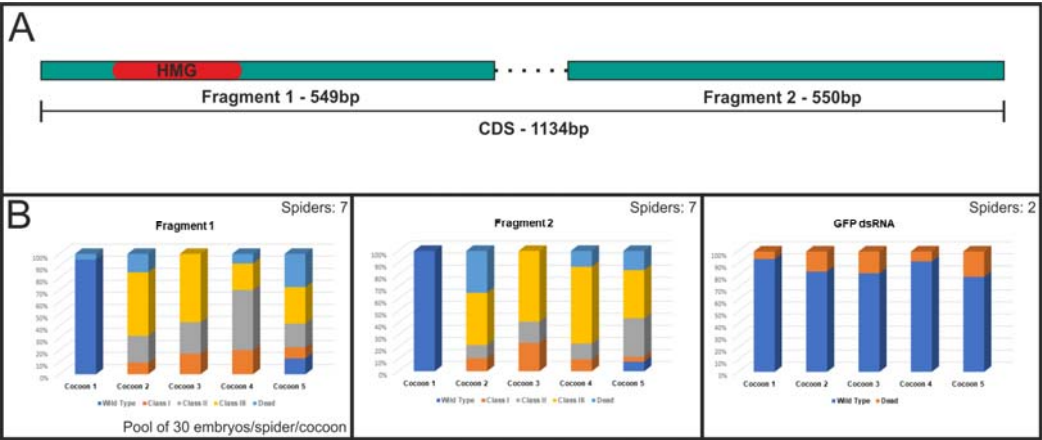
Figure 6. Summary of the regulation of spider segmentation

The interaction of *Sox21b-1* is presented in relation to genes involved in spider embryogenesis. We found that *fkf* expression requires *Sox21b-1* in the most anterior part of the head (OC, Ch, Pp segments). In the prosoma *Sox21b-1* has a gap gene like function and is required for the expression of *hairy*, while *Distal-less* (12) and *hunchback* (11) also act as gap genes during prosomal segmentation (L1-L4). The molecular control of segmentation in the SAZ involves a feedback loop between *Dl* and *Wnt8*, which acts upstream of *cad* and also controls the dynamic expression of *hairy* (15). We can infer from our results that *Sox21b-1* acts upstream of these genes in the SAZ.



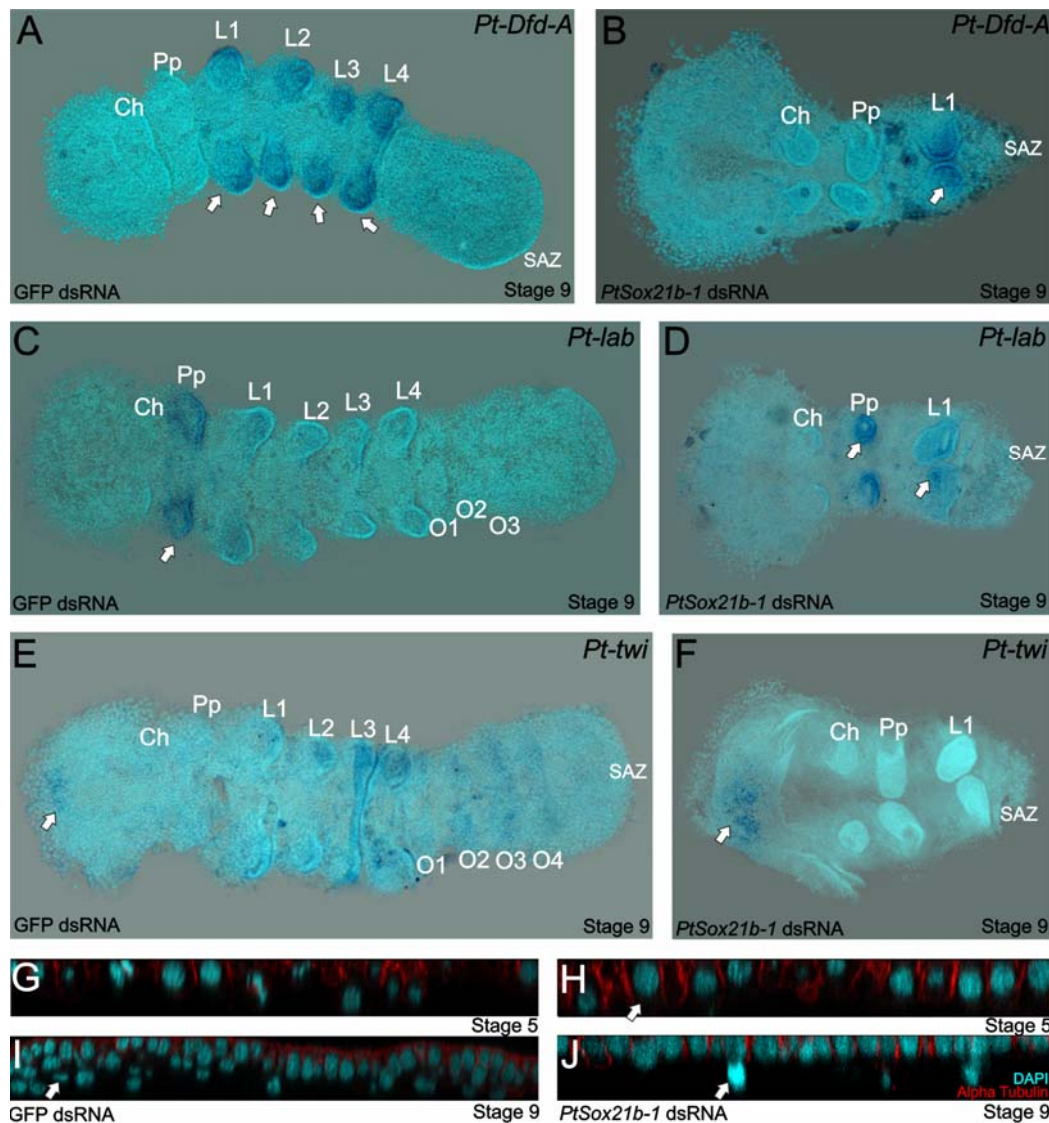
Supplementary Figure 1. RAxML phylogeny of eumetazoan Sox genes.

Phylogenetic tree made with RAxML algorithm showing the relationship between *Anopheles gambiae* (Ag), *Mus musculus* (Mm), *D. melanogaster* (Dm), *P. tepidariorum* (Pt) and *S. mimosarum* (Sm) Sox proteins based on HMG domain sequences.



Supplementary Figure 2. dsRNA design and phenotypical class frequencies for each fragment and GFP control injections

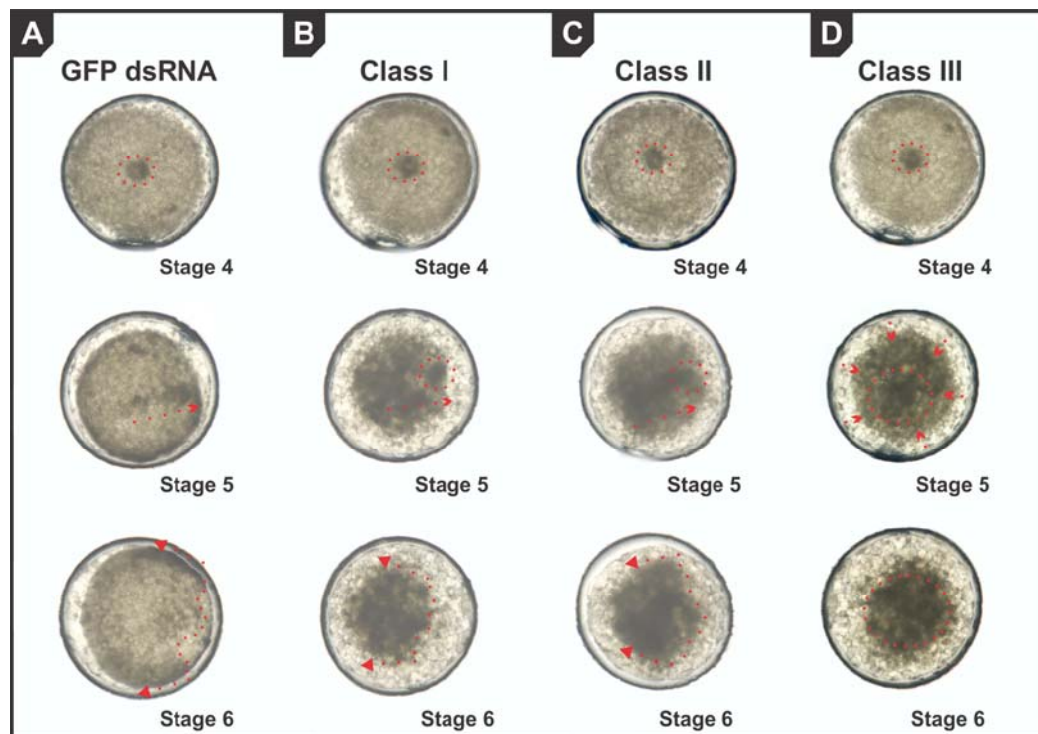
A) Two non-overlapping fragments were designed for the *Sox21b-1* coding sequence. Fragment 1 contains the HMG conserved domain (549 bp) and fragment 2 has no conserved domains (550 bp). **B)** Frequencies for each fragment, cocoon number and phenotypic class. Seven spiders were injected for each *Pt-Sox21b-1* fragment and two spiders for the GFP dsRNA controls. For the phenotypic class frequencies, 30 embryos per spider per cocoon were pooled, DAPI stained and analysed (total n= 210 for each).



Supplementary Figure 3. Homeotic and mesodermal gene expression at stage 9 in *Sox21b-1* pRNAi embryos

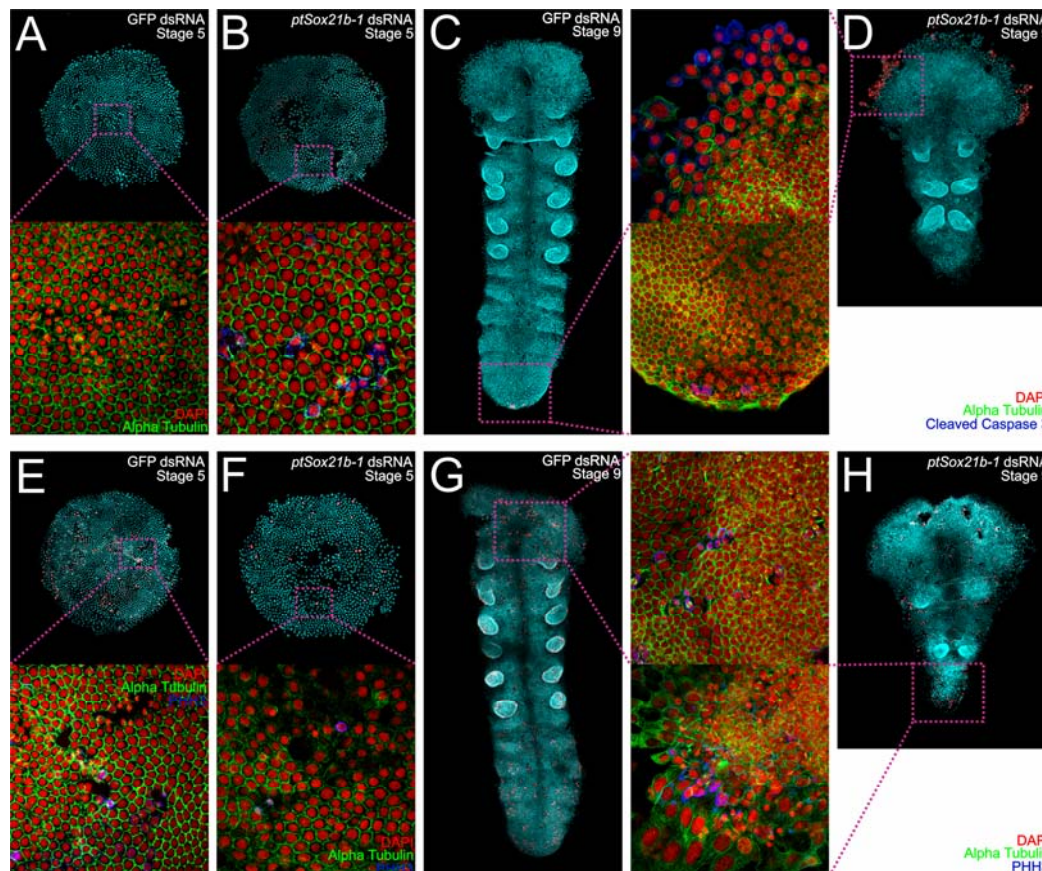
A) Ventral view showing *Pt-Dfd-A* expression in the limb buds of L1 to L4 segments in the control embryos (white arrows). **B)** Expression of *Pt-Dfd-A* is also observed in L1 in *Sox21b-1* pRNAi embryos (n = 9) (white arrow in **B**). **C)** *Pt-lab* is expressed in the pedipalpal segment and faintly in L1 segment in control embryos (white arrow in **D**). In *Sox21b-1* pRNAi embryos, *Pt-lab* expression can still be observed in the pedipalpal and L1 segments (n = 10) (white arrows in **D**). **E)** The mesodermal marker *Pt-twi* is expressed in the anterior-most medial region of the head, limb buds of L1 to L4, and with a striped pattern in the O1 to O4 segments. **F)** In *Sox21b-1* knockdown embryos, only the head expression is maintained (n= 14) (white arrow in **F**).

800 **G-J** show orthogonal projections of the cumulus (stage 5) and the head (stage
 801 9) at 40x magnification of whole mount control embryos (left panels) and
 802 *Sox21b-1* knockdown embryos (right panels), respectively. In control embryos
 803 the formation of subectodermal layers are visible, which are lost in the
 804 knockdown embryos. DAPI stained nuclei are shown in cyan and the
 805 membrane marker alpha-Tubulin in red. Anterior is to the left in all panels.
 806



Supplementary Figure 4. Snapshots from live imaged videos in control and *Sox21b-1* knockdown embryos

Ventral view of the germ disc in **A**) GFP dsRNA control embryos, showing the cumulus formation (red dotted lines), cumulus migration (red dotted arrow) and dorsal field opening (red dotted line and arrows). **B**) Class I *Sox21b-1* knockdown embryos showing cumulus formation, the partial migration of mesenchymal cells and limited dorsal field opening, which is also seen but more severely disrupted in Class II embryos (**C**), and absent in class III (**D**). Anterior is to the left, opposite to cumulus migration.



Supplementary Figure 5. Cell death and cell proliferation in *Sox21b-1* knockdown embryos

Ventral view of stage 5 control embryos stained for Cleaved-Caspase3 (A) and PHH3 (E). Cell death is not detectable in control embryos, but a high level of proliferation can be seen. In *Sox21b-1* knockdown embryos, clusters of cells undergoing cell death can be found (B), as well as a decrease in proliferation in the knockdown embryos compared to controls (n= 15 for each staining) (F). Embryos at stage 9 stained for Cleaved-Caspase 3 (C) and PHH3 (G) show that only a small amount of cell death occurs in the SAZ, and that there is proliferation detectable throughout the entire embryo. Cell death is visible in the head extraembryonic layer in *Sox21b-1* pRNAi embryos (D), and less proliferation is detected in stage 9 knockdown embryos (n= 15 for each staining). Anterior is to the top. Magnifications are 100X and 400x respectively.

836 **Table S1. List of primers used in this paper.**

Oligonucleotide name	Sequence (5' to 3')
In situ Hybridization probes - Cloning	
Caudal_fw	CCCATGCGGAGTTATGGACA
Caudal_rv	GTCCTGGTTCTGCCTGGATT
Dfd_A_fw	CCCCTGTAAGTTATGGCCC
Dfd_A_rv	AGCACTGGGTGCTGTTTCT
Dpp_fw	ATGCGCCAGCGCATTGTTGGCT
Dpp_rv	ACGGCAACCACATCCTTCAACAAC
Delta_fw	CTGTCGTTTGGGTGGCAAG
Delta_rv	CCCCATTGAGGCATGGTTCT
Engrailed_fw	ATGATACCAATGAGAACTCGA
Engrailed_rv	CCATTAATTGCAATGCCAGT
Ets4_fw	AGGTCCACCTCCCTATGT
Ets4_rv	ACGCTCAACGTACAGGA
Fkh_fw	CATGCCCATGTCCCTCAAC
Fkh_rv	AAGCGTTTTTGGCGCCTTAG
Hairy_fw	AAATACGGCCACAGTCAGGG
Hairy_rv	ATCCGAGCTTATGCTCACCG
Hedgehog_fw	GTGCCTGGCCGCATTAGTG
Hedgehog_rv	TGAGTCACCATCGAAACATC
Labial_fw	GGACAACTACGTGCAGGACA
Labial_rv	AGCTGAAACAGACGCTCCTC
Sox21b-1_fw	ATGCAAGCTCCGCAAATCGTACAAAA
Sox21b-1_rv	TTACATCTGTAATGGCATGCCACG
Twist_fw	ACGTTAGGACGAATCCACTG
Twist_rv	CTGGGCTCTCTGAACCTG
Wnt8_fw	CTATGCAGACAGCGTTGCTATTG
Wnt8_rv	GGTGAAATTTTATTGTAGATTAGCTGG
dsRNA synthesis	
DI_dsRNA_fw	TAATACGACTCACTATAGGATGTAAGCGAGTTCTGGACTCAAG ACA
DI_dsRNA_rv	TAATACGACTCACTATAGGCACGTTCTCCATTAGAGCACGGC TTG
GFP_dsRNA_fw	TAATACGACTCACTATAGGCGTGTCCGGCGAGGGCGAGGG
GFP_dsRNA_rv	TAATACGACTCACTATAGGAGGACCATGTGATCGCGCT
Sox21b1_dsRNA_F1_fw	TAATACGACTCACTATAGGATGCAAGCTCCGCAAATCGTAC
Sox21b1_dsRNA_F1_rv	TAATACGACTCACTATAGGAGAAGAGGCAGGATAGCCGC
Sox21b1_dsRNA_F2_fw	TAATACGACTCACTATAGGTCAAGTGTCTGGATCAGCAGC
Sox21b1_dsRNA_F2_rv	TAATACGACTCACTATAGGTTACATCTGTAATGGCATGCCAC
Wnt8_dsRNA_fw	TAATACGACTCACTATAGGCTATGCAGACAGCGTTGCTATTG
Wnt8_dsRNA_rv	TAATACGACTCACTATAGGGGTGAAATTTTATTGTAGATTAGCT GG

837
838
839
840
841
842

Table S2. Phenotypic frequencies for each fragment (1 and 2) and GFP (control) dsRNA – 30 embryos per cocoon for each spider were pooled and the characteristics were divided into three phenotypical classes.

Fragment 1					
	Wild Type	Class I	Class II	Class III	Dead
Cocoon 1	205				5
Cocoon 2	0	21	47	110	32
Cocoon 3	0	35	54	121	
Cocoon 4	0	42	105	46	17
Cocoon 5	27	19	41	63	39
Fragment 2					
	Wild Type	Class I	Class II	Class III	Dead
Cocoon 1	210				
Cocoon 2	0	20	20	91	79
Cocoon 3	0	49	37	124	
Cocoon 4	0	20	27	133	30
Cocoon 5	16	8	67	84	35
GFP dsRNA					
	Wild Type	Dead			
Cocoon 1	56	4			
Cocoon 2	53	7			
Cocoon 3	50	10			
Cocoon 4	56	4			
Cocoon 5	48	12			