N-linked glycosylation restricts the function of Short gastrulation to bind and shuttle BMPs

Negreiros, E.[1]; Herszterg, S. [1,2]; Hwa, K.[3]; Câmara, A. [1,4]; Dias, W.B. [4]; Carneiro, K. [1]; Bier, E. [3]; Todeschini, A. [4*] ; Araujo, H. [1,5*]

* corresponding authors

Address

Federal University of Rio de Janeiro
Av. Carlos Chagas Filho, 373
Bloco F – Sala F2-031
Ilha do Fundão – RJ - Brazil
CEP: 21941-902
e-mail: haraujo@histo.ufrj.br
adrianet@biof.ufrj.br

[1]Institute for Biomedical Sciences
[2] present address: The Francis Crick Institute, London, UK.
[3] Division of Biological Sciences, University of California at San Diego
[4] Institute of Biophysics Carlos Chagas Filho
[5] Instituto Nacional de Ciência e Tecnologia em Entomologia Molecular, Brasil (INCT-ENEM)

© 2018. Published by The Company of Biologists Ltd.

Development • Accepted manuscript

Summary Statement

N-glycosylation restricts Short gastrulation function during Drosophila development by controlling extracellular protein distribution. This adds another layer of complexity to regulation of Bone Morphogenetic Protein signals.

Keywords

Glycosylation, Bone Morphogenetic Protein, Short gastrulation, Chordin, Morphogen

Abstract

Disorders of N-linked glycosylation are increasingly reported in the literature. However, targets responsible for the associated developmental and physiological defects are largely unknown. Bone Morphogenetic Proteins (BMPs) act as highly dynamic complexes to regulate several functions during development. The range and strength of BMP activity depend on interactions with glycosylated protein complexes in the extracellular milieu. Here we investigate the role of glycosylation for the function of the conserved extracellular BMP antagonist Short gastrulation (Sog). We identify conserved N-glycosylated sites and describe the effect of mutating these residues on BMP pathway activity in Drosophila. Functional analysis reveals that loss of individual Sog glycosylation sites enhances BMP antagonism and/or increases the spatial range of Sog effects in the tissue. Mechanistically, we provide evidence that N-terminal and stem glycosylation controls extracellular Sog levels and distribution. The identification of similar residues in vertebrate Chordin proteins suggests that N-glycosylation may be an evolutionarily conserved process that adds complexity to the regulation of BMP activity.

Development • Accepted manuscript

Introduction

Glycosylation of secreted proteins in multicellular organisms increases structural diversity extending the possibilities of cell responses to the microenvironment. Protein glycosylation depends on a series of reactions taking place in different cellular compartments. Activated nucleotide-sugar precursors are synthetized in the cytoplasm, and imported into the lumen of the endoplasmic reticulum (ER) and Golgi compartments, where these sugar substrates are used as the building blocks. N-linked glycosylation is initiated in the ER by adding a conserved oligosaccharide en bloc from a dolichol-linked precursor oligosaccharide onto newly translated proteins. After an initial trimming in the ER the oligosaccharide chain is processed and modified in the Golgi apparatus (Stanley et al., 2017). A series of glycosyltransferases and multiprotein complexes regulate these steps. Improper sugar trimming is linked to errors in protein folding and protein degradation in the ER. Notably, a series of congenital disorders of glycosylation have been reported, with the majority affecting primarily N-glycan assembly (Freeze et al., 2014).

A great challenge is to understand how glycosylation modulates the function of their molecular targets. Critical targets of glycosylation are the Bone Morphogenetic Proteins (BMPs) and their binding partners, which control embryonic development and tissue patterning (Bier and De Robertis, 2015). Extracellular proteoglycans and glycoproteins modulate BMP signaling in several species. For instance, *Xenopus* and mammalian N-acetylgalactosaminyltransferases inhibit BMP signaling (Herr et al., 2008). In zebrafish, embryos injected with morpholinos against beta1,4-galactosyltransferase have reduced activation of the BMP-dependent transcription factors Smad1/5/8 (Machingo et al., 2006). Likewise, mutation of glycosylation sites in mouse Twisted gastrulation (Tsg), an important regulator of BMP activity, reduces its binding to BMPs (Billington et al., 2011). Glycosylation also controls the secretion and folding of human BMP-2 (Hang et al., 2014), BMP receptor recognition, specificity, and binding strength (Lowery et al., 2014, Saremba et al., 2008).

Development • Accepted manuscript

In Drosophila, activity of the BMPs Decapentaplegic (Dpp), Glass bottom boat (Gbb) and Screw (Scw) depend on glycan-based interactions. Early work has shown that genes involved in proteoglycan biosynthesis such as *division abnormally delayed* (*dally*), *sugarless* (*sgl*), and *sulfateless* (*sfl*) regulate Dpp function (Häcker et al., 2005, Selleck, 2000). Furthermore, glycosylation can transform Drosophila Dpp from a long-range to a short-range signal (Humphreys et al., 2013). However, the role of N-linked glycosylation on Drosophila BMP function is less explored. N-linked glycans in Drosophila are less complex than in vertebrates (North et al., 2006, Aoki et al., 2007, Gagneux and Varki, 1999). Nonetheless, several enzymes in the N-glycosylation pathway have been reported to impair Drosophila development (Wandall et al., 2003, Tian and Ten Hagen, 2006, Yamamoto-Hino et al., 2015).

Dpp activity depends on the formation of protein complexes with Tsg proteins, Tolloid (Tld) metalloproteases and the dedicated antagonist Short gastrulation (Sog). Complex formation is required to regulate Dpp activity range during formation of the veins in the pupal wing and to specify the amnioserosa, the dorsal-most extra-embryonic tissue in the embryo. In the pupal wing, Dpp and Gbb ligands are transported from the longitudinal veins in a complex with Sog, the Tsg-family protein Crossveinless and the Tolloid-related (Tlr) metalloprotease for signaling in the posterior crossvein (PCV) forming area (Serpe et al., 2005, Serpe et al., 2008, Ray and Wharton, 2001). Longitudinal veins also require Sog and Dpp antagonism (Yu et al., 1996) and may depend on shuttling peak amounts of BMPs to the center of the provein domains by the action of Sog, which is produced by the intervein cells (Araujo et al., 2003, Negreiros et al., 2010). During embryogenesis, laterally secreted Sog binds Dpp and Scw in a complex with Tsg that Inhibits Dpp locally and facilitates long-range ligand diffusion towards the dorsal midline. Away from the Sog source, Tld cleaves Sog, delivering peak BMP levels for receptor activation and dorsal amnioserosa formation (Peluso et al., 2011, Sawala et al., 2012, Mizutani et al., 2006, Umulis et al., 2006, Eldar et al., 2002). Sog is a N-glycosylated protein that presents evolutionarily conserved structure and function (François and Bier, 1995, Marqués et al., 1997). Sog and the vertebrate Chordin homologs regulate dorsal-ventral patterning across phyla by antagonizing BMP activity and regulating BMP spread (Bier and De Robertis, 2015). Furthermore, Sog (Srinivasan et al., 2002) and Chd (Plouhinec et al., 2013) form morphogen gradients in the extracellular space in

Development • Accepted manuscript

invertebrate and vertebrate embryos respectively. Since Sog ultimately regulates the pattern of BMP activity in tissues, modifications that alter Sog binding and distribution will potentially have great impact on BMP function. Here we investigate whether Sog glycosylation is important for Sog function. We show that Sog glycosylation mutants alter BMP activity in cells, wings and embryos, consistent with a role in regulating BMP activity during Drosophila development.

Results

BMP signaling is impaired in *frc-* wings

fringe connection (*frc*) encodes a UDP-sugar transporter that transports a broad range of UDP-sugars for the synthesis of glycans, including N-linked types, GAGs, and mucins (Selva et al., 2001, Goto et al., 2001). *frc* is thus at the basis of glycoprotein and glucosaminoglycan synthesis. Loss-of-function *frc* alleles are pupal lethal and *frc*- wings of adult escapers show notches at the margin, due to impairment of the Notch pathway (Goto et al., 2001). *frc*- wings also display veins of uneven width which may result from the impairment of other signaling pathways as well (Fig.1A,B). Given the implication of the BMP pathway in wing vein patterning, we investigated whether *frc* regulates BMP pathway activity.

During pupal wing development, *dpp* is expressed in the center of broad and irregular (4-7 cells wide) longitudinal vein competent, or provein domains (Blair, 2007). Dpp is required for vein fate maintenance and refinement, and formation of the vein proper requires BMP activity to be restricted to the center of the proveins by the action of both its receptor *thickveins* (*tkv)* and *sog* (de Celis, 1997, Yu et al., 1996). Sog is expressed in the adjacent intervein domains from where it diffuses into the provein territory, likely to control the transport of Dpp outwards and along the veins (Yu et al., 1996, Araujo et al., 2003, Negreiros et al., 2010). BMP activity in the provein domains can be detected with anti-phosphorylated Mad (pMad) antisera. From 24 hours after puparium formation (h APF) to wing expansion, pMad becomes restricted to quasi-continuous 3-4 cells wide domains along the longitudinal veins, as well as in the posterior cross vein (PCV) (Fig. 1C) (de Celis, 1997, Conley et al.,

Development • Accepted manuscript

2000). We observed that, in *frc-* wings, pMad staining is patchy along longitudinal veins and enlarged at the PCV, consistent with the resulting veins of uneven width and extra PCV tissue in adult wings (Fig. 1D). This pattern could result from uneven BMP diffusion and/or uneven BMP receptor activation within the provein domain, based on modified interactions with glycosylated molecules.

Sog distribution is also altered in *frc-* wings. We have shown that Sog diffusion from intervein-producing cells into the provein domains takes place only at the dorsal surface of wild type wings and is restricted to the basolateral domain of the dorsal epithelium (Negreiros et al., 2010). Furthermore, the use of antisera against N- and C-terminal Sog epitopes (8A and 8B, respectively) suggests that only C-terminal Sog fragments are able to enter the provein domain (Araujo et al., 2003, Negreiros et al., 2010) (Figs. 1F and 1H). However, in *frc-* wings, both the N- and C-terminal epitopes of Sog are detected in the provein domains (Fig. 1E,G). These results suggest that glycosylation is necessary to control the distribution of Sog between the intervein and vein-competent domains and that Sog interactions based on glycosylated residues could be modified in *frc-* mutants. Since *frc* modifies the loss-of-vein phenotype of enhancer piracy *sog* lines (Fig. S1) and Sog is a glycoprotein (Marqués et al., 1997), we decided to investigate whether Sog glycosylation regulates BMP function.

Sog and Chordin putative glycosylation sites are conserved

It has been known for a long time that Sog is a glycoprotein (Marqués et al., 1997). However, how glycosylation contributes to Sog function has not been investigated. The evolutionary conservation of putative glycosylation sites suggests a positive selection for these sites in protostome and deuterostome lineages. We have aligned Sog sequences from the 12 sequenced Drosophila species and found that, among the six predicted glycosylation sites, three are conserved in all species analyzed (Fig. 2A; Fig. S2B). The first conserved site is located after the first cystein rich (CR) domain, in close proximity to a Tld cleavage site (Peluso et al., 2011, Marqués et al., 1997) (Fig. 2B). The second and third sites lie in the stem and after the second CR domain, respectively (Fig. 2B). Putative glycosylation sites in locations corresponding to the second site are consistently detected in vertebrate species as well (Fig. S2B).

Development • Accepted manuscript

To investigate a functional role for Sog glycosylation, we generated single, double and triple Sog glycosylation mutant constructs, in which we abolished the conserved putative N-glycosylation sites (Asn-X-Ser/Thr) by mutating the Asn residues to Gln. We refer to these constructs as SogN mutants (N1, N2, N3, N23 and N123) (Fig. 2C). Expression of V5/His-tagged double and triple mutants in S2 cells revealed that these mutants have a reduced molecular weight when compared to wild type Sog, as seen on SDS-PAGE (Fig. 2C). Treatment of cells with tunicamycin, a specific inhibitor of N-linked glycosylation, abolishes this difference (Fig. 2C). This strongly suggests that the identified sites are N-glycosylated and that a decrease in the number of sugar side chains is responsible for the differential migration pattern of SogN mutants.

Since loss of glycosylation is frequently associated with impaired protein folding (Caramelo and Parodi, 2015) we tested whether SogN mutants are appropriately secreted. We transfected S2 cells with equivalent amounts of wild type or mutated epitope tagged *sog* constructs. After 48h of induction, we collected cell (C) or medium (M) samples for Western blot and Elisa. Secretion of Sog into the medium was slightly decreased only for constructs bearing the second mutated site (N2 and N123) and was unaffected for the other mutants (Fig. 2D,E). Therefore, eventual phenotypes resulting from *sogN* expression (see below) are unlikely to result from impaired protein secretion.

Sog glycosylation mutants display increased BMP binding

Next, we analyzed whether SogN mutants displayed differences in BMP binding properties and cleavage by metalloproteases. Full-length Sog binds both Dpp homodimers and heterodimers, but shows a greater affinity for Dpp heterodimers (Fig. 3A, compare lanes 2 and 4) (Shimmi et al., 2005). This preference is modified by the presence of Tsg, which increases the affinity of full-length Sog for Dpp homodimers (Fig. 3A, lanes 5 and 6). This behavior is distinct for the N-terminal Supersog cleavage product, which is able to bind Dpp alone (Fig. 3B, lanes 4 and 6). Interestingly, the Sog triple glycosylation mutant N123 shows increased binding to Dpp homodimers in the absence of Tsg (Fig. 3C), suggesting that Sog glycosylation modulates Sog binding to BMPs. Among the single

Development • Accepted manuscript

glycosylation mutants, SogN1 was the sole mutant showing enhanced binding to Dpp homodimers in the absence of another BMP or Tsg (Fig. 3D, lane 7).

Sog glycosylation could also affect Sog cleavage by Tld metalloproteases and impact on the consequent release of BMPs for receptor binding. Unlike vertebrate Chordin, Sog cleavage by Tld metalloproteases relies on the presence of BMPs (Marqués et al., 1997, Peluso et al., 2011). We found that SogN mutants display a cleavage pattern similar to wild type Sog (Fig. 3D, right panel and Yu et al., 2000). Therefore, loss of Sog glycosylation does not affect cleavage by Tld metalloproteases, although a decrease in cleavage efficiency cannot be ruled out in these experiments.

Loss of glycosylation sites enhances Sog function *in vivo*

The differences in BMP binding described above suggested that loss of glycosylation could modify Sog function *in vivo*. To test this prediction, we assayed the effects of overexpressing *sog* glycosylation mutants during embryogenesis and pupal wing development. During pupal wing development, Dpp is expressed in the longitudinal veins where it establishes the vein proper domain (de Celis, 1997, Sotillos and De Celis, 2005). In addition, Dpp is transported from longitudinal veins to initiate BMP signals and formation of the PCV (Conley et al., 2000, Serpe et al., 2005, Ralston and Blair, 2005, Shimmi et al., 2005) Shimmi et al., 2005). Ubiquitous overexpression of wild type *sog* in the wing leads to various degrees of longitudinal vein truncation and PCV loss (Yu et al., 1996, Serpe et al., 2005)(Fig. S3 and Fig. 4B). Vein-restricted *sog* overexpression by using the shortvein Gal4 driver (*shv*-Gal4) (Sotillos and De Celis, 2005) also results in longitudinal vein truncation to various degrees, as well as partial or complete PCV loss (Fig. 4A,C). We observed that overexpression of *sogN* mutants leads to more severe longitudinal vein truncation phenotypes when compared to wild type *sog*, with *sogN1* and *sogN2* showing the greatest effect (Fig. 4A). Consistent with inhibition of the BMP pathway exerted by *sog*, overexpression of *sogWT* and *sogN* also increases the severity of longitudinal vein truncation in the *dpp*[shv] background (Fig. S4).

sog overexpression also results in PCV loss. We find that ubiquitous overexpression of *sogN* mutants in the wing (with MS1096-Gal4 driver) increases

Development • Accepted manuscript

PCV loss when compared to *sogWT* overexpression (Fig. 4B). This difference is exacerbated when the overexpression is restricted to the longitudinal proveins by using the *shv*-Gal4 driver, with *sogN1* and *sogN2* again showing the strongest effects (Fig. 4C). These different effects on the PCV produced by local (ubiquitous, MS1096-Gal4) versus at-a-distance (longitudinal proveins, *shv*-Gal4) overexpression of *sog* mutants indicate that the transport of BMP signaling complexes in the extracellular space is affected in the *sogN1* and *sogN2* mutants.

Accordingly, overexpression of *sogWT* and *sogN* mutants in longitudinal proveins affects the pMad pattern in the pupal wing. In the longitudinal proveins, pMad staining is decreased and uneven in *sogWT* overexpression wings when compared to control wings (Fig. 4D,E) consistent with inhibition of BMP signaling by Sog in these domains. In contrast, the PCV appears enlarged and with increased pMad levels (Fig. 4E), indicating increased BMP transport from the longitudinal proveins to the PCV domain. Overexpression of *sogN* mutants seems to exacerbate these defects, with the PCV often appearing detached from the longitudinal proveins (Fig. 4F-H), consistent with the phenotypes observed in adult wings (Fig. 4A,C) and with increased BMP binding (Fig. 3D). Interestingly, the pMad pattern resulting from *sog* overexpression (decreased and non-uniform levels in longitudinal proveins and enlarged PCV) is reminiscent of the pMad pattern observed in *frc-* wings (compare Fig. 4F-H to Fig. 1C,D).

During embryogenesis, *sog* is expressed in the lateral neuroectoderm where it inhibits *dpp* from auto-activating (Biehs et al., 1996). Sog also diffuses dorsally to concentrate BMP activity to the dorsal-most region of the embryo (Mizutani et al., 2005, Shimmi et al., 2005, Ashe and Levine, 1999, Decotto and Ferguson, 2001). Peak BMP activity leads to dorsal expression of *zen*, *RACE* and *rho*, as well as the formation of the amnioserosa, which is absent in loss-of-function *sog* mutants (Ray et al., 1991, Yu et al., 2000). After gastrulation, amnioserosal cells can be detected with antibodies against the Kruppel (Kr) transcription factor. To assay the effects of *sogN* mutants in the embryo we drove early ubiquitous expression of *sogWT and sogN* mutants with a maternal Gal4 driver (matα-Gal4). *sogWT* overexpression does not significantly alter the domain of *zen* expression or the number of Kr+ cells, as previously shown (Yu et al., 2000) (Fig. 5A,B,F,G). However, overexpression of

Development • Accepted manuscript

sogN1 and *sogN3* reduces the domain of *zen* expression as well as the number of Kr+ cells (Fig. 5C,E,H,J,K,L), indicating that these mutants exert an inhibitory effect on BMP signaling.

Taking into account that, in the wing, *sogN* mutants displayed different effects locally versus at-a-distance, we performed a similar analysis in the embryo. We induced localized expression of *sogN* mutants in the anterior tip of the embryo by using the *bcd*-Gal4 driver (*bcdGCN*-Gal4, (Yu et al., 2000), Fig. 5M-R). Anterior overexpression of *sogN1* and *sogN3* decreases the domain of *zen* expression to the same extent as *sogWT* (Fig. 5O,P,R). However, *sogN2* inhibits *zen* expression further away from the anterior tip (Fig. 5O,Q,S). In addition, in *bcd>sogN2* the *zen* domain is broad in the middle of the embryo, suggesting that SogN2 shuttles the active BMP complex towards the posterior end to a greater extent than wild type Sog (Fig. 5Q).

Altogether, the analysis of *sog* overexpression in the wing and embryo shows that the effects of individual mutations in putative Sog glycosylation sites are context-dependent, with SogN1 and SogN2 showing the greatest effects in the wing and SogN2 in the embryo. Nevertheless, it is important to point out that in all contexts the loss of putative glycosylation sites increases the well-established effects of Sog to inhibit and/or shuttle BMPs.

Glycan removal inhibits Sog retrieval

Precise morphogen activity requires controlling the levels and spatial distribution of extracellular morphogens and their antagonists. In Drosophila, endocytosis controls the amount of Sog protein in the extracellular space and thus BMP activity (Srinivasan et al., 2002, Negreiros et al., 2010). Sog suppresses the action of excess Gbb in the wing and Scw in the embryo, but not Dpp (Neul and Ferguson, 1998, Nguyen et al., 1998, Yu et al., 2000). Accordingly, extracellular Sog is able to bind the BMPs Scw and Gbb, and with lower affinity to Dpp. However, in the presence of Tsg, Sog binds to and antagonizes both Dpp heterodimers and homodimers (Ross et al., 2001, Shimmi et al., 2005). Importantly, Sog retrieval from the extracellular space also increases in the presence of Tsg (Fig. 6A), regardless of the BMP involved. Sog retrieval from the extracellular space is dependent on BMP

Development • Accepted manuscript

and Integrin receptors (BMP receptor type I Tkv, βPS integrin receptor myospheroid, αPS1 integrin receptor multiple edematous wing and αPS2 receptor inflated), since *tkv*, *mys, mew* and *if* knockdowns decrease intracellular Sog amount in S2 cells (Fig. 6B). This indicates that Sog retrieval in the presence of Tsg requires an interaction with a Dpp homo or heterodimers and with Integrin receptors.

Loss of putative glycosylation sites decreases Sog retrieval in S2 cells (Fig. 6C). In the presence of Dpp and Tsg, SogWT-V5 is retrieved from the medium by cells that do not express V5-tagged Sog. In the same condition, intracellular levels of SogN1, SogN2 and the triple mutant SogN123 decrease compared to SogWT (Fig. 6C). This indicates that glycosylated residues at the Sog N-terminus and stem are required for an interaction with membrane receptors that retrieve Sog from the extracellular space. Furthermore, Sog retrieval results in internal vesicle-bound Sog. Co-localization analyses shows that a significant amount of intracellular Sog positive punctae correspond to Rab5 positive endocytic vesicles (Fig. 6E). Notably, the percentage of cells displaying Sog associated to endocytic vesicles decreases by mutating any of the three putative glycosylation sites N1, N2 or N3 (Fig. 6F-H and 6I). Therefore, a likely interpretation of these results is that loss of Sog glycosylation leads to a decrease in receptor-based Sog endocytosis, thus increasing the amount of extracellular Sog. Interestingly, retrieval of wild type and mutant Sog is equivalent in the presence of Tsg+Dpp+Gbb (Fig. 6J).

Discussion

N-linked glycosylation modulates Sog function

Here we have undertaken a biochemical and functional analysis of glycosylation in the BMP antagonist Sog. Our results confirm that Sog is glycosylated and show that three sites target for N-glycosylation are evolutionarily conserved. While appropriate glycosylation may be important for protein folding in general (Caramelo and Parodi, 2015, Jayaprakash and Surolia, 2017) and thus may impact the total amount of active glycoprotein produced, our results indicate that Sog glycosylation controls a higher level of Sog function. Our *in vivo* functional analysis shows that loss of conserved glycosylation sites enhances extracellular Sog activity in all the contexts tested. Specifically, mutating conserved glycosylation sites

Development • Accepted manuscript

increases BMP antagonism in the wing (N1 and N2) and embryo (N1 and N3), consistent with an increase in Dpp binding in S2 cells (N1). Furthermore, loss of Sog glycosylation sites also increases the range of Sog effects in the pupal wing (N1 and N2) and embryo (N2), and decreases the retrieval of Sog protein from the extracellular space in S2 cells (N1 and N2, to a smaller extent N3). Therefore, our results consistently point to a role for the first glycosylation site (N1) in dampening BMP antagonism and for the first and second glycosylation sites (N1 and N2) in restricting BMP shuttling by Sog.

The presence of glycans on extracellular proteins frequently alters the degree and quality of interactions with the extracellular matrix and with other secreted proteins (Tauscher et al., 2016, Dennis, 2017, Jayaprakash and Surolia, 2017). By analyzing the position of the putative Sog glycosylation sites one might gain insight to explain how glycan addition may modify Sog binding to interacting partners. It has been suggested that Chd and Sog assume a horseshoe-like conformation that enables cooperative BMP binding (Troilo et al., 2015, Larraín et al., 2000, Shimmi et al., 2005). In this arrangement, specific N- and C-terminal regions interact with BMPs (CR1, CR3 and CR4) (Sawala et al., 2012, Troilo et al., 2015), Tsg proteins (CR1 and surrounding) (Yu et al., 2004), heparan sulfate proteoglycans (HSPGs, CR1 and CR4) (Jasuja et al., 2004), integrin receptors (N-terminal) (Araujo et al., 2003, Larraín et al., 2000) and collagen (CR1 and CR4) (Sawala et al., 2012) (Fig. 6D). Taking into account that most glycans are present at the external face of Drosophila glycoproteins, it is unlikely that Sog glycosylation directly alters BMP dimer binding at internal sites (Jayaprakash and Surolia, 2017). Alternatively, glycosylation could modify Sog conformation, indirectly changing the interaction with BMPs. Another possibility is that Sog glycosylation provides the structural basis for interaction of the Sog/BMP complex with the extracellular matrix and transmembrane receptors. These interactions could secondarily affect BMP binding and metalloprotease cleavage dynamics as well as movement of the BMP complex in the tissue. Structural analysis of the Sog protein should allow testing this prediction in the future.

The most consistent results gathered in this manuscript concern the N1 and N2 sites. The N-terminal, Cystein Repeat 1 (CR1) containing Sog domain,

Development • Accepted manuscript

harbors a Tld/Tlr site proximal to N1 (Marqués et al., 1997). Metalloprotease cleavage at this site generates an N-terminal fragment termed Supersog, which is a stronger BMP antagonist than full-length Sog (Yu et al., 2000), and is able to bind Dpp homodimers, unlike full-length Sog that only binds BMP heterodimers (Yu et al., 2000, Shimmi et al., 2005) (Fig. 3). Therefore, one mechanism by which SogN1 could increase BMP antagonism would be to favor the generation of Supersog fragments. However, our results contradict this interpretation, since loss of this first glycosylation site does not favor N-terminal Sog cleavage or modify the Sog cleavage pattern. Alternatively, loss of N1 glycosylation may modify interactions that take place near this site. For instance, by favoring BMP binding, either by increasing binding strength or decreasing the requirement for accessory proteins such as Tsg, ultimately resulting in less BMP available for receptor activation. This interpretation is consistent with the ability of SogN1 to bind Dpp in the absence of Tsg, unlike wild type Sog that requires either a BMP heterodimer for binding or the presence of Tsg to bind Dpp homodimers.

Stem (N2) Sog glycosylation controls the BMP activity range

An essential and conserved feature of Sog and its vertebrate Chd homologs is the ability to form highly dynamic complexes with BMP, Tsg and Tld proteins. During embryonic dorsal-ventral patterning Sog and Chd form multiprotein complexes that shuttle BMPs, concentrating BMP activity away from the site of Sog or Chd synthesis in a source-and-sink model of morphogen gradient formation (Bier and De Robertis, 2015). This model relies on the ability of Sog and Chd to diffuse in the extracellular space (Srinivasan et al., 2002, Plouhinec et al., 2013) and to interact with BMP, Tsg and Tld proteins (Mizutani et al., 2005, Umulis et al., 2006, Plouhinec et al., 2013, Reversade and De Robertis, 2005, Ambrosio et al., 2008, Inomata et al., 2008). In the pupal wing, formation of similar Sog/BMP complexes and shuttling BMPs away from the Sog source is also important for PCV formation (Serpe et al, 2008).

By altering the N2 Sog glycosylation site we observe BMP inhibition in the embryo along a greater extension relative to wild type Sog. A slight increase in *zen* expression away from the source of anterior *sogN2* expression is also observed, indicating that BMPs are shuttled further way from the SogN2 source in the process (Figure 5Q). This effect is strongest in the wing, where *sogN2* expression in the

Development • Accepted manuscript

provein territory leads to inhibition of PCV formation close to the source (generating detached PCV and pMAD staining) and increased pMAD staining away from the source (Fig. 4G). These results can be interpreted as indication that Sog glycosylation at the N2 site restricts Sog diffusion and consequently how far from the source it is able to shuttle BMPs. Remarkably, the N2 site is the best conserved among Chordins (Fig. S2B), indicating that the vertebrate BMP activity range may depend on Chd stem glycosylation as well.

Interestingly, in *frc*- wings Sog distribution is abnormal, with high amounts entering the provein domain, away from the intervein Sog-producing source. An uneven pattern of pMAD inside the provein territory is also observed in *frc*- pupal wings, akin to the effect of overexpressing *sogN* inside the provein. Therefore, despite the generalized decrease in glycosylation expected in *frc*- wings, if the altered pMAD pattern in *frc*- results in great part from non-glycosylated Sog entering the provein domain, then an important function of Sog glycosylation would be to control the even spread of BMPs along the center of the provein domain, leading to the formation straight, continuous veins (Yu et al., 1996).

Other developmental processes that require the precise control of BMP spread have been reported to depend on glycan residues. In the female germline stem cell (GSC) niche, the collagen IV Viking and the proteoglycan Dally bind to extracellular Dpp near anterior GSCs and prevent it from spreading further to the posterior, allowing the differentiation of cystoblasts (Guo and Wang, 2009, Wang et al., 2008, Hayashi et al., 2009, Harris and Ashe, 2011). Similarly, during embryonic dorsal closure, *mummy,* that codes for a key enzyme during synthesis of O- and N-linked glycans, is required to define a highly restricted BMP activity field (Humphreys et al., 2013, Schimmelpfeng et al., 2006, Tonning et al., 2006). Therefore, glycosylated Sog, Viking, Dally and Mummy may be classified as a novel type of BMP regulators that can transform BMP from a long-range to a short-range signal.

Glycans control extracellular Sog levels

The resident time of a secreted morphogen in the extracellular millieu has a direct impact on the range it is able to exert its effects (Gonzalez-Gaitan and

Development • Accepted manuscript

Jülicher, 2014). As a secreted morphogen, extracellular Sog levels are regulated by endocytosis. This has proven important for Sog spread during embryogenesis (Srinivasan et al., 2002) and wing development (Araujo et al., 2003, Negreiros et al., 2010). Our S2 cell results show that Sog retrieval from the extracellular space requires Dpp (Tkv) and αPSβPS Integrin receptors (Fig. 6B). Accordingly, intracellular Sog levels are low inside *mew*- and *mys*- clones, implying that αPS1βPS integrin receptors regulate Sog levels in the pupal wing (Negreiros et al., 2010). By mutating Sog glycosylation sites N1 and N2 we show that the amount of Sog retrieved from the extracellular space reduces. These sites are located inside the region of the Sog molecule responsible for Integrin binding (Araujo et al., 2003). Interestingly, Integrin receptors bind Xenopus Chordin, controlling the amount of Chordin that is internalized (Larraín et al., 2000). Taking into account the structural similarities between Sog and Chd, the possibility of a conserved mechanism of glycan-based Sog and Chordin endocytosis should be investigated.

Addition of sugar side chains as a general mechanism to modify BMP function

Several molecules that interact with BMP ligands are modified by glycans, and BMPs themselves are subject to glycan addition (Hang et al., 2014, Lowery et al., 2014, Tauscher et al., 2016). Among these molecules are proteoglycans (Guo and Wang, 2009, Häcker et al., 2005, Selleck, 2000), Tsg family proteins (Billington et al., 2011), as well as Sog/Chd (Marqués et al., 1997). Therefore, disorders of glycosylation or changes in glycan availability will have a compounded impact on BMP function that increases in complexity as increases the number of glycosylated partners in a given biological context. Here we have shown that Sog function is modified at specific, evolutionarily conserved glycosylation sites. The slight variability of effects observed in the different contexts is in consonance with the fact that the population of sugars attached to each glycosylated asparagine in a mature glycoprotein depends on the cell type, physiological and metabolic status of the cell in which the glycoprotein is expressed (Dennis et al., 2009; Dennis JW 2017). Therefore, altering the Sog glycosylation status could add great diversity to Sog glycoprotein function and consequently BMP activity during development.

Development • Accepted manuscript

Materials and methods

Lines used in this study were: loss-of-function *frc*[00073], crossed into a TM6-Tb balancer to identify homozygous pupae, *dpp*[shv], and the MS1096- and matα Gal4 drivers, obtained from the Bloomington Indiana Stock Center. The shv-GAL4 driver was a gift of José DeCelis. All drivers have been described previously (Yu et al., 1996, Sotillos and De Celis, 2005).

Constructs for S2 cell expression

A *Drosophila melanogster sog* cDNA cloned into pBluescript (Yu et al., 1996) was mutated at codons 179 (N1), 520 (N2), or 821 (N3), replacing Asn for Gln (CAA) (Retrogen, USA). Production of double (N23) mutated construct was performed by pBS-SogN1 and pBS-SogN2 digestion with SmaI/Xba and the triple (N123) mutated construct was performed by pBS-SogN1 and pBS-SogN23 digestion with SmaI/NotI. Wild type and mutant *sog* sequences were amplified and inserted into KpnI/NotI sites in pMT-V5/HisC (Invitrogen) to produce C-terminal tagged inducible constructs.

Transgenic constructs

Constructs for Gal4/UAS mediated expression in flies were generated in pTiger, a derivative of pUASp (Ferguson et al., 2012). For the pTiger-SogWT construct, pBS-SogWT and pTiger were digested with KpnI and XbaI and ligated using T4 DNA ligase (NEB). For the pTiger-SogN1, N2 and N3 constructs, SogN1, N2 and N3 sequences were amplified by PCR from pBS-SogN1, N2 and N3, digested with KpnI and XbaI and introduced into KpnI/XbaI digested pTiger using T4 DNA ligase (NEB). pTiger-Sog constructs were sent to Bestgene for injection and integrated into the *attP*VK00027-docking site using the ΦC31 system (Markstein et al., 2008). At least two lines for each construct were generated and presented consistent effects.
All constructs were sequenced before use in cell transfection or transformation. Further details and primers used for production of S2 cell and transgenic constructs are available upon request.

Development • Accepted manuscript

S2 cell culture and transfection

S2 cells were cultured in Schneider's cell medium (GIBCO, Carlsbad, CA) containing 10% heat inactivated fetal bovine serum. Cells were transfected with CellFectin (Invitrogen) according to manufacturer's instructions. For pMT-sog-V5/His constructs protein expression was induced by treatment of 0.7 mM $CuSO_4$ 24 h after transfection. Sog protein was harvested from medium 48h after $CuSO_4$ treatment. Vectors for sog-myc and tsg-His expression are described in (Yu et al., 2000). tld-HA, dpp-HA (Marqués et al., 1997), gbb-FLAG and Mad-FLAG (Ross et al., 2001) were a generous gift from M. O'Connor. Transfection product proteins were produced in the absence of glycosylation by incubating with tunicamycin (10ug/ml). RNA interference for *mys*, *mew*, *if*, *sax* and *tkv* was performed as in Kang and Bier 2010, with double stranded RNA synthesized from primers listed in Supplementary Material.

Elisa

96-well plates (Nunc) were sensitized with 100 µl/well (4µg/ml) of anti-His, the capture antibodie, overnight at 4°C. Wells were washed 3 times with PBS pH 7.2; Tween-20 0.025% (PBST) and blocked with PBS supplemented with 10% FBS (PBS/10% SFB) in a volume of 200 µl/well. Plates were allowed to stand for 2 h at 37°C and washed 3 times with PBST. Subsequently, 50 µl of supernatants from S2 cell cultures were added to the wells. Plates were covered and incubated overnight at 4°C. After washing as plates 4 times with PBST, 100µl (4 µg/ml) of the detection antibodies (anti-V5, Invitrogen) was added to the wells. Plates were incubated for 1h at room temperature and washed 6 times with PBST. Subsequently, 100µl of streptoavidin-alkaline phosphatase (1µg/ml) diluted in PBS supplemented with 10% FBS were added to each well. Plates were incubated for 3h at room temperature. Subsequently, as plates were washed 8 times with PBST and so on in each well, 100µl of the 1.0mg/mL solution substrate bis-azine ethyl benzothiazole sulfonic acid (Sigma) diluted in 20 mM Tris, 100 mM $MgCl_2$. Readings were performed on the Beckman Coulter AD340 reader with 405nm filter.

Development • Accepted manuscript

Immunoblotting and immunoprecipitation

Medium or cell lysates from transfected S2 cells collected 72h after transfection. Cells were harvested and lysed with 200 □l ice cold lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.1% NP-40, protease inhibitor cocktail - Complete, Behringer Manheim) and centrifuged for 20 min at 4 °C to remove cell debris. After preclearing of total cell lysates with protein-G Sepharose beads for 20 min, the lysates were incubated with 30 µl of antibody bound agarose beads (crosslinked with monoclonal anti-V5 antibody, Invitrogen) at 4 °C for 3 h. Beads were washed 3 times with lysis buffer, then boiled in SDS-PAGE sample buffer at 95 °C for 5 min, and electrophoresed on 10% Bis-Tris gel and detected by immunoblotting onto PVDF membranes using anti-V5 (1:2000).

Immunofluorescence and in situ hybridization in pupal wings and embryos

24-48h pupal wings were fixed for 10 min in 4% PFA in PBS, pH 7.0 and permeabilized in PBST buffer (PBS + 0.1% Tween20), blocked with 5% normal goat serum, incubated with primary antibody overnight at 4°C (rabbit anti-phospho-Smad1/5, Cell Signalling 41D10, 1:500; mouse anti-βIntegrin 1:500 DSHB; anti-Sog 8A or 8B antisera 1:500) and subsequently incubated with Alexa fluor secondary antibodies. Embryos were fixed and devitelinized in methanol using standard proceedures, after dechorionation in bleach. Embryos were blocked with 5% normal goat serum in PBST for 1 h and incubated with the primary anti-Kr antibodies (1: 1000, kind donation of Francisco Lopes) overnight at 4 °C. Secondary antibodies used were Alexa Fluor 488 goat anti-rabbit and Alexa Fluor 568 goat anti-mouse (1:500; ThermoFisher Scientific, USA), incubated for 1 h at room temperature. Hoechst 33342 was used at 1 □g/ml for nuclear stains. The fluorescence was detected by confocal microscopy using a Leica SP5-AOBS microscope (Leica Microsystems, Wetzlar, Germany). In situ hybridization was performed as previously described (Araujo and Bier, 2000).

Development • Accepted manuscript

Protein sequence analysis

Short gastrulation homolog protein sequences were downloaded from Flybase or NCBI for the 12 *Drosophila* sequenced genomes, using the 2015 release. Sequence accession numbers are listed in Supplemental Material. Protein sequences were aligned using ClustalW in the MacVector (MacVector, Cary, NC, USA) program. Glycosylation sites were predicted in each protein using NetNGlyc 1.0.

Rab5/Sog co-localization analysis

Sog endocytosis was analyzed by adding medium from control, wild type Sog or N mutant Sog expressing cells, grown in the presence of Dpp-HA and Tsg-His, to non-transfected cells. Rab5 and Sog (wild type or mutant) were detected in the receiving cells with rabbit anti-Rab5 (1:500, Abcam, UK) and mouse anti-V5 (1:500) antisera and subsequently incubated with Alexa fluor secondary antibodies. Co-localization was analyzed with Image J software in single confocal optical sections. The number of cells with at least 2 Rab5+V5+ punctae was quantified for each condition.

Acknowledgements

We are indebted to Mihaela Serpe and Michael O'Connor for kindly providing BMP pathway constructs and flies. We thank Attilio Pane and members of the Araujo and Todeschini labs for suggestions and critical reading of the manuscript. Monoclonal antibodies against Integrin subunits were obtained from the Developmental Studies Hybridoma Bank, created by the NICHD of the NIH and maintained at The University of Iowa, Department of Biology, Iowa City, IA 52242. Stocks obtained from the Bloomington Drosophila Stock Center (NIH P40OD018537) were used in this study.

Development • Accepted manuscript

Competing interests

The authors declare that they have no conflict of interest.

Author contributions

EN, SH, KY, AC, KC and HA performed experiments. All authors analyzed the data. AT, EB and HA developed the approach. SH, KC, EB, AT and HA prepared or edited the manuscript.

Funding

This research was funded by grants to HA (CNPq-Brazil, 477157/2013-0 and FAPERJ E26/010.001/2015) and AT (Faperj, CNPq and CAPES) and EB (NIH R01 NS29870, R01 GM117321, and Paul G. Allen Frontiers Group Distinguished Investigators Award). AC and EN were recipients of CAPES and FAPERJ fellowships respectively. SH was a recipient of CsF Joven Talento Fellowship from CNPq/Brazil.

Development • Accepted manuscript

References

Ambrosio, A. L., Taelman, V. F., Lee, H. X., Metzinger, C. A., Coffinier, C. and De Robertis, E. M. (2008) 'Crossveinless-2 Is a BMP feedback inhibitor that binds Chordin/BMP to regulate Xenopus embryonic patterning', *Dev Cell,* 15(2), pp. 248-60.

Aoki, K., Perlman, M., Lim, J. M., Cantu, R., Wells, L. and Tiemeyer, M. (2007) 'Dynamic developmental elaboration of N-linked glycan complexity in the Drosophila melanogaster embryo', *J Biol Chem,* 282(12), pp. 9127-42.

Araujo, H., Negreiros, E. and Bier, E. (2003) 'Integrins modulate Sog activity in the Drosophila wing', *Development,* 130(16), pp. 3851-64.

Ashe, H. L. and Levine, M. (1999) 'Local inhibition and long-range enhancement of Dpp signal transduction by Sog', *Nature,* 398(6726), pp. 427-31.

Biehs, B., François, V. and Bier, E. (1996) 'The Drosophila short gastrulation gene prevents Dpp from autoactivating and suppressing neurogenesis in the neuroectoderm', *Genes Dev,* 10(22), pp. 2922-34.

Bier, E. and De Robertis, E. M. (2015) 'EMBRYO DEVELOPMENT. BMP gradients: A paradigm for morphogen-mediated developmental patterning', *Science,* 348(6242), pp. aaa5838.

Billington, C. J., Fiebig, J. E., Forsman, C. L., Pham, L., Burbach, N., Sun, M., Jaskoll, T., Mansky, K., Gopalakrishnan, R., O'Connor, M. B., Mueller, T. D. and Petryk, A. (2011) 'Glycosylation of Twisted Gastrulation is Required for BMP Binding and Activity during Craniofacial Development', *Front Physiol,* 2, pp. 59.

Blair, S. S. (2007) 'Wing vein patterning in Drosophila and the analysis of intercellular signaling', *Annu Rev Cell Dev Biol,* 23, pp. 293-319.

Caramelo, J. J. and Parodi, A. J. (2015) 'A sweet code for glycoprotein folding', *FEBS Lett,* 589(22), pp. 3379-87.

Conley, C. A., Silburn, R., Singer, M. A., Ralston, A., Rohwer-Nutter, D., Olson, D. J., Gelbart, W. and Blair, S. S. (2000) 'Crossveinless 2 contains cysteine-rich domains and is required for high levels of BMP-like activity during the formation of the cross veins in Drosophila', *Development,* 127(18), pp. 3947-59.

de Celis, J. F. (1997) 'Expression and function of decapentaplegic and thick veins during the differentiation of the veins in the Drosophila wing', *Development,* 124(5), pp. 1007-18.

Decotto, E. and Ferguson, E. L. (2001) 'A positive role for Short gastrulation in modulating BMP signaling during dorsoventral patterning in the Drosophila embryo', *Development,* 128(19), pp. 3831-41.

Dennis, J. W. (2017) 'Genetic code asymmetry supports diversity through experimentation with posttranslational modifications', *Curr Opin Chem Biol,* 41, pp. 1-11.

Eldar, A., Dorfman, R., Weiss, D., Ashe, H., Shilo, B. Z. and Barkai, N. (2002) 'Robustness of the BMP morphogen gradient in Drosophila embryonic patterning', *Nature,* 419(6904), pp. 304-8.

Ferguson, S. B., Blundon, M. A., Klovstad, M. S. and Schupbach, T. (2012) 'Modulation of gurken translation by insulin and TOR signaling in Drosophila', *J Cell Sci,* 125(Pt 6), pp. 1407-19.

Development • Accepted manuscript

François, V. and Bier, E. (1995) 'Xenopus chordin and Drosophila short gastrulation genes encode homologous proteins functioning in dorsal-ventral axis formation', *Cell,* 80(1), pp. 19-20.
Freeze, H. H., Chong, J. X., Bamshad, M. J. and Ng, B. G. (2014) 'Solving glycosylation disorders: fundamental approaches reveal complicated pathways', *Am J Hum Genet,* 94(2), pp. 161-75.
Gagneux, P. and Varki, A. (1999) 'Evolutionary considerations in relating oligosaccharide diversity to biological function', *Glycobiology,* 9(8), pp. 747-55.
Gonzalez-Gaitan, M. and Jülicher, F. (2014) 'The role of endocytosis during morphogenetic signaling', *Cold Spring Harb Perspect Biol,* 6(7), pp. a016881.
Goto, S., Taniguchi, M., Muraoka, M., Toyoda, H., Sado, Y., Kawakita, M. and Hayashi, S. (2001) 'UDP-sugar transporter implicated in glycosylation and processing of Notch', *Nat Cell Biol,* 3(9), pp. 816-22.
Guo, Z. and Wang, Z. (2009) 'The glypican Dally is required in the niche for the maintenance of germline stem cells and short-range BMP signaling in the Drosophila ovary', *Development,* 136(21), pp. 3627-35.
Hang, Q., Zhou, Y., Hou, S., Zhang, D., Yang, X., Chen, J., Ben, Z., Cheng, C. and Shen, A. (2014) 'Asparagine-linked glycosylation of bone morphogenetic protein-2 is required for secretion and osteoblast differentiation', *Glycobiology,* 24(3), pp. 292-304.
Harris, R. E. and Ashe, H. L. (2011) 'Cease and desist: modulating short-range Dpp signalling in the stem-cell niche', *EMBO Rep,* 12(6), pp. 519-26.
Hayashi, Y., Kobayashi, S. and Nakato, H. (2009) 'Drosophila glypicans regulate the germline stem cell niche', *J Cell Biol,* 187(4), pp. 473-80.
Herr, P., Korniychuk, G., Yamamoto, Y., Grubisic, K. and Oelgeschläger, M. (2008) 'Regulation of TGF-(beta) signalling by N-acetylgalactosaminyltransferase-like 1', *Development,* 135(10), pp. 1813-22.
Humphreys, G. B., Jud, M. C., Monroe, K. M., Kimball, S. S., Higley, M., Shipley, D., Vrablik, M. C., Bates, K. L. and Letsou, A. (2013) 'Mummy, A UDP-N-acetylglucosamine pyrophosphorylase, modulates DPP signaling in the embryonic epidermis of Drosophila', *Dev Biol,* 381(2), pp. 434-45.
Häcker, U., Nybakken, K. and Perrimon, N. (2005) 'Heparan sulphate proteoglycans: the sweet side of development', *Nat Rev Mol Cell Biol,* 6(7), pp. 530-41.
Inomata, H., Haraguchi, T. and Sasai, Y. (2008) 'Robust stability of the embryonic axial pattern requires a secreted scaffold for chordin degradation', *Cell,* 134(5), pp. 854-65.
Jasuja, R., Allen, B. L., Pappano, W. N., Rapraeger, A. C. and Greenspan, D. S. (2004) 'Cell-surface heparan sulfate proteoglycans potentiate chordin antagonism of bone morphogenetic protein signaling and are necessary for cellular uptake of chordin', *J Biol Chem,* 279(49), pp. 51289-97.
Jayaprakash, N. G. and Surolia, A. (2017) 'Role of glycosylation in nucleating protein folding and stability', *Biochem J,* 474(14), pp. 2333-2347.
Larraín, J., Bachiller, D., Lu, B., Agius, E., Piccolo, S. and De Robertis, E. M. (2000) 'BMP-binding modules in chordin: a model for signalling regulation in the extracellular space', *Development,* 127(4), pp. 821-30.
Lowery, J. W., Amich, J. M., Andonian, A. and Rosen, V. (2014) 'N-linked glycosylation of the bone morphogenetic protein receptor type 2 (BMPR2) enhances ligand binding', *Cell Mol Life Sci,* 71(16), pp. 3165-72.

Development • Accepted manuscript

Machingo, Q. J., Fritz, A. and Shur, B. D. (2006) 'A beta1,4-galactosyltransferase is required for Bmp2-dependent patterning of the dorsoventral axis during zebrafish embryogenesis', *Development,* 133(11), pp. 2233-41.
Markstein, M., Pitsouli, C., Villalta, C., Celniker, S. E. and Perrimon, N. (2008) 'Exploiting position effects and the gypsy retrovirus insulator to engineer precisely expressed transgenes', *Nat Genet,* 40(4), pp. 476-83.
Marqués, G., Musacchio, M., Shimell, M. J., Wünnenberg-Stapleton, K., Cho, K. W. and O'Connor, M. B. (1997) 'Production of a DPP activity gradient in the early Drosophila embryo through the opposing actions of the SOG and TLD proteins', *Cell,* 91(3), pp. 417-26.
Mizutani, C. M., Meyer, N., Roelink, H. and Bier, E. (2006) 'Threshold-dependent BMP-mediated repression: a model for a conserved mechanism that patterns the neuroectoderm', *PLoS Biol,* 4(10), pp. e313.
Mizutani, C. M., Nie, Q., Wan, F. Y., Zhang, Y. T., Vilmos, P., Sousa-Neves, R., Bier, E., Marsh, J. L. and Lander, A. D. (2005) 'Formation of the BMP activity gradient in the Drosophila embryo', *Dev Cell,* 8(6), pp. 915-24.
Negreiros, E., Fontenele, M., Câmara, A. R. and Araujo, H. (2010) 'alphaPS1betaPS integrin receptors regulate the differential distribution of Sog fragments in polarized epithelia', *Genesis,* 48(1), pp. 31-43.
Neul, J. L. and Ferguson, E. L. (1998) 'Spatially restricted activation of the SAX receptor by SCW modulates DPP/TKV signaling in Drosophila dorsal-ventral patterning', *Cell,* 95(4), pp. 483-94.
Nguyen, M., Park, S., Marqués, G. and Arora, K. (1998) 'Interpretation of a BMP activity gradient in Drosophila embryos depends on synergistic signaling by two type I receptors, SAX and TKV', *Cell,* 95(4), pp. 495-506.
North, S. J., Koles, K., Hembd, C., Morris, H. R., Dell, A., Panin, V. M. and Haslam, S. M. (2006) 'Glycomic studies of Drosophila melanogaster embryos', *Glycoconj J,* 23(5-6), pp. 345-54.
Oelgeschläger, M., Larraín, J., Geissert, D. and De Robertis, E. M. (2000) 'The evolutionarily conserved BMP-binding protein Twisted gastrulation promotes BMP signalling', *Nature,* 405(6788), pp. 757-63.
Peluso, C. E., Umulis, D., Kim, Y. J., O'Connor, M. B. and Serpe, M. (2011) 'Shaping BMP morphogen gradients through enzyme-substrate interactions', *Dev Cell,* 21(2), pp. 375-83.
Plouhinec, J. L., Zakin, L., Moriyama, Y. and De Robertis, E. M. (2013) 'Chordin forms a self-organizing morphogen gradient in the extracellular space between ectoderm and mesoderm in the Xenopus embryo', *Proc Natl Acad Sci U S A,* 110(51), pp. 20372-9.
Ralston, A. and Blair, S. S. (2005) 'Long-range Dpp signaling is regulated to restrict BMP signaling to a crossvein competent zone', *Dev Biol,* 280(1), pp. 187-200.
Ray, R. P., Arora, K., Nüsslein-Volhard, C. and Gelbart, W. M. (1991) 'The control of cell fate along the dorsal-ventral axis of the Drosophila embryo', *Development,* 113(1), pp. 35-54.
Ray, R. P. and Wharton, K. A. (2001) 'Context-dependent relationships between the BMPs gbb and dpp during development of the Drosophila wing imaginal disk', *Development,* 128(20), pp. 3913-25.
Reversade, B. and De Robertis, E. M. (2005) 'Regulation of ADMP and BMP2/4/7 at opposite embryonic poles generates a self-regulating morphogenetic field', *Cell,* 123(6), pp. 1147-60.

Development • Accepted manuscript

Ross, J. J., Shimmi, O., Vilmos, P., Petryk, A., Kim, H., Gaudenz, K., Hermanson, S., Ekker, S. C., O'Connor, M. B. and Marsh, J. L. (2001) 'Twisted gastrulation is a conserved extracellular BMP antagonist', *Nature,* 410(6827), pp. 479-83.
Saremba, S., Nickel, J., Seher, A., Kotzsch, A., Sebald, W. and Mueller, T. D. (2008) 'Type I receptor binding of bone morphogenetic protein 6 is dependent on N-glycosylation of the ligand', *FEBS J,* 275(1), pp. 172-83.
Sawala, A., Scarcia, M., Sutcliffe, C., Wilcockson, S. G. and Ashe, H. L. (2015) 'Peak BMP Responses in the Drosophila Embryo Are Dependent on the Activation of Integrin Signaling', *Cell Rep,* 13(7), pp. 1519-1520.
Sawala, A., Sutcliffe, C. and Ashe, H. L. (2012) 'Multistep molecular mechanism for bone morphogenetic protein extracellular transport in the Drosophila embryo', *Proc Natl Acad Sci U S A,* 109(28), pp. 11222-7.
Schimmelpfeng, K., Strunk, M., Stork, T. and Klämbt, C. (2006) 'Mummy encodes an UDP-N-acetylglucosamine-dipohosphorylase and is required during Drosophila dorsal closure and nervous system development', *Mech Dev,* 123(6), pp. 487-99.
Selleck, S. B. (2000) 'Proteoglycans and pattern formation: sugar biochemistry meets developmental genetics', *Trends Genet,* 16(5), pp. 206-12.
Selva, E. M., Hong, K., Baeg, G. H., Beverley, S. M., Turco, S. J., Perrimon, N. and Häcker, U. (2001) 'Dual role of the fringe connection gene in both heparan sulphate and fringe-dependent signalling events', *Nat Cell Biol,* 3(9), pp. 809-15.
Serpe, M., Ralston, A., Blair, S. S. and O'Connor, M. B. (2005) 'Matching catalytic activity to developmental function: tolloid-related processes Sog in order to help specify the posterior crossvein in the Drosophila wing', *Development,* 132(11), pp. 2645-56.
Serpe, M., Umulis, D., Ralston, A., Chen, J., Olson, D. J., Avanesov, A., Othmer, H., O'Connor, M. B. and Blair, S. S. (2008) 'The BMP-binding protein Crossveinless 2 is a short-range, concentration-dependent, biphasic modulator of BMP signaling in Drosophila', *Dev Cell,* 14(6), pp. 940-53.
Shimmi, O., Umulis, D., Othmer, H. and O'Connor, M. B. (2005) 'Facilitated transport of a Dpp/Scw heterodimer by Sog/Tsg leads to robust patterning of the Drosophila blastoderm embryo', *Cell,* 120(6), pp. 873-86.
Sotillos, S. and De Celis, J. F. (2005) 'Interactions between the Notch, EGFR, and decapentaplegic signaling pathways regulate vein differentiation during Drosophila pupal wing development', *Dev Dyn,* 232(3), pp. 738-52.
Srinivasan, S., Rashka, K. E. and Bier, E. (2002) 'Creation of a Sog morphogen gradient in the Drosophila embryo', *Dev Cell,* 2(1), pp. 91-101.
Stanley**,** P., Taniguchi, N., Aebi, M. In: Essentials of Glycobiology (ed. Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH). 3rd edition. Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press; 2015-2017. Chapter 9.
Tauscher, P. M., Gui, J. and Shimmi, O. (2016) 'Adaptive protein divergence of BMP ligands takes place under developmental and evolutionary constraints', *Development,* 143(20), pp. 3742-3750.
Tian, E. and Ten Hagen, K. G. (2006) 'Expression of the UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase family is spatially and temporally regulated during Drosophila development', *Glycobiology,* 16(2), pp. 83-95.

Development • Accepted manuscript

Tonning, A., Helms, S., Schwarz, H., Uv, A. E. and Moussian, B. (2006) 'Hormonal regulation of mummy is needed for apical extracellular matrix formation and epithelial morphogenesis in Drosophila', *Development,* 133(2), pp. 331-41.

Troilo, H., Barrett, A. L., Wohl, A. P., Jowitt, T. A., Collins, R. F., Bayley, C. P., Zuk, A. V., Sengle, G. and Baldock, C. (2015) 'The role of chordin fragments generated by partial tolloid cleavage in regulating BMP activity', *Biochem Soc Trans,* 43(5), pp. 795-800.

Umulis, D. M., Serpe, M., O'Connor, M. B. and Othmer, H. G. (2006) 'Robust, bistable patterning of the dorsal surface of the Drosophila embryo', *Proc Natl Acad Sci U S A,* 103(31), pp. 11613-8.

Wandall, H. H., Pedersen, J. W., Park, C., Levery, S. B., Pizette, S., Cohen, S. M., Schwientek, T. and Clausen, H. (2003) 'Drosophila egghead encodes a beta 1,4-mannosyltransferase predicted to form the immediate precursor glycosphingolipid substrate for brainiac', *J Biol Chem,* 278(3), pp. 1411-4.

Wang, X., Harris, R. E., Bayston, L. J. and Ashe, H. L. (2008) 'Type IV collagens regulate BMP signalling in Drosophila', *Nature,* 455(7209), pp. 72-7.

Yamamoto-Hino, M., Muraoka, M., Kondo, S., Ueda, R., Okano, H. and Goto, S. (2015) 'Dynamic regulation of innate immune responses in Drosophila by Senju-mediated glycosylation', *Proc Natl Acad Sci U S A,* 112(18), pp. 5809-14.

Yu, K., Kang, K. H., Heine, P., Pyati, U., Srinivasan, S., Biehs, B., Kimelman, D. and Bier, E. (2004) 'Cysteine repeat domains and adjacent sequences determine distinct bone morphogenetic protein modulatory activities of the Drosophila Sog protein', *Genetics,* 166(3), pp. 1323-36.

Yu, K., Srinivasan, S., Shimmi, O., Biehs, B., Rashka, K. E., Kimelman, D., O'Connor, M. B. and Bier, E. (2000) 'Processing of the Drosophila Sog protein creates a novel BMP inhibitory activity', *Development,* 127(10), pp. 2143-54.

Yu, K., Sturtevant, M. A., Biehs, B., François, V., Padgett, R. W., Blackman, R. K. and Bier, E. (1996) 'The Drosophila decapentaplegic and short gastrulation genes function antagonistically during adult wing vein development', *Development,* 122(12), pp. 4033-44.

Development • Accepted manuscript

Figures

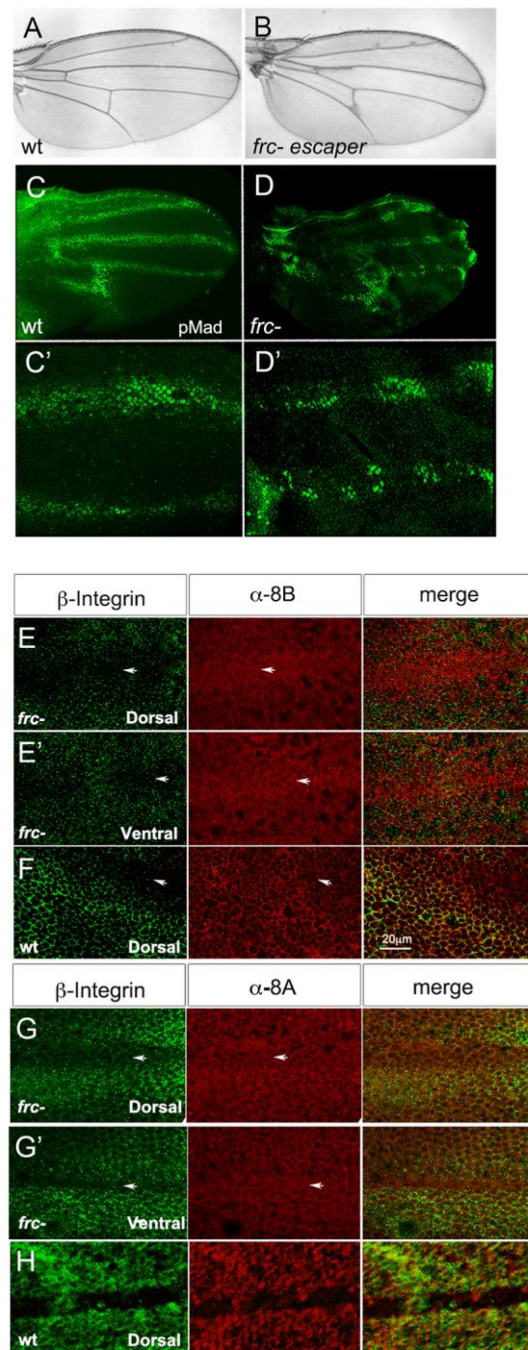

Figure 1: *frc* mutants affect BMP activity and modify Sog distribution in the pupal wing. A) Wild type and B) *frc-* adult wing. C,D) Wild type (C,C') and *frc-* (D,D-) 24h apf wings stained for pMAD. High magnification shows continuity of pMAD staining in wild type (C') and a patchy pattern in *frc-* provein staining (D'). E-H) Double immunolabelling for β-integrin and anti-8B (E,F) or anti-8A (J,H) Sog antisera

Development • Accepted manuscript

in *frc-* (E,G) or wild type (F,H) 24h apf wings. Shown are the dorsal (E,F,G,H) and ventral (E',G') wing epithelia. Note Sog staining inside the provein territory in both the dorsal and ventral wing epithelia in *frc-*.

Development • Accepted manuscript

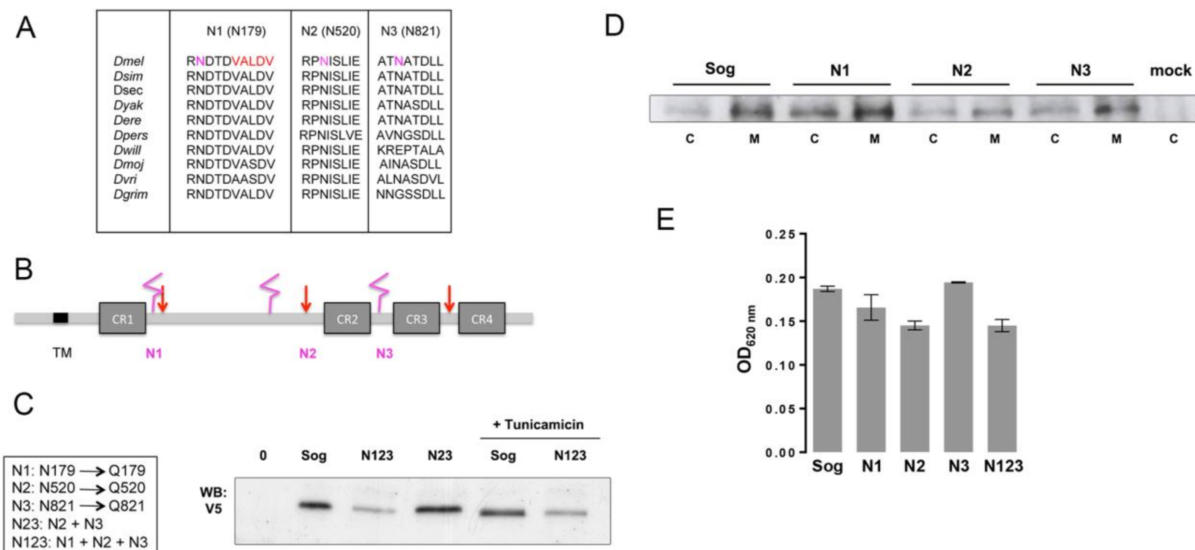

Figure 2: *Drosophila sog* has three conserved glycosylation sites. A) Alignment of Sog sequences from *Drosophila sp*. shows that three putative glycosylation sites are conserved (Arg residue in pink in *D. melanogaster*). In red a Tld/Tlr cleavage site sequence close to N1 is shown. B) Sog protein scheme with the location of the four conserved Cystein Rich (CR) domains, Tld/Tlr cleavage sites (red arrows) and putative glycosylation sites (pink). C-E) Analysis of Sog constructs expressed in S2 cells. C) Glycosylation mutants produced by site-directed mutagenesis and their effect on Sog migration in SDS-PAGE. N23 and N123 migrate faster that wild type Sog. All mutants and wild type Sog bear a C-terminal V5/His tag. Treatment with Tunicamycin decreases wild type Sog Mw and confirms that Sog is glycosylated. D) SDS-PAGE for Sog protein in cells (C) and extracellular medium (M) shows that Sog is secreted and that only the N2 mutation slightly decreases Sog secretion. E) Elisa for S2 cell secreted Sog confirms the analysis in D.

Development • Accepted manuscript

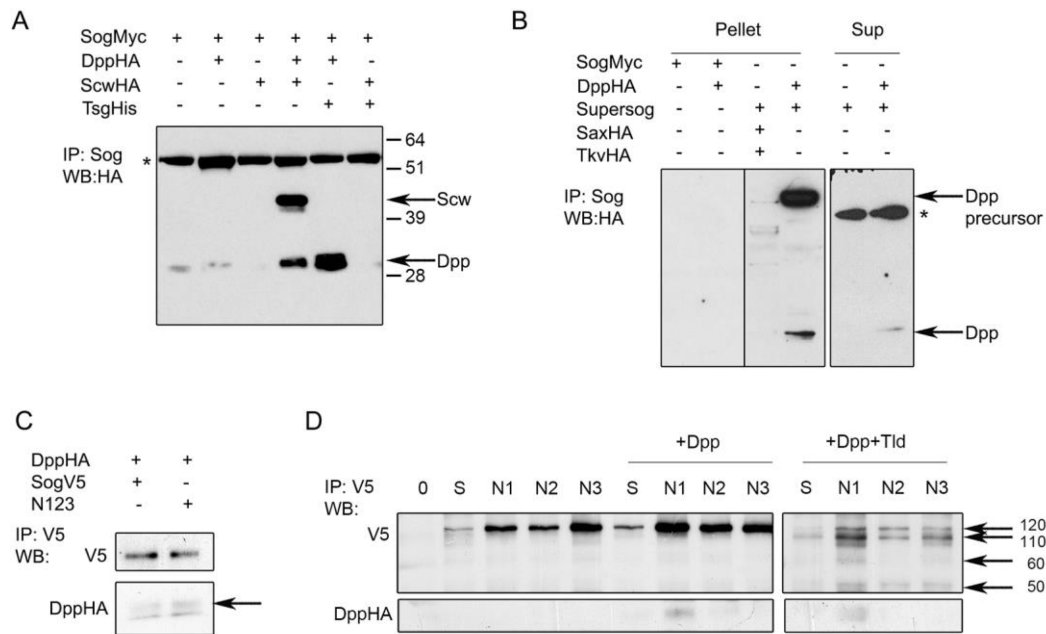

Figure 3: Loss of Sog glycosylation modifies Dpp binding. A) S2 cells transfected with wild type Sog-myc and different combinations of Tsg and BMP ligands. Co-IP with cell supernatants shows that Sog binds only BMP heterodimers in the absence of Tsg, but binds Dpp homodimers in the presence of Tsg. B) S2 cells transfected with wild type Sog-myc or the N-terminal Supersog fragment reveal that Supersog binds Dpp alone, either using cell pellets or supernatants. C) S2 cells transfected with wild type *sog*-V5 or *sogN123*-V5, and *dpp*-HA. Co-IP for V5 shows that the N123 mutant binds more Dpp than wild type Sog. D) S2 cells transfected with wild type *sog* (S) or single glycosylation mutants (N1, N2, N3) plus *dpp*-HA or *dpp*-HA and *tld*-HA. Co-IP for V5. SogN1 binds Dpp alone, while other mutants and wild type Sog do not. Arrows point to the different Sog fragments produced by the Tld metalloprotease and show that the cleavage pattern is similar among all constructs.

Development • Accepted manuscript

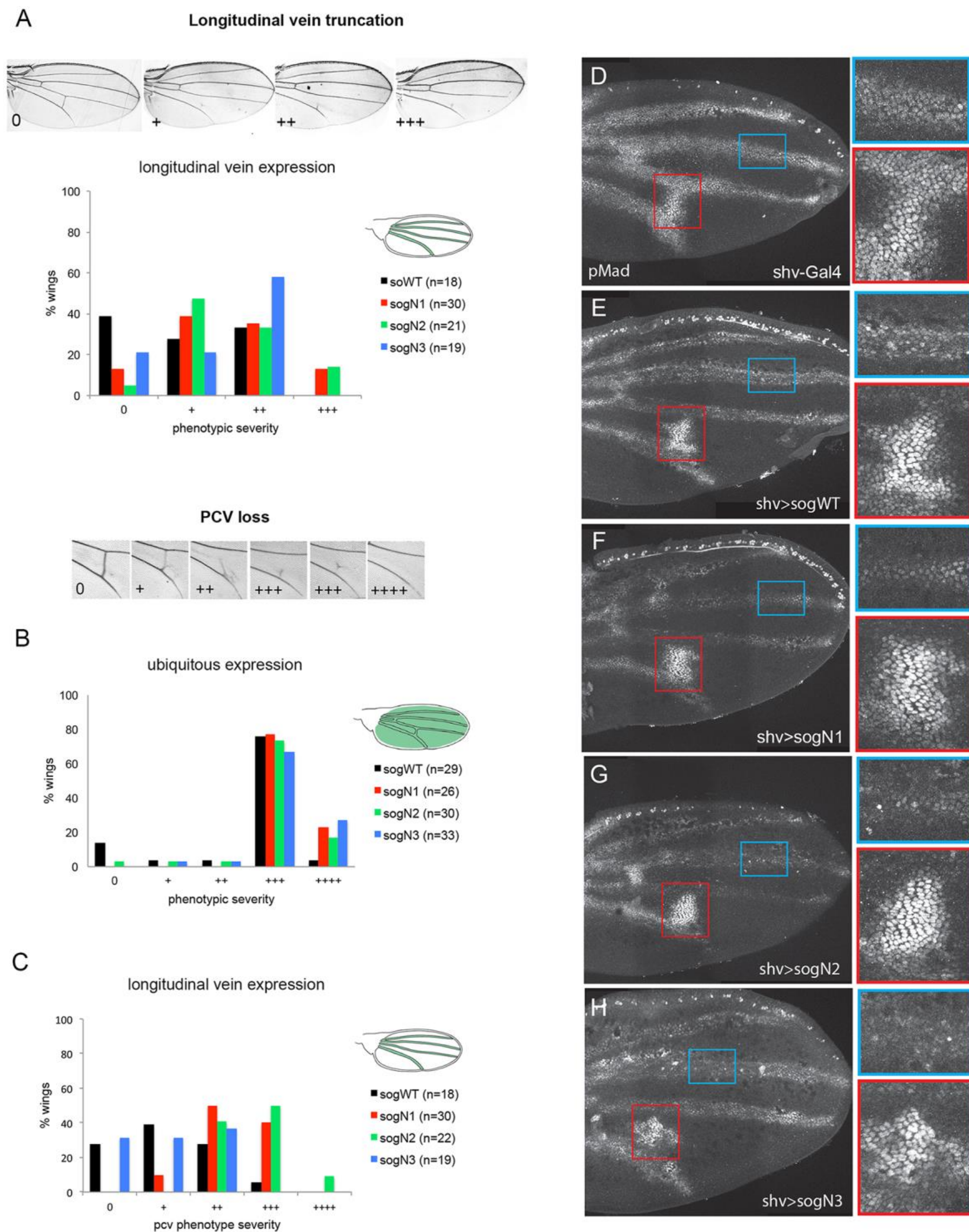

Figure 4. Loss of glycosylation sites enhances Sog function in the wing. A-C) The adult wing venation pattern produced by Gal4/UAS expression of wild type Sog or Sog glycosylation mutants was quantified. Expression induced by the shv-Gal4 driver in the longitudinal veins (A,C) or ubiquitous expression in the wing with the MS1096-Gal4 driver (B) leads to different degrees of longitudinal vein loss (A) or

Development • Accepted manuscript

posterior crossvein (PCV) phenotypes (B,C). D-H) 24hAPF control pupal wing (D) or wings expressing wild type *sog* (E) or *sogN* mutants (F-H), driven by shv-Gal4 and stained for pMad.

Development • Accepted manuscript

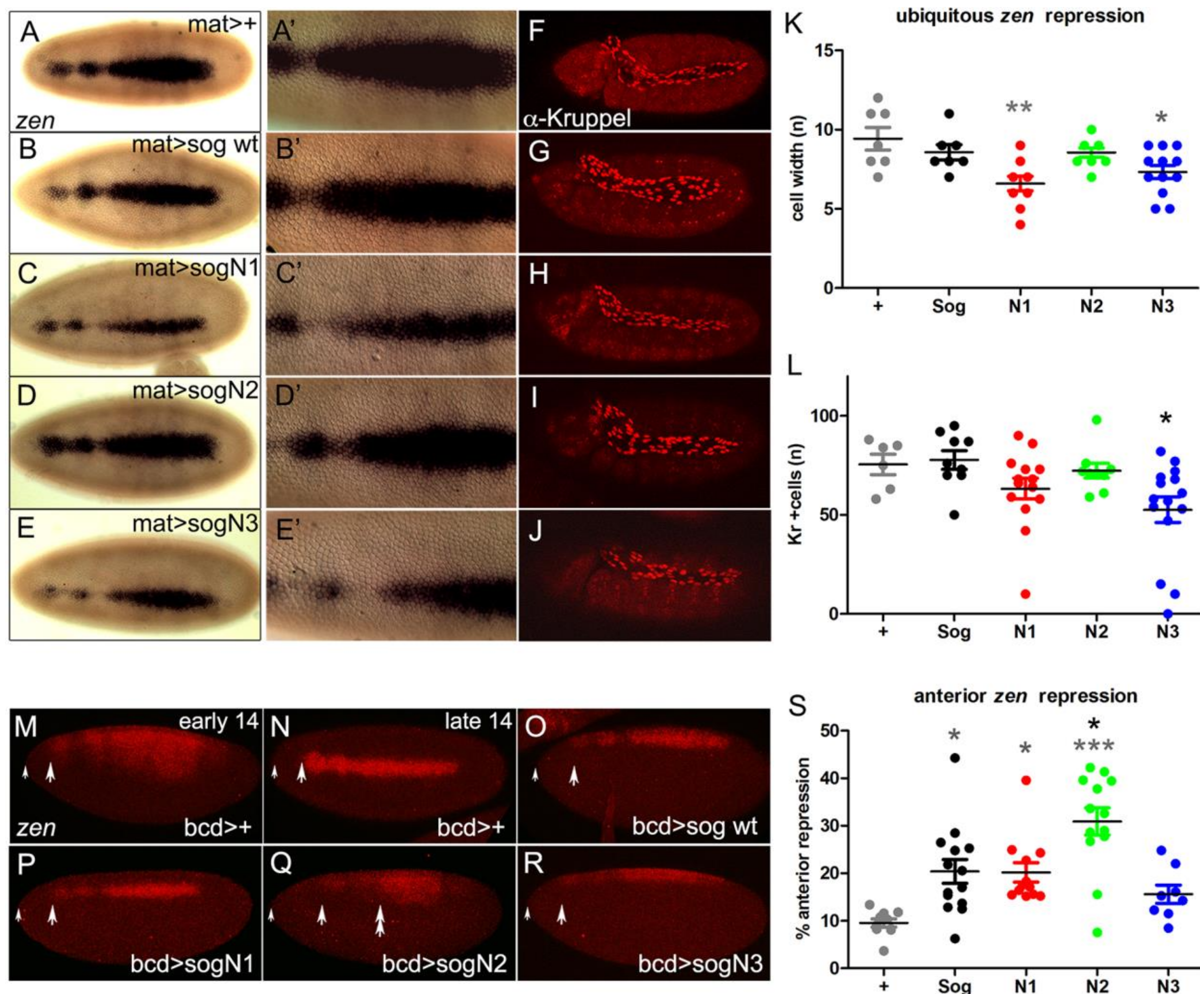

Figure 5. Loss of glycosylation enhances Sog function in the embryo. A-L) Embryonic pattern produced by expression of wild type *sog* or *sog* glycosylation mutants *sogN1, sogN2* or *sogN3*, with the ubiquitous matαGAL4 driver. Effects were analyzed by A-E) dorsal *zen* expression, quantified in K), or F-J) Kruppel protein in amnioserosal cells, quantified in L). M-S) Embryonic pattern produced by localized expression of wild type *sog* or *sog* glycosylation mutants *sogN1, sogN2* or *sogN3*, with the anteriorly restricted bcdGAL4/GCN driver. Effects were analyzed by dorsal *zen* expression. The domain of *zen* repression from the anterior tip of the embryo was measured and quantified in S). A small arrow points to the anterior tip of each embryo. Arrows in M-R point to the most anterior region of *zen* expression detected. Double arrow shows the region where *zen* is expanded relative to wild type *sog* by anterior *sogN2* expression (Q). Statistically significant differences based on Student's t-test (***P≤0.001, **P≤0.01, *P≤0.05).

Development • Accepted manuscript

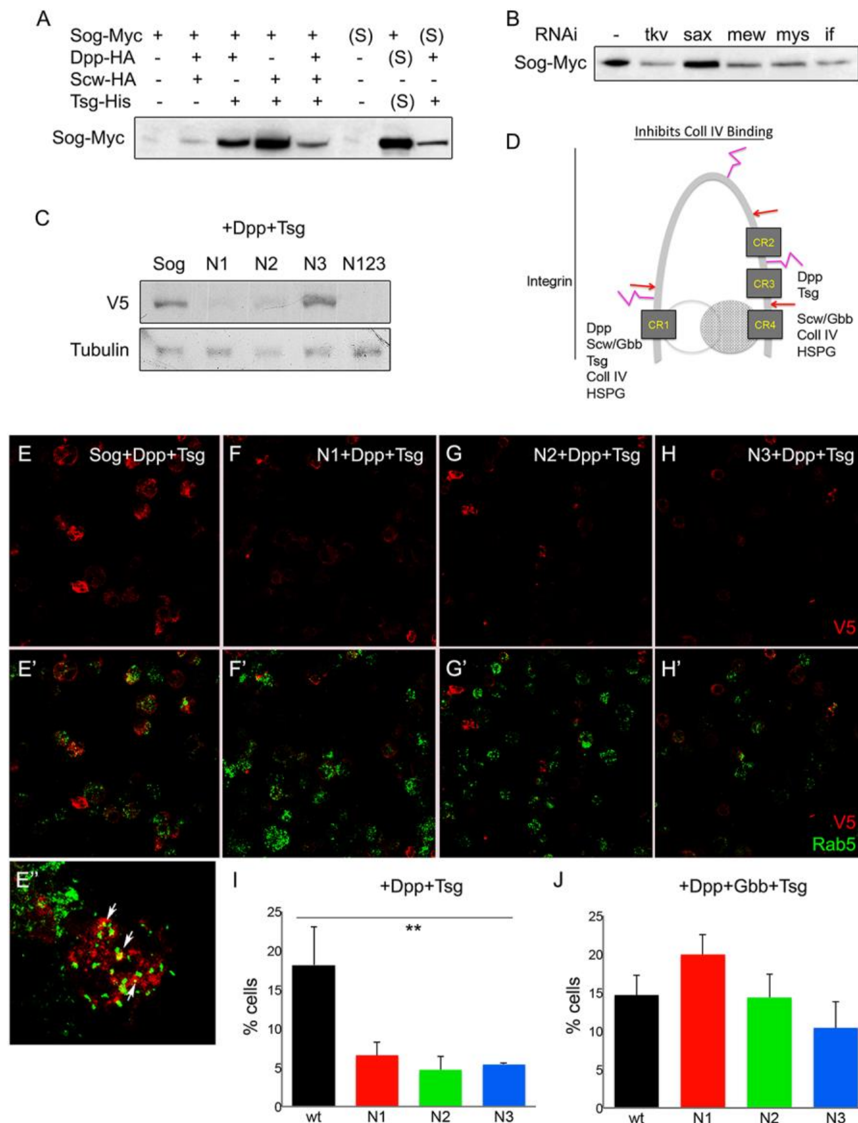

Figure 6. Sog levels are controlled by glycan and receptor-based retrieval from the extracellular space. A) S2 cells transfected with wild type *sog-myc* construct and different combinations of constructs producing Tsg and BMP ligands. Cellular amounts of Sog are detected with anti-myc antiserum. Sog amounts increase in the presence of Dpp plus Tsg. These cellular levels correspond to Sog retrieved from the medium, since equivalent intracellular amounts are observed when medium (S) from construct-expressing cells is added to cells that do not express *sog-myc.* B) Intracellular Sog levels retrieved from Sog-myc+Dpp-HA+Tsg-His medium decrease by knocking down the Dpp receptor *tkv* or αPS1 (*mew*), αPS2 (*if*) or βPS (*mys*) integrin expression. C) S2 cells transfected with wild type *sog-V5* or *sogN1-V5*, *sogN2-V5*, *sogN3-V5* or tripple *sogN123-V5* mutant constructs in the presence of

Development • Accepted manuscript

constructs producing Tsg and Dpp. Sog retrieval decreases by mutating the N1 and N2 sites. D) Predicted conformation of Sog protein with regions depicted for binding to established extracellular partners. E-J) Medium from S2 cells transfected with *dpp-HA* and *tsg-His* plus wild type *sog-V5* or *sogN1-V5*, *sogN2-V5*, or *sogN3-V5*, was added to naive cells and imunolabelled with anti-V5 (red) and anti-Rab5 (green, E'-H). E'') High magnification for wild type Sog-V5 and Rab5 shows partial co-localization, indicative of Sog endocytosis. These cells were quantified in (I), showing that loss of N1, N2 and N3 glycosylation sites decreases the number of cells with V5+Rab5+ endocytic punctae. J) In cells expressing *dpp-HA, gbb-His* and *tsg-His* plus wild type *sog-V5* or *sogN1-V5*, *sogN2-V5*, or *sogN3-V5,* the number of cells with V5+Rab5+ punctae is equivalent. Statistically significant differences based on Student's t-test, $^{**}P \leq 0.01$.

Development • Accepted manuscript

Supplementary Figures

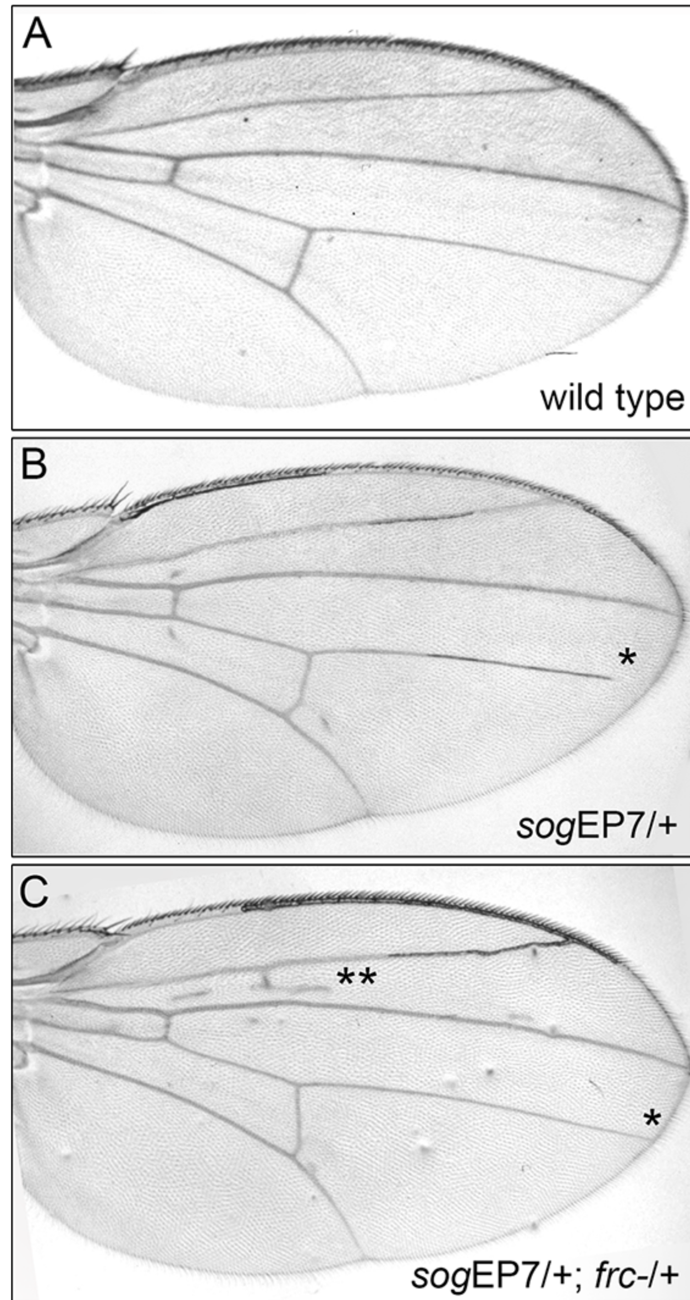

Figure S1. *frc* modifies the loss-of-vein phenotype of an enhancer piracy *sog* line. (A) Wild type wing; (B) Enhancer piracy sog-EP7 wing (Yu et al, 1996) with longitudinal vein loss and extra crossvein material. C) EP7/+; *frc*-/+ adult wing with a reduction in vein loss and extra veins throughout the wing.

Development • Supplementary information

Figure. S2. Protein alignment for *Drosophila sp.* Short gastrulation and for vertebrate Chordin homologs. A) Protein alignment for vertebrate Chordin putative glycosylation sites. B) Full protein alignment for Sog from the 12 sequenced Drosophila species.

	N1 (217)	N2 (351)	N3 (365)	N4 (434)
Human	DSNGSVL	QANVSAQ	LPNLTVQ	LGNGSLIY
Mouse	DPTGNIL	QANTSAQ	LPSLTDQ	LGNGSLIY
Chicken	DPEGTVL	HANITME	LSDLSAH	HENGTLEY
Fish	DTDGNTA	RANITAD	LADLNSR	HPNGSLDY
Frog	DLDGSVL	YANISAQ	LPDLSSR	HENGTLEY

Figure S2A. Predicted glycosylation sites for vertebrate Chordins. All predicted glycosylation sites all located in the stem region, between CR1 and CR2. An additional sites is predicted after CR3 in *Xenopus laevis* (frog) and *Gallus gallus* (chicken).

Figure S2B. Aligment of full-length Sog from Drosophila species. The CR domains are highligthed in yellow. A pink N (Arg) marks the predicted glycosylated residue and the accompanying sequence recognition motif highlighted in pink. Red indicates sequences target for cleavage by metaloproteases.

```
sog-PA Dmel      1 MANKLRK--------SNAIEWATATGTVPLLER-----SCCHSEDA------ALEPQASK  41
sog-PB Dsim      1 MANKLRK--------SNAIEWATATGTVPLLER-----SCCHSADA------PLEPQANR  41
sog-PA Dsec      1 MANKLRK--------SNAIEWATATGTVPLLER-----SCCHSEDA------PLEPQANR  41
sog-PA Dyak      1 MANKLRK--------SSGIEWATATGTVPLLER-----SCCHSEDA------QMEPQACR  41
sog-PA Dere      1 MANKLRK--------STGIEWATATGTVPLLER-----SCCHSEDADADE--EMEPQASR  45
sog-PA Dpers     1 MANKQRPSVSGSGSISGGVAWATATGTVPNLESSNSRKSCTDSTSTSASASTSTSRNSSN  60
sog-PA Dwill     1 ------------------------------------------------------------   0
sog-PA Dmoj      1 MNNNQTN---------IATALATATGTVQTLER-----HANISCSNCCN-TSSNSNSNCN  45
sog-PB Dvril     1 MNNKQTN---------AATAWATATGTVQLLDR-----PAVAGAAATDSSTACNSSSSSS  46
sog-PA Dgrim     1 MINNSTN---------VATALATATGTVQTLDR-----GQQQQQQQPGAAANSTATYGRG  46

sog-PA Dmel     42 TS-------HREQAPILRHLSQLSH---------LLIIAGLLIVCLAGVTEGRRHAPLMF  85
sog-PB Dsim     42 TS-------HREQAPILRHLSQLSH---------LLIIAGLLIVCLAGVTEGRRHAPLMF  85
sog-PA Dsec     42 TS-------HREQAPILRHLSQLSH---------LLIIAGLLIVCLAGVTEGRRHAPLMF  85
sog-PA Dyak     42 TTTSHRVPATREQSPILRHLSQRSQRSHLSHL--LIIIAGLLIVCLAGVTEGRRHAPLMF  99
sog-PA Dere     46 TS--HRVPTAREQPPILRRLSQLSQLRRLSHLSHLLIIAGLLIVCFAGVTEGRRHAPLMF 103
sog-PA Dpers    61 SSTSTSTSSSTTSASASASASQSWPMGRLSHQLRHLLIVGLLLCLTGGLVEGRRHAPLMF 120
sog-PA Dwill     1 ----------------------------------------------------------MF   2
sog-PA Dmoj     46 NSSSSSSNNSNTNNATYYYQRRLQHLLLLA----LLVCVASLNSCWLQPVQARRHAPLMF 101
sog-PB Dvril    47 SNNNCNSCTCNITGATYRYHRRLQQLLLVA----LLLCLSSLSSCWPQAVQARRHAPLMF 102
sog-PA Dgrim    47 ISSDCSS-TSNSGIPAYHYHRKLQHLLLVA----LLICLSSA-CCWPQAVQARRHAPLMF 100
                                                                             **

sog-PA Dmel     86 EESDTGRRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAIPKKRR 141
sog-PB Dsim     86 EESDTGRRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAIPKKRR 141
sog-PA Dsec     86 EESDTGRRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAIPKKRR 141
sog-PA Dyak    100 EESDTGRRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAIPKKRR 155
sog-PA Dere    104 EESDTGRRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAIPKKRR 159
sog-PA Dpers   121 EESDTGRRSNRPAGEGKVTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAIPKKRR 180
sog-PA Dwill     3 EESDTGRRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAMPKKRR  58
sog-PA Dmoj    102 EEADTARRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVSIPKKRR 157
sog-PB Dvril   103 EEADTARRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAIPKKRR 158
sog-PA Dgrim   101 EEADTARRSNRP-----------------------ADVYGPTPVCVCV-CVCVCIPKKRR 136
                   **.** ******                       **.  *  *  *. * **..*****

sog-PA Dmel    142 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPNEEEERNMKHY 201
sog-PB Dsim    142 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPNEEEERNMKHY 201
sog-PA Dsec    142 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPNEEEERNMKHY 201
sog-PA Dyak    156 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPNEEEERNMKHY 215
sog-PA Dere    160 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPNEEEERNMKHY 219
sog-PA Dpers   181 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPNEEEERNMKHY 240
sog-PA Dwill    59 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVASDVPVPSEEEERNMKHY 118
sog-PA Dmoj    158 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDAASDVPVPSEEEERNMKHY 217
sog-PB Dvril   159 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPSEEEERNMKHY 218
sog-PA Dgrim   137 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPSEEEERNMKHY 196
                   ***************************************** * ***** **********
```

Development • Supplementary information

Development: doi:10.1242/dev.167338: Supplementary information

```
sog-PA Dmel     202 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSSRIGRPRAIQF  261
sog-PB Dsim     202 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSSRIGRPRAIQF  261
sog-PA Dsec     202 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSSRIGRPRAIQF  261
sog-PA Dyak     216 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSSRIGRPRAIQF  275
sog-PA Dere     220 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSSRIGRPRAIQF  279
sog-PA Dpers    241 AALLTGRTSYFLKGEEMKSMYTTYNPHNIVATARFLFHKKNLYYSFYTSSRIGRPRAIQF  300
sog-PA Dwill    119 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSSKVGRPRAIQF  178
sog-PA Dmoj     218 AALLTGRTSYFLKGEEMKSMYTTYNPQNIVATARFLFHKKNLYYSFYTSTKIGRPRAIQF  277
sog-PB Dvril    219 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSSRIGRPRAIQF  278
sog-PA Dgrim    197 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSARIGRPRAIQF  256
                    **************************.*.********************...********

sog-PA Dmel     262 VDDAGVILEEHQLETTLAGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  321
sog-PB Dsim     262 VDDAGVILEEHQLETTLAGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  321
sog-PA Dsec     262 VDDAGVILEEHQLETTLAGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  321
sog-PA Dyak     276 VDDAGVILEEHQLETTLAGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  335
sog-PA Dere     280 VDDAGVILEEHQLETTLAGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  339
sog-PA Dpers    301 VDDAGVILEEHQLETTLTGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  360
sog-PA Dwill    179 VDDGGVILEEHQLETTLTGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  238
sog-PA Dmoj     278 VDDRGDILEEHQLETTLSGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  337
sog-PB Dvril    279 VDDAGVILEEHQLETTLSGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  338
sog-PA Dgrim    257 VDDAGVILEEHQLETTLSGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  316
                    *** * ***********.******************************************

sog-PA Dmel     322 QQAELALAGKVAKYTALQTELFSSLLEAP--LPDGKTDPQLAGAGGTAIVSTSSGAASSM  379
sog-PB Dsim     322 QQAELALAGKVAKYTALQTELFSSLLEAP--LPDGKTDPQLAGAGGTAIVSTSSGAASSM  379
sog-PA Dsec     322 QQAELALAGKVAKYTALQTELFSSLLEAP--LPDGKTDPQLAGAGGTAIVSTSSGAASSM  379
sog-PA Dyak     336 QQAELALAGKVAKYTALQTELFSSLLEAP--LPDGKTDPQLAGAGGTAIVSTSSGAASSM  393
sog-PA Dere     340 QQAELALAGKVAKYTALQTELFSSLLEAP--LPDGKTDPQLAGAGGTAIVSTSSGAASSM  397
sog-PA Dpers    361 HQSELALAGKIAKYTALQTELFSSLLEPPQSQPGGKFDPMLTGAGGTAIVSTSSGAASSM  420
sog-PA Dwill    239 HQSELALAGKIAKYTALQSELFSSLLEPS-----NIQDPQLAGAGGTAIVSTSSGVASSM  293
sog-PA Dmoj     338 HQAELALAGKIAKYTALQTELFSSLLEP--AATNGKTDPQLAGAGGTAIVSTSSGAASSM  395
sog-PB Dvril    339 HQAELALAGKIAKYTALQTELFSSLLEPP-ALVGGKPDPQLAGAGGTAIVSTSSGAASSM  397
sog-PA Dgrim    317 HQSELALAGKIAKYTALQTELFSSLLEPP-LTASGKPDPQLAGAGGTAIVSTSSGAASSM  375
                      .*.*******.*******.********       .  ** *.************* ****

sog-PA Dmel     380 HLTLVFNGVFGAEEYADAALSVKIELAERKEVIFDEIPRVRKPSAEINVLELSSPISIQN  439
sog-PB Dsim     380 HLTLVFNGVFGAEEYADAALSVKIELAERKEVIFDEIPRVRKPSAEINVLELSSPISIQN  439
sog-PA Dsec     380 HLTLVFNGVFGAEEYADAALSVKIELAERKEVIFDEIPRVRKPSAEINVLELSSPISIQN  439
sog-PA Dyak     394 HLTLVFNGVFGAEEYADAALSVKIELAERKEVIFDEIPRVRKPSAEINVLELSSPISIQN  453
sog-PA Dere     398 HLTLVFNGVFGAEEYADAALSVKIELAERKEVIFDEIPRVRKPSAEINVLELSSPISIQN  457
sog-PA Dpers    421 HLTLVFNGIFGAEEYADAALSVKIELPERKEIIFDEVPRVRKPSAEINVLELSSPISIQN  480
sog-PA Dwill    294 HLTLIFNGVFGSEEFADAALSVRIELPERKEMIFDEVPRVRKPSAEINVLELSSPISIQN  353
sog-PA Dmoj     396 HLTLVYNGIFGVEEYADAGLSVRIELPERKELILDEVSRVRKPSAEINVLELSSPISIQN  455
sog-PB Dvril    398 HLTLVFNGIFGAEEYADAALSVRIELPERKELIFDEVPRVRKPSAEINVLELSSPISIQN  457
sog-PA Dgrim    376 HLTLVFNGIFGTEEYADSTLSVRIELPERKEVIFDELSRVRKPSAEINVLELSSPISIQN  435
                    ****..**.** **.**. ***.*** ****.* **. **********************
```

Development • Supplementary information

```
sog-PA Dmel     440 LRLMSRGKLLLTVESKKYPHLRIQGHIVTRASCEIFQTLLAP--HSA----------ESS 487
sog-PB Dsim     440 LRLMSRGKLLLTVESKKYPHLRIQGHIVTRASCEIFQTLLAP--HSA----------ESS 487
sog-PA Dsec     440 LRLMSRGKLLLTVESKKYPHLRIQGHIVTRASCEIFQTLLAP--HSA----------ESS 487
sog-PA Dyak     454 LRLMSRGKLLLTVESKKYPHLRIQGHIVTRASCEIFQTLLAP--HSA----------ESS 501
sog-PA Dere     458 LRLMSRGKLLLTVESKKYPHLRIQGHIVTRASCEIFQTLLAP--HSA----------ESS 505
sog-PA Dpers    481 LRLMSRGKLLLTVESKKYPQLRIQGHIVTRASCEIFQTLLAP--HNG----------ESA 528
sog-PA Dwill    354 LRLMSRGKLLLIVESKKYPQLRIQGHIVTRASCEIFQTLLAP--HASGSGSTETTALASN 411
sog-PA Dmoj     456 LRLMSRGKLLLTVESKKYPQLRIQGHIVTRASCEIFQTLLAP-SHGN---------ELS- 504
sog-PB Dvril    458 LRLMSRGKLLLTVESKKYPQLRIQGHIVTRASCEIFQTLLAP-PHGN---------ELS- 506
sog-PA Dgrim    436 LRLMSRGKLLLTVESKKYPQLRIQGHIVTRASCEIFQTLLAPPMHGG---------ELST 486
                    *********** *******.**********************  *             * 

sog-PA Dmel     488 TKSSGLAWVYLNTDGSLAYNIETEHVNTRDRPNISLIEEQGKRKAKLEDLTPSFNFNQAI 547
sog-PB Dsim     488 TKSSGLAWVYLNTDGSLAYNIETEHVNTRDRPNISLIEEQGKRKAKLEDLTPSFNFNQAI 547
sog-PA Dsec     488 TKSSGLAWVYLNTDGSLAYNIETEHVNTRDRPNISLIEEQGKRKAKLEDLTPSFNFNQAI 547
sog-PA Dyak     502 TKSSGLAWVYLNTDGSLAYNIETEHVNTRDRPNISLIEEQGKRKAKLEDLTPSFNFNQAI 561
sog-PA Dere     506 TKSSGLAWVYLNTDGSLAYNIETEHVNTRDRPNISLIEEQGKRKAKLEDLTPSFNFNQAI 565
sog-PA Dpers    529 TRSSGLAWVYLNTDGSLAYNIETDHVNTRDRPNISLVEEQGKRKAKLEDLTPSFNFNQAI 588
sog-PA Dwill    412 SRSSGLAWVYLNTDGSLAYNIETDHVNTKDRPNISLIEE-GKRKAKLEDLTPSFNFNQAI 470
sog-PA Dmoj     505 TRSSGLAWVYLNTDGSLAYNIETDHVNTRDRPNISLIEEQGNRKAKLEDLTPSFNFNQAI 564
sog-PB Dvril    507 TRSSGLAWVYLNTDGSLAYNIETDHVNTRDRPNISLIEEQGKRKAKLEDLTPSFNFNQAI 566
sog-PA Dgrim    487 TRSSGLAWVYLNTDGSLAYNIETDHVNTRDRPNISLIEEQGKRKAKLEDLTPSFNFNQAI 546
                    ..*********************.****.*******.** * ******************

sog-PA Dmel     548 GSVEKLGPKVLESLYAGELGVNVATEHETSLIRGRLVPRPVADARDSAEPILLKRQ-EHT 606
sog-PB Dsim     548 GSVEKLGPKVLESLYAGELGVNVATEHETSLIRGRLVPRPVADARDSAEPILLKRQ-EHS 606
sog-PA Dsec     548 GSVEKLGPKVLESLYAGELGVNVATEHETSLIRGRLVPRPVADARDSAEPILLKRQ-EHS 606
sog-PA Dyak     562 GSVEKLGPKVLESLYAGELGVNVATEHETSLIRGRLVPRPVADARDSAEPILLKRQ-EHT 620
sog-PA Dere     566 GSVEKLGPKVLESLYAGELGVNVATEHETSLIRGRLVPRPVADARDSAEPILLKRQ-EHP 624
sog-PA Dpers    589 GSVEKLGPKVLESLYAGELGVNVGTEHETSLIRGRLVPRPVADARDSAEPILLKRQ-EHL 647
sog-PA Dwill    471 GSVEKLGPKVLESLYDGELVVNVATEHDTSLIRGRLVPRPVADARDSAEPILLKRP-EHS 529
sog-PA Dmoj     565 GSVEKLGPKVLESLYAGELGVNVATDHEASLIRGRLVPRPVADARDSAEPILLKAHVEHG 624
sog-PB Dvril    567 GSVEKLGPKVLESLYAGELGVNVATEHEASLIRGRLVPRPVADARDSAEPILLKPHGEHA 626
sog-PA Dgrim    547 GTVEKLGPKVLESLYAGELGVNVATEHEASLIRGRLVPRPVADARDSAEPILLKQHMEHG 606
                    *.************* *** *** *.*..*************************   ** 

sog-PA Dmel     607 -------DAQNPHAVGMAWMSIDNECNLHYEVTLNGVPA--QDLQLYLEEKPIEAIGAPV 657
sog-PB Dsim     607 -------DAQNPHAVGMAWMSIDNECNLHYEVTLNGVPA--QDLQLYLEEKPIEAIGAPV 657
sog-PA Dsec     607 -------DAQNPHAVGMAWMSIDNECNLHYEVTLNGVPA--QDLQLYLEEKPIEAIGAPV 657
sog-PA Dyak     621 Q------DAQNPHAVGMAWMSIDNECNLHYEVTLNGVPA--QDLQLYLEEKPIEAIGAPV 672
sog-PA Dere     625 -------DAQNPHAVGMAWMSIDNECNLHYEVTLNGVPA--QDLQLYLEEKPIEAIGAPV 675
sog-PA Dpers    648 AA---SGSNQSGHAMGMAWMSIDNECNLHYELTLSGVPP--QDLQLYLEEKPIEAIGAPV 702
sog-PA Dwill    530 P------SVGGEHSMGMAWMSIDNECNLHYELTLSGLPS--QDMQLYLEEKPIEAIGAPV 581
sog-PA Dmoj     625 QN--------QGQGMGMAWMAIDNECNLHYELTLSGVPN--QELQLYLEEKPIEAIGAPV 674
sog-PB Dvril    627 S---------PGLGMGMAWMSIDNECNLHYELTLSGMPNAGQDLQLYLEEKPIEAIGAPV 677
sog-PA Dgrim    607 SSASASASASAGHSMGMAWMSIDNECNLHYEITLSGVPT--QELQLYLEEKPIEAIGAPV 664
                                  .*****.**********.** *.*   *..****************
```

Development • Supplementary information

```
sog-PA Dmel     658 TRKLLEEFNGSYLEGFFLSMPSAELIKLEMSVCYLEVHSKHSKQLLLRGKLKSTKVPGHC  717
sog-PB Dsim     658 TRKLLEEFNGSYLEGFFLSMPSAELIKLEMSVCYLEVHSKHSKQLLLRGKLKSTKVPGHC  717
sog-PA Dsec     658 TRKLLEEFNGSYLEGFFLSMPSAELIKLEMSVCYLEVHSKLSKQLLLRGKLKSTKVPGHC  717
sog-PA Dyak     673 TRKLLEEFNGSYLEGFFLSMPSAELIKLEMSVCYLEVHAKHSKQLLLRGKLKSTKVPGHC  732
sog-PA Dere     676 TRKLLEEFNGSYLEGFFLSMPSAELIKLEMSVCYLEVHSKHSKQLLLRGKLKSTKVPGHC  735
sog-PA Dpers    703 TRKLLEEFNGSYLEGFFLSMPSAELIKLEMSVCYLELHSKHSKQLLLRGKLKSTKVPAHC  762
sog-PA Dwill    582 TRKLLEEFNSSYMEGFGLGMPSAELIKLETSVCYLELQSKHSKQLLLRGKLKSTKVPAHC  641
sog-PA Dmoj     675 TRKLLKEFTGNYMEGFQLSMPSAELIKLETSVCYLELHSKVNPRLLLRGKLKSTKVPAHC  734
sog-PB Dvril    678 TRKLLEEFNGAYMEGFVLGMPSAELIKLETSVCYLELHAKHSHQLLLRGKLKSTKVPAHC  737
sog-PA Dgrim    665 TRKLLEEFNGQYMEGFLLGMPSAELIKLETSVCYLELHAKHSHQLLLRGKLKSTKVPAHC  724
                    ***** **.  *.*** * ********** ******...*   .************* **

sog-PA Dmel     718 FPVYTDNNVPVPGDHNDNHLVNGETKCFHSGRFYNESEQWRSAQDSCQMCACLRGQSSCE  777
sog-PB Dsim     718 FPVYTDNNVPVPGDHNDNHLVNGETKCFHSGRFYNESEQWRSAQDSCQMCACLRGQSSCE  777
sog-PA Dsec     718 FPVYTDNNVPVPGDHNDNHLVNGETKCFHSGRFYNESEQWRSAQDSCQMCACLRGQSSCE  777
sog-PA Dyak     733 FPVYTDNNVPVPGDHNDNHLVNGETKCFHSGRFYNESEQWRSAQDACQMCACLRGQSNCE  792
sog-PA Dere     736 FPVYTDNNVPVPGDHNDNHLVNGETKCFHSGRFYNESEQWRSAQDSCQMCACLRGQSSCE  795
sog-PA Dpers    763 FPVYTDNNVPVPGDHNDNHLINAETKCFHSGRFYNESEQWRSAQDACQMCACLRGQSNCE  822
sog-PA Dwill    642 FPIYTDNNVPVPGDQNDNHLINTETKCFHSGRFYNESEQWRSAQDTCQMCACLRGQANCE  701
sog-PA Dmoj     735 FPIYTDNNVPVPGDQNGNQMLSTESNCYHSGRFYNESEQWRSAQDTCHMCACQRGQANCE  794
sog-PB Dvril    738 FPIYTDNNVPVPGDHNDNQMLTAESKCFHSGRFYNESEQWRSAQDTCHMCACQRGQANCE  797
sog-PA Dgrim    725 FPIYTDNNVPVPGDHNDNQMLTAETKCFHSGRFYNESEQWRSAQDTCHMCACQRGQSNCE  784
                    **.***********.* *...  *. *.*****************.*.**** ***. **

sog-PA Dmel     778 VIKCPALKCKST-EQLLQRDGECCPSCVPKK---------EAADYSAQ------------  815
sog-PB Dsim     778 VIKCPALKCKST-EQLLQREGECCPSCVPKK---------EAADYSAQ------------  815
sog-PA Dsec     778 VIKCPALKCKST-EQLLQREGECCPSCVPKK---------EAADYSAQ------------  815
sog-PA Dyak     793 VIKCPALKCKAT-EQLLQREGECCPSCVPKK---------EAADYSAQ------------  830
sog-PA Dere     796 VIKCPALKCKST-EQLLQREGECCPICVSKK---------EAADYSSQ------------  833
sog-PA Dpers    823 PQRS----------------------SVSAS---------ASASASS-------------  838
sog-PA Dwill    702 IIKCPALKCKS-------------PNEQLIQ---------REGECCPS------------  727
sog-PA Dmoj     795 LIKCPALKCKAG-EQLLQREGECCPSCVAK----PADWAQAQGQLSSS------------  837
sog-PB Dvril    798 LIKCPALKCKTGSEQLLQRDGECCPSCVARR---DADWHTAGSSSSSSSSSSTGSGSGS  854
sog-PA Dgrim    785 PIKCPALKCKPLTEQLLQREGECCPTCVARRGENGADWHMAAGATTSSTS---------I  835
                      ..                                         .              

sog-PA Dmel     816 -SSPATNATDLLQQ--RRGCRLGEQFHPAGASWHPFLPPNGFDTCTTCSCDPLTLEIRCP  872
sog-PB Dsim     816 -SSPATNATDLLQQ--RRGCRLGEQFHPAGASWHPFLPPNGFDTCTTCSCDPLTLEIRCP  872
sog-PA Dsec     816 -SSPATNATDLLQQ--RRGCRLGEQFHPAGASWHPFLPPNGFDTCTTCSCDPLTLEIRCP  872
sog-PA Dyak     831 -SSPATNASDLLQQ--RRGCRLGEQFHPAGASWHPFLPPNGFDTCTTCSCDPLTLEIRCP  887
sog-PA Dere     834 -SSPATNATDLLQQ--RRGCRLGEQFHLAGASWHPFLPPNGFDTCTTCSCDPLTLEIRCP  890
sog-PA Dpers    839 -SLPAVNGSDLLLQQQRRGCRLGDQFHAAGASWHPFLPPNGFDTCTTCSCDPLTLEIRCP  897
sog-PA Dwill    728 -CMPKREPTALASES-GRGCRLGDQFHPAGASWHPFLPPNGFDTCTTCSCDPLTLEVRCP  785
sog-PA Dmoj     838 ---PAINASDLLQQ--RRGCRLGEQFHAAGASWHPFLPPNGFDTCTTCSCDAQTLEVRCP  892
sog-PB Dvril    855 GSSPALNASDVLQQ--RRGCRLGEQFHAAGASWHPFLPPNGFDTCTTCSCDALTLEVRCP  912
sog-PA Dgrim    836 AAHSNNGSSDLLQP--HRGCRLGDQFHAAGASWHPFLPPNGFDTCTTCSCDILTLEVRCP  893
                            . .      ******.*** ***********************  ***.***
```

Development • Supplementary information

```
sog-PA Dmel     873 RLVCPPLQCSEKLAYRPDKKACCKICPEGKQSSS--NGHKTTPNNPNVLQDQAMQRSPSH  930
sog-PB Dsim     873 RLVCPPLQCSEKLAYRPDKKACCKICPEGKQSSS--NGHKTAPNNPNVLQDQAMQRSPSH  930
sog-PA Dsec     873 RLVCPPLQCSEKLAYRPDKKACCKICPEVKQSSS--NGHKTAPNNPNVLQDQAMQRSPSH  930
sog-PA Dyak     888 RLVCPPLQCSEKLAYRPDKKACCKICPEGKQSSS--NGHKAAPNNPNVLQDQAMQRSPSH  945
sog-PA Dere     891 RLVCPPLQCSEKLAYRPDKKACCKICPEGKQSSS--NGHKAAPNNPNVLQDQAMQRSPSH  948
sog-PA Dpers    898 RLVCPPLQCSEKLAYRPDKKACCKVCPEGKQSSP--AGGKAAPNNPNVLQDQAMQRSPQH  955
sog-PA Dwill    786 RMVCPPLQCSEKLAYRPDKKACCKVCPEGKQSSS--NSDKSMPNNPNVLHDQAMPRTPGH  843
sog-PA Dmoj     893 RMVCPPLQCSEKLAYRPDKKACCKVCPEGKQSNHGASSKHQAPANHNELHDQAVQRTAAH  952
sog-PB Dvril    913 RMVCPPLQCSEKLAYRPDKKACCKVCPEGKQSNHGQSGKH-TPSNPNVLLDQAVQRTAAH  971
sog-PA Dgrim    894 RMVCPPLQCIEKLAFRPDKKACCKVCPEGKQSTHGQSSKH-TPSNPNVLHDQAMQHSAAH  952
                    *.******* ****.*********.*** ***       .  * * * * ***. ..  *

sog-PA Dmel     931 SAEEVLANGGCKVVNKVYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ  990
sog-PB Dsim     931 SAEEVLANGGCKVVNKVYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ  990
sog-PA Dsec     931 SAEEVLANGGCKVVNKVYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ  990
sog-PA Dyak     946 SAEEVLANGGCKVVNKVYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ 1005
sog-PA Dere     949 SAEEVLANGGCKVVNKVYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ 1008
sog-PA Dpers    956 SAEDVLAAGGCKVVNKVYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ 1015
sog-PA Dwill    844 TAEEILAAGGCKVVNKIYENGREWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ  903
sog-PA Dmoj     953 NAEEVLAAGGCKVVNKVYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ 1012
sog-PB Dvril    972 NAEEVLAAGGCKVVNKIYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ 1031
sog-PA Dgrim    953 TAEEVLNAGGCKVVNKIYENGQEWHPVLMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ 1012
                     **..*  ********.****.****.*********************************

sog-PA Dmel     991 Q---TRVTSKRRLFEKP---------DAAAPAIDECCSTQCRRSRRHHKRQPHHQQRSSS 1038
sog-PB Dsim     991 Q---TRVTSKRRLFEKP---------DAAAPAIDECCSTQCRRSRRHHKRQPHHQQRSSS 1038
sog-PA Dsec     991 Q---TRVTSKRRLFEKP---------DAAAPAIDECCSTQCRRSRRHHKRQPHHQQRSSS 1038
sog-PA Dyak    1006 Q---TRVTSKRRLFEKP---------DAAAPAIDECCSTQCRRSRRHHKRQPHHQQRSSS 1053
sog-PA Dere    1009 Q---TRVTSKRRLFEKP---------DAAAPAIDECCSTQCRRSRRHHKRQPHHQQRSSS 1056
sog-PA Dpers   1016 Q---TRVSSKRRIFEKP---------DS-APTLDECCSTQCRRSRRHHKRQPHHQQRSNA 1062
sog-PA Dwill    904 ----TRVSSKRRPFEKPGN-----GNDGSGAPIDECCSTQCRRSRRHHKRQPHHQ-----  949
sog-PA Dmoj    1013 ----TRVSSKRRLFDKQTGVAEL--VGGVAQPIDECCSTQCRRSRRHHKRQPHHQQRATS 1066
sog-PB Dvril   1032 QQQQTRVSGKRRLFDKPTAGSELGGVAGAAAPIDECCSTQCRRSRRHHKRQPHHQQRAAS 1091
sog-PA Dgrim   1013 ----TRVSGKRRLFDKPGSEAATG-AGGGPQPIDECCSTQCRRSRRHHKRQPHHQQRAAS 1067
                        ***. *** *.*                .**********************

sog-PA Dmel    1039 ----- 1038
sog-PB Dsim    1039 ----- 1038
sog-PA Dsec    1039 ----- 1038
sog-PA Dyak    1054 ----- 1053
sog-PA Dere    1057 ----- 1056
sog-PA Dpers   1063 AASSS 1067
sog-PA Dwill    950 -----  949
sog-PA Dmoj    1067 ----- 1066
sog-PB Dvril   1092 S---- 1092
sog-PA Dgrim   1068 S---- 1068
```

Development • Supplementary information

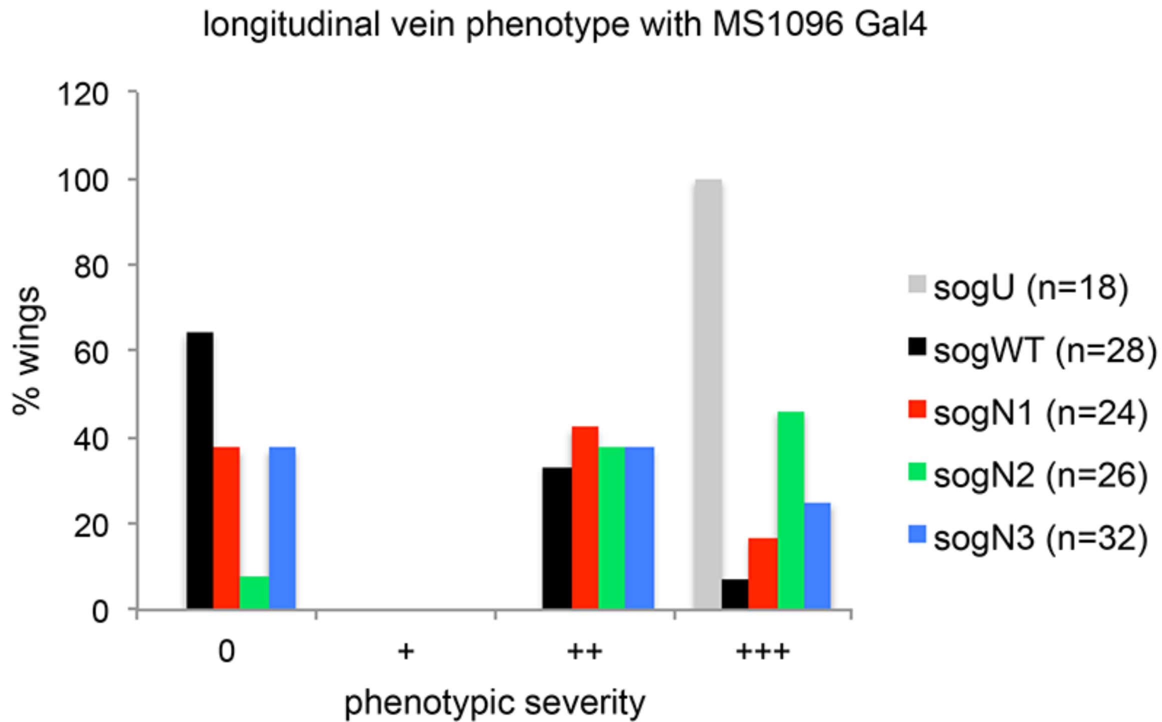

Figure S3. Ubiquitous expression of *sogN* mutants leads to vein truncation. Longitudinal vein truncation phenotypes resulting from ubiquitous overexpression of wild type *sog* and *sogN* mutants in the wing primordia, driven by MS1096-Gal4. The uncleavable *sog-u* mutant (Peluso et al, 2011) was also included in the analysis.

Development • Supplementary information

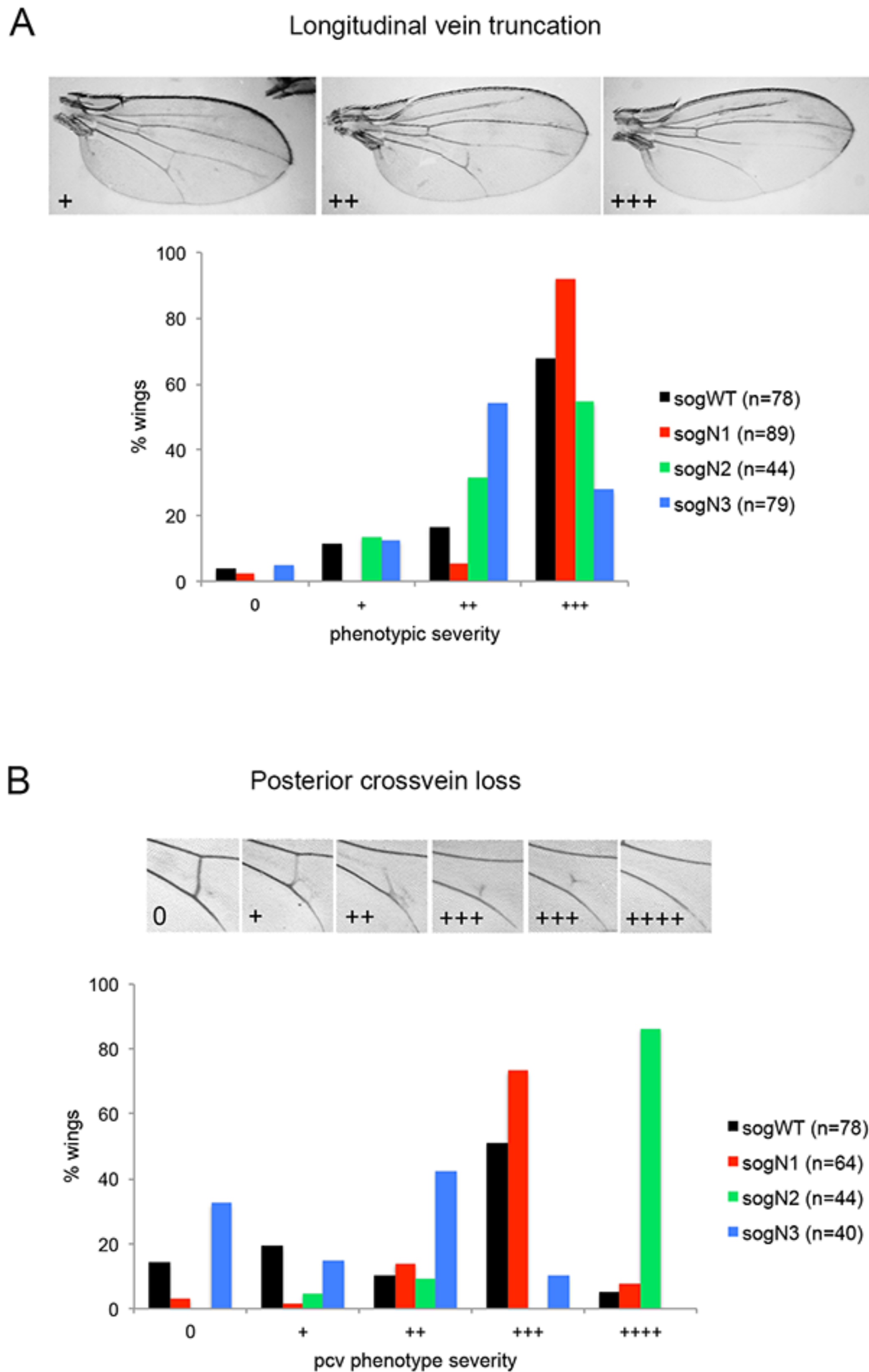

Fig. S4. Longitudinal vein expression of *sogN* mutants increases a *dpp* loss-of-function phenotype. Longitudinal vein truncation phenotypes (A) or PCV phenotypes (B) resulting from

Development • Supplementary information

longtudinal vein overexpression of wild-type *sog* and *sogN* mutants in the wing primordia, driven by shv-Gal4. Expression was driven in a *dpp[shv]* heterozygous background, revealing that all constructs induce a strong vein loss phenotype. 100% of *dpp[shv]/*+ wings have a wild type (0) wing venation pattern.

Supplementary Material and Methods

Sequence accession numbers were: NP-476736 (D. *melanogaster*), FBpp0320361 (XP_016039421.1, *D. simulans*), FBpp0137934 (XP_001978067, *D. erecta*), FBpp0261045 (XP_0002100689, *D. yakuba*), FBpp0393205 (*D. virilis*), FBpp0170731 (XP_002010774, *D. mojavensis*), FBgn 0131639 (*D. grimshawi*), FBpp0254766 (XP_002071108, *D. willistoni*), FBp0180550 (XP_002022952, *D. persinilis*), FBpp0204020 (XP_002044153 (*D. sechellia*). Chordin sequences were downloaded from NCBI for *Mus musculus (NP_001264970 isoform 2 precursor and NP_034023 isoform 1 precursor), Homo sapiens (NP_001291401 isoform 2 precursor and NP_003732 isoform 1 precursor), Gallus gallus (NP_990311), Danio rerio (NP_571048) and Xenopus laevis (NP_001081778)*.

Primers used in this study for double-stranded RNA production and RNA interference assays (Fig. 6):

Gene	forward primer	reverse primer
tkvH	TAATACGACTCACTATAGGGAGAGAGAA GCTGCGCAAGC	TAATACGACTCACTATAGGGAGAGCTGGTTT GCCAGGGGT
tkvM	TAATACGACTCACTATAGGGAGAGAACC ATTGCCAAGCAGATTCAGAT	TAATACGACTCACTATAGGGAGATGAATGA CATCCAGTTCCGAGTTGT
saxH	TAATACGACTCACTATAGGGAGAGTGAA TGTGGTCTGCTGTG	TAATACGACTCACTATAGGGAGACTCCCG CTTCCAGGACT
saxM	TAATACGACTCACTATAGGGAGACGCGA TGCCGATGGTCAGGTGCAGGAG	TAATACGACTCACTATAGGGAGACCTCGT CCAATGCACTCGATCAGGG
mysH	TAATACGACTCACTATAGGGAGACCTCT TCGGTGGAGATGAA	TAATACGACTCACTATAGGGAGAGGATTTG GTCGCTTGTGG
mysM	TAATACGACTCACTATAGGGAGAATGAA GGACAACGCCACTGGAGATG	TAATACGACTCACTATAGGGAGACTCAATG GGATTTGGTCGCTTGTGG
mewH	TAATACGACTCACTATAGGGAGAGCAAG GAGCTGCATAACAT	TAATACGACTCACTATAGGGAGACGGCTC CACCCGGATAA
mewM	TAATACGACTCACTATAGGGAGAAACAC GGACTTGGATGATAACTCCT	TAATACGACTCACTATAGGGAGACAATATC GGATTGAGGCGTACCAAA
ifH	TAATACGACTCACTATAGGGAGACTGTT AATGCCGAAACCAG	TAATACGACTCACTATAGGGAGAATCTGG CTCGCAAATATTGT
ifM	TAATACGACTCACTATAGGGAGATGCGT GGACACAGCACACAGGGTCA	TAATACGACTCACTATAGGGAGACCTCATT CTCGCGGCCATAGTAGGG

Development • Supplementary information

Supplementary Figures

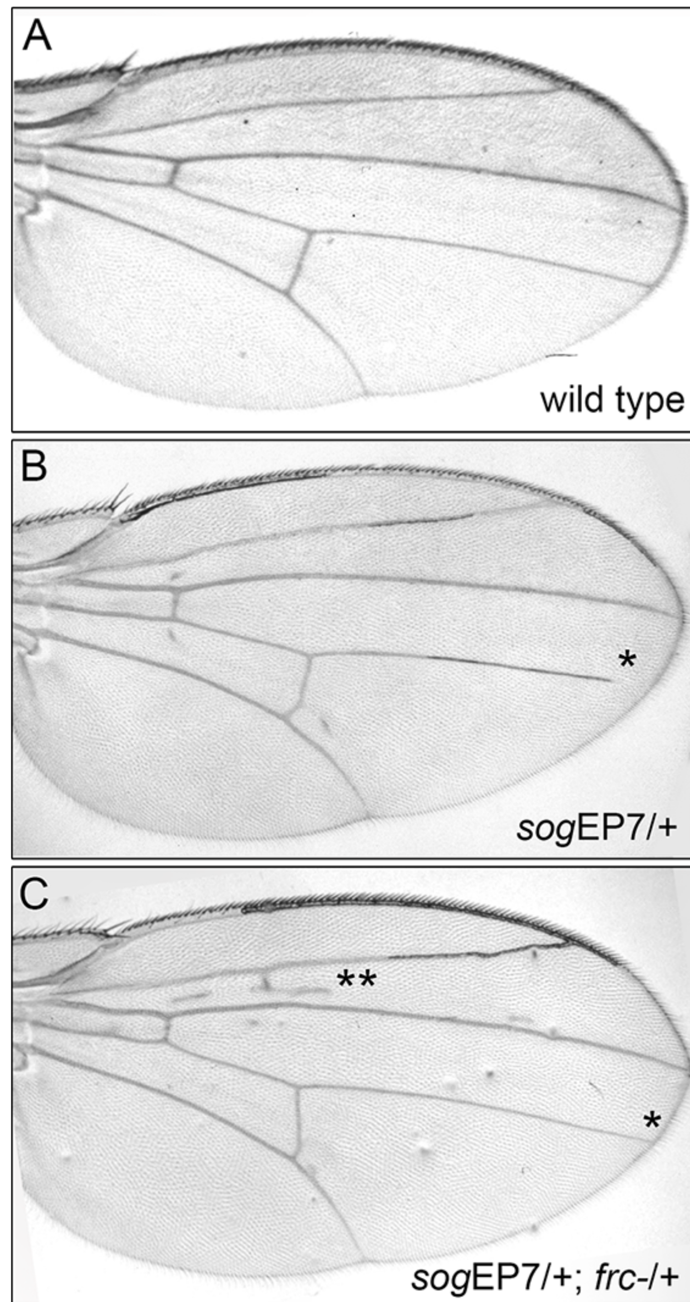

Figure S1. *frc* modifies the loss-of-vein phenotype of an enhancer piracy *sog* line. (A) Wild type wing; (B) Enhancer piracy sog-EP7 wing (Yu et al, 1996) with longitudinal vein loss and extra crossvein material. C) EP7/+; *frc*-/+ adult wing with a reduction in vein loss and extra veins throughout the wing.

Development • Supplementary information

Figure. S2. Protein alignment for *Drosophila sp.* Short gastrulation and for vertebrate Chordin homologs. A) Protein alignment for vertebrate Chordin putative glycosylation sites. B) Full protein alignment for Sog from the 12 sequenced Drosophila species.

	N1 (217)	N2 (351)	N3 (365)	N4 (434)
Human	DSNGSVL	QANVSAQ	LPNLTVQ	LGNGSLIY
Mouse	DPTGNIL	QANTSAQ	LPSLTDQ	LGNGSLIY
Chicken	DPEGTVL	HANITME	LSDLSAH	HENGTLEY
Fish	DTDGNTA	RANITAD	LADLNSR	HPNGSLDY
Frog	DLDGSVL	YANISAQ	LPDLSSR	HENGTLEY

Figure S2A. Predicted glycosylation sites for vertebrate Chordins. All predicted glycosylation sites all located in the stem region, between CR1 and CR2. An additional sites is predicted after CR3 in *Xenopus laevis* (frog) and *Gallus gallus* (chicken).

Development • Supplementary information

Figure S2B. Aligment of full-length Sog from Drosophila species. The CR domains are highligthed in yellow. A pink N (Arg) marks the predicted glycosylated residue and the accompanying sequence recognition motif highlighted in pink. Red indicates sequences target for cleavage by metaloproteases.

```
sog-PA Dmel      1 MANKLRK--------SNAIEWATATGTVPLLER-----SCCHSEDA------ALEPQASK  41
sog-PB Dsim      1 MANKLRK--------SNAIEWATATGTVPLLER-----SCCHSADA------PLEPQANR  41
sog-PA Dsec      1 MANKLRK--------SNAIEWATATGTVPLLER-----SCCHSEDA------PLEPQANR  41
sog-PA Dyak      1 MANKLRK--------SSGIEWATATGTVPLLER-----SCCHSEDA------QMEPQACR  41
sog-PA Dere      1 MANKLRK--------STGIEWATATGTVPLLER-----SCCHSEDADADE--EMEPQASR  45
sog-PA Dpers     1 MANKQRPSVSGSGSISGGVAWATATGTVPNLESSNSRKSCTDSTSTSASASTSTSRNSSN  60
sog-PA Dwill     1 ------------------------------------------------------------   0
sog-PA Dmoj      1 MNNNQTN---------IATALATATGTVQTLER-----HANISCSNCCN-TSSNSNSNCN  45
sog-PB Dvril     1 MNNKQTN---------AATAWATATGTVQLLDR-----PAVAGAAATDSSTACNSSSSSS  46
sog-PA Dgrim     1 MINNSTN---------VATALATATGTVQTLDR-----GQQQQQQQPGAAANSTATYGRG  46

sog-PA Dmel     42 TS-------HREQAPILRHLSQLSH---------LLIIAGLLIVCLAGVTEGRRHAPLMF  85
sog-PB Dsim     42 TS-------HREQAPILRHLSQLSH---------LLIIAGLLIVCLAGVTEGRRHAPLMF  85
sog-PA Dsec     42 TS-------HREQAPILRHLSQLSH---------LLIIAGLLIVCLAGVTEGRRHAPLMF  85
sog-PA Dyak     42 TTTSHRVPATREQSPILRHLSQRSQRSHLSHL--LIIIAGLLIVCLAGVTEGRRHAPLMF  99
sog-PA Dere     46 TS--HRVPTAREQPPILRRLSQLSQLRRLSHLSHLLIIAGLLIVCFAGVTEGRRHAPLMF 103
sog-PA Dpers    61 SSTSTSTSSSTTSASASASASQSWPMGRLSHQLRHLLIVGLLLCLTGGLVEGRRHAPLMF 120
sog-PA Dwill     1 ----------------------------------------------------------MF   2
sog-PA Dmoj     46 NSSSSSSNNSNTNNATYYYQRRLQHLLLLA----LLVCVASLNSCWLQPVQARRHAPLMF 101
sog-PB Dvril    47 SNNNCNSCTCNITGATYRYHRRLQQLLLVA----LLLCLSSLSSCWPQAVQARRHAPLMF 102
sog-PA Dgrim    47 ISSDCSS-TSNSGIPAYHYHRKLQHLLLVA----LLICLSSA-CCWPQAVQARRHAPLMF 100
                                                                             **

sog-PA Dmel     86 EESDTGRRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAIPKKRR 141
sog-PB Dsim     86 EESDTGRRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAIPKKRR 141
sog-PA Dsec     86 EESDTGRRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAIPKKRR 141
sog-PA Dyak    100 EESDTGRRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAIPKKRR 155
sog-PA Dere    104 EESDTGRRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAIPKKRR 159
sog-PA Dpers   121 EESDTGRRSNRPAGEGKVTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAIPKKRR 180
sog-PA Dwill     3 EESDTGRRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAMPKKRR  58
sog-PA Dmoj    102 EEADTARRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVSIPKKRR 157
sog-PB Dvril   103 EEADTARRSNRPA----VTECQFGKVLRELGSTWYADLGPPFGVMYCIKCECVAIPKKRR 158
sog-PA Dgrim   101 EEADTARRSNRP-----------------------ADVYGPTPVCVCV-CVCVCIPKKRR 136
                   **.** ******                       **.  *  *  *. * **..*****

sog-PA Dmel    142 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPNEEEERNMKHY 201
sog-PB Dsim    142 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPNEEEERNMKHY 201
sog-PA Dsec    142 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPNEEEERNMKHY 201
sog-PA Dyak    156 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPNEEEERNMKHY 215
sog-PA Dere    160 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPNEEEERNMKHY 219
sog-PA Dpers   181 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPNEEEERNMKHY 240
sog-PA Dwill    59 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVASDVPVPSEEEERNMKHY 118
sog-PA Dmoj    158 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDAASDVPVPSEEEERNMKHY 217
sog-PB Dvril   159 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPSEEEERNMKHY 218
sog-PA Dgrim   137 IVARVQCRNIKNECPPAKCDDPISLPGKCCKTCPGDRNDTDVALDVPVPSEEEERNMKHY 196
                   ***************************************** * ***** **********
```

Development • Supplementary information

```
sog-PA Dmel    202 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSSRIGRPRAIQF  261
sog-PB Dsim    202 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSSRIGRPRAIQF  261
sog-PA Dsec    202 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSSRIGRPRAIQF  261
sog-PA Dyak    216 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSSRIGRPRAIQF  275
sog-PA Dere    220 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSSRIGRPRAIQF  279
sog-PA Dpers   241 AALLTGRTSYFLKGEEMKSMYTTYNPHNIVATARFLFHKKNLYYSFYTSSRIGRPRAIQF  300
sog-PA Dwill   119 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSSKVGRPRAIQF  178
sog-PA Dmoj    218 AALLTGRTSYFLKGEEMKSMYTTYNPQNIVATARFLFHKKNLYYSFYTSTKIGRPRAIQF  277
sog-PB Dvril   219 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSSRIGRPRAIQF  278
sog-PA Dgrim   197 AALLTGRTSYFLKGEEMKSMYTTYNPQNVVATARFLFHKKNLYYSFYTSARIGRPRAIQF  256
                   **************************.*.*******************...********

sog-PA Dmel    262 VDDAGVILEEHQLETTLAGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  321
sog-PB Dsim    262 VDDAGVILEEHQLETTLAGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  321
sog-PA Dsec    262 VDDAGVILEEHQLETTLAGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  321
sog-PA Dyak    276 VDDAGVILEEHQLETTLAGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  335
sog-PA Dere    280 VDDAGVILEEHQLETTLAGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  339
sog-PA Dpers   301 VDDAGVILEEHQLETTLTGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  360
sog-PA Dwill   179 VDDGGVILEEHQLETTLTGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  238
sog-PA Dmoj    278 VDDRGDILEEHQLETTLSGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  337
sog-PB Dvril   279 VDDAGVILEEHQLETTLSGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  338
sog-PA Dgrim   257 VDDAGVILEEHQLETTLSGTLSVYQNATGKICGVWRRVPRDYKRILRDDRLHVVLLWGNK  316
                   *** * ***********.******************************************

sog-PA Dmel    322 QQAELALAGKVAKYTALQTELFSSLLEAP--LPDGKTDPQLAGAGGTAIVSTSSGAASSM  379
sog-PB Dsim    322 QQAELALAGKVAKYTALQTELFSSLLEAP--LPDGKTDPQLAGAGGTAIVSTSSGAASSM  379
sog-PA Dsec    322 QQAELALAGKVAKYTALQTELFSSLLEAP--LPDGKTDPQLAGAGGTAIVSTSSGAASSM  379
sog-PA Dyak    336 QQAELALAGKVAKYTALQTELFSSLLEAP--LPDGKTDPQLAGAGGTAIVSTSSGAASSM  393
sog-PA Dere    340 QQAELALAGKVAKYTALQTELFSSLLEAP--LPDGKTDPQLAGAGGTAIVSTSSGAASSM  397
sog-PA Dpers   361 HQSELALAGKIAKYTALQTELFSSLLEPPQSQPGGKFDPMLTGAGGTAIVSTSSGAASSM  420
sog-PA Dwill   239 HQSELALAGKIAKYTALQSELFSSLLEPS-----NIQDPQLAGAGGTAIVSTSSGVASSM  293
sog-PA Dmoj    338 HQAELALAGKIAKYTALQTELFSSLLEP--AATNGKTDPQLAGAGGTAIVSTSSGAASSM  395
sog-PB Dvril   339 HQAELALAGKIAKYTALQTELFSSLLEPP-ALVGGKPDPQLAGAGGTAIVSTSSGAASSM  397
sog-PA Dgrim   317 HQSELALAGKIAKYTALQTELFSSLLEPP-LTASGKPDPQLAGAGGTAIVSTSSGAASSM  375
                     .*.*******.*******.********        .   ** *.************* ****

sog-PA Dmel    380 HLTLVFNGVFGAEEYADAALSVKIELAERKEVIFDEIPRVRKPSAEINVLELSSPISIQN  439
sog-PB Dsim    380 HLTLVFNGVFGAEEYADAALSVKIELAERKEVIFDEIPRVRKPSAEINVLELSSPISIQN  439
sog-PA Dsec    380 HLTLVFNGVFGAEEYADAALSVKIELAERKEVIFDEIPRVRKPSAEINVLELSSPISIQN  439
sog-PA Dyak    394 HLTLVFNGVFGAEEYADAALSVKIELAERKEVIFDEIPRVRKPSAEINVLELSSPISIQN  453
sog-PA Dere    398 HLTLVFNGVFGAEEYADAALSVKIELAERKEVIFDEIPRVRKPSAEINVLELSSPISIQN  457
sog-PA Dpers   421 HLTLVFNGIFGAEEYADAALSVKIELPERKEIIFDEVPRVRKPSAEINVLELSSPISIQN  480
sog-PA Dwill   294 HLTLIFNGVFGSEEFADAALSVRIELPERKEMIFDEVPRVRKPSAEINVLELSSPISIQN  353
sog-PA Dmoj    396 HLTLVYNGIFGVEEYADAGLSVRIELPERKELILDEVSRVRKPSAEINVLELSSPISIQN  455
sog-PB Dvril   398 HLTLVFNGIFGAEEYADAALSVRIELPERKELIFDEVPRVRKPSAEINVLELSSPISIQN  457
sog-PA Dgrim   376 HLTLVFNGIFGTEEYADSTLSVRIELPERKEVIFDELSRVRKPSAEINVLELSSPISIQN  435
                   ****..**.** **.**. ***.*** ****.* **. *********************
```

Development • Supplementary information

```
sog-PA Dmel     440 LRLMSRGKLLLTVESKKYPHLRIQGHIVTRASCEIFQTLLAP--HSA----------ESS 487
sog-PB Dsim     440 LRLMSRGKLLLTVESKKYPHLRIQGHIVTRASCEIFQTLLAP--HSA----------ESS 487
sog-PA Dsec     440 LRLMSRGKLLLTVESKKYPHLRIQGHIVTRASCEIFQTLLAP--HSA----------ESS 487
sog-PA Dyak     454 LRLMSRGKLLLTVESKKYPHLRIQGHIVTRASCEIFQTLLAP--HSA----------ESS 501
sog-PA Dere     458 LRLMSRGKLLLTVESKKYPHLRIQGHIVTRASCEIFQTLLAP--HSA----------ESS 505
sog-PA Dpers    481 LRLMSRGKLLLTVESKKYPQLRIQGHIVTRASCEIFQTLLAP--HNG----------ESA 528
sog-PA Dwill    354 LRLMSRGKLLLIVESKKYPQLRIQGHIVTRASCEIFQTLLAP--HASGSGSTETTALASN 411
sog-PA Dmoj     456 LRLMSRGKLLLTVESKKYPQLRIQGHIVTRASCEIFQTLLAP-SHGN---------ELS- 504
sog-PB Dvril    458 LRLMSRGKLLLTVESKKYPQLRIQGHIVTRASCEIFQTLLAP-PHGN---------ELS- 506
sog-PA Dgrim    436 LRLMSRGKLLLTVESKKYPQLRIQGHIVTRASCEIFQTLLAPPMHGG---------ELST 486
                    *********** *******.**********************  *             *

sog-PA Dmel     488 TKSSGLAWVYLNTDGSLAYNIETEHVNTRDRPNISLIEEQGKRKAKLEDLTPSFNFNQAI 547
sog-PB Dsim     488 TKSSGLAWVYLNTDGSLAYNIETEHVNTRDRPNISLIEEQGKRKAKLEDLTPSFNFNQAI 547
sog-PA Dsec     488 TKSSGLAWVYLNTDGSLAYNIETEHVNTRDRPNISLIEEQGKRKAKLEDLTPSFNFNQAI 547
sog-PA Dyak     502 TKSSGLAWVYLNTDGSLAYNIETEHVNTRDRPNISLIEEQGKRKAKLEDLTPSFNFNQAI 561
sog-PA Dere     506 TKSSGLAWVYLNTDGSLAYNIETEHVNTRDRPNISLIEEQGKRKAKLEDLTPSFNFNQAI 565
sog-PA Dpers    529 TRSSGLAWVYLNTDGSLAYNIETDHVNTRDRPNISLVEEQGKRKAKLEDLTPSFNFNQAI 588
sog-PA Dwill    412 SRSSGLAWVYLNTDGSLAYNIETDHVNTKDRPNISLIEE-GKRKAKLEDLTPSFNFNQAI 470
sog-PA Dmoj     505 TRSSGLAWVYLNTDGSLAYNIETDHVNTRDRPNISLIEEQGNRKAKLEDLTPSFNFNQAI 564
sog-PB Dvril    507 TRSSGLAWVYLNTDGSLAYNIETDHVNTRDRPNISLIEEQGKRKAKLEDLTPSFNFNQAI 566
sog-PA Dgrim    487 TRSSGLAWVYLNTDGSLAYNIETDHVNTRDRPNISLIEEQGKRKAKLEDLTPSFNFNQAI 546
                    ..*********************.****.*******.** * ******************

sog-PA Dmel     548 GSVEKLGPKVLESLYAGELGVNVATEHETSLIRGRLVPRPVADARDSAEPILLKRQ-EHT 606
sog-PB Dsim     548 GSVEKLGPKVLESLYAGELGVNVATEHETSLIRGRLVPRPVADARDSAEPILLKRQ-EHS 606
sog-PA Dsec     548 GSVEKLGPKVLESLYAGELGVNVATEHETSLIRGRLVPRPVADARDSAEPILLKRQ-EHS 606
sog-PA Dyak     562 GSVEKLGPKVLESLYAGELGVNVATEHETSLIRGRLVPRPVADARDSAEPILLKRQ-EHT 620
sog-PA Dere     566 GSVEKLGPKVLESLYAGELGVNVATEHETSLIRGRLVPRPVADARDSAEPILLKRQ-EHP 624
sog-PA Dpers    589 GSVEKLGPKVLESLYAGELGVNVGTEHETSLIRGRLVPRPVADARDSAEPILLKRQ-EHL 647
sog-PA Dwill    471 GSVEKLGPKVLESLYDGELVVNVATEHDTSLIRGRLVPRPVADARDSAEPILLKRP-EHS 529
sog-PA Dmoj     565 GSVEKLGPKVLESLYAGELGVNVATDHEASLIRGRLVPRPVADARDSAEPILLKAHVEHG 624
sog-PB Dvril    567 GSVEKLGPKVLESLYAGELGVNVATEHEASLIRGRLVPRPVADARDSAEPILLKPHGEHA 626
sog-PA Dgrim    547 GTVEKLGPKVLESLYAGELGVNVATEHEASLIRGRLVPRPVADARDSAEPILLKQHMEHG 606
                    *.************* *** *** *.*..*************************   **

sog-PA Dmel     607 -------DAQNPHAVGMAWMSIDNECNLHYEVTLNGVPA--QDLQLYLEEKPIEAIGAPV 657
sog-PB Dsim     607 -------DAQNPHAVGMAWMSIDNECNLHYEVTLNGVPA--QDLQLYLEEKPIEAIGAPV 657
sog-PA Dsec     607 -------DAQNPHAVGMAWMSIDNECNLHYEVTLNGVPA--QDLQLYLEEKPIEAIGAPV 657
sog-PA Dyak     621 Q------DAQNPHAVGMAWMSIDNECNLHYEVTLNGVPA--QDLQLYLEEKPIEAIGAPV 672
sog-PA Dere     625 -------DAQNPHAVGMAWMSIDNECNLHYEVTLNGVPA--QDLQLYLEEKPIEAIGAPV 675
sog-PA Dpers    648 AA---SGSNQSGHAMGMAWMSIDNECNLHYELTLSGVPP--QDLQLYLEEKPIEAIGAPV 702
sog-PA Dwill    530 P------SVGGEHSMGMAWMSIDNECNLHYELTLSGLPS--QDMQLYLEEKPIEAIGAPV 581
sog-PA Dmoj     625 QN--------QGQGMGMAWMAIDNECNLHYELTLSGVPN--QELQLYLEEKPIEAIGAPV 674
sog-PB Dvril    627 S---------PGLGMGMAWMSIDNECNLHYELTLSGMPNAGQDLQLYLEEKPIEAIGAPV 677
sog-PA Dgrim    607 SSASASASASAGHSMGMAWMSIDNECNLHYEITLSGVPT--QELQLYLEEKPIEAIGAPV 664
                                  .*****.**********.** *.*   *..****************
```

Development • Supplementary information

```
sog-PA Dmel     658 TRKLLEEFNGSYLEGFFLSMPSAELIKLEMSVCYLEVHSKHSKQLLLRGKLKSTKVPGHC  717
sog-PB Dsim     658 TRKLLEEFNGSYLEGFFLSMPSAELIKLEMSVCYLEVHSKHSKQLLLRGKLKSTKVPGHC  717
sog-PA Dsec     658 TRKLLEEFNGSYLEGFFLSMPSAELIKLEMSVCYLEVHSKLSKQLLLRGKLKSTKVPGHC  717
sog-PA Dyak     673 TRKLLEEFNGSYLEGFFLSMPSAELIKLEMSVCYLEVHAKHSKQLLLRGKLKSTKVPGHC  732
sog-PA Dere     676 TRKLLEEFNGSYLEGFFLSMPSAELIKLEMSVCYLEVHSKHSKQLLLRGKLKSTKVPGHC  735
sog-PA Dpers    703 TRKLLEEFNGSYLEGFFLSMPSAELIKLEMSVCYLELHSKHSKQLLLRGKLKSTKVPAHC  762
sog-PA Dwill    582 TRKLLEEFNSSYMEGFGLGMPSAELIKLETSVCYLELQSKHSKQLLLRGKLKSTKVPAHC  641
sog-PA Dmoj     675 TRKLLKEFTGNYMEGFQLSMPSAELIKLETSVCYLELHSKVNPRLLLRGKLKSTKVPAHC  734
sog-PB Dvril    678 TRKLLEEFNGAYMEGFVLGMPSAELIKLETSVCYLELHAKHSHQLLLRGKLKSTKVPAHC  737
sog-PA Dgrim    665 TRKLLEEFNGQYMEGFLLGMPSAELIKLETSVCYLELHAKHSHQLLLRGKLKSTKVPAHC  724
                    ***** **.  *.*** * ********** ******...*   .************* **

sog-PA Dmel     718 FPVYTDNNVPVPGDHNDNHLVNGETKCFHSGRFYNESEQWRSAQDSCQMCACLRGQSSCE  777
sog-PB Dsim     718 FPVYTDNNVPVPGDHNDNHLVNGETKCFHSGRFYNESEQWRSAQDSCQMCACLRGQSSCE  777
sog-PA Dsec     718 FPVYTDNNVPVPGDHNDNHLVNGETKCFHSGRFYNESEQWRSAQDSCQMCACLRGQSSCE  777
sog-PA Dyak     733 FPVYTDNNVPVPGDHNDNHLVNGETKCFHSGRFYNESEQWRSAQDACQMCACLRGQSNCE  792
sog-PA Dere     736 FPVYTDNNVPVPGDHNDNHLVNGETKCFHSGRFYNESEQWRSAQDSCQMCACLRGQSSCE  795
sog-PA Dpers    763 FPVYTDNNVPVPGDHNDNHLINAETKCFHSGRFYNESEQWRSAQDACQMCACLRGQSNCE  822
sog-PA Dwill    642 FPIYTDNNVPVPGDQNDNHLINTETKCFHSGRFYNESEQWRSAQDTCQMCACLRGQANCE  701
sog-PA Dmoj     735 FPIYTDNNVPVPGDQNGNQMLSTESNCYHSGRFYNESEQWRSAQDTCHMCACQRGQANCE  794
sog-PB Dvril    738 FPIYTDNNVPVPGDHNDNQMLTAESKCFHSGRFYNESEQWRSAQDTCHMCACQRGQANCE  797
sog-PA Dgrim    725 FPIYTDNNVPVPGDHNDNQMLTAETKCFHSGRFYNESEQWRSAQDTCHMCACQRGQSNCE  784
                    **.***********.* *...   *. *.****************.*.**** ***. **

sog-PA Dmel     778 VIKCPALKCKST-EQLLQRDGECCPSCVPKK---------EAADYSAQ------------  815
sog-PB Dsim     778 VIKCPALKCKST-EQLLQREGECCPSCVPKK---------EAADYSAQ------------  815
sog-PA Dsec     778 VIKCPALKCKST-EQLLQREGECCPSCVPKK---------EAADYSAQ------------  815
sog-PA Dyak     793 VIKCPALKCKAT-EQLLQREGECCPSCVPKK---------EAADYSAQ------------  830
sog-PA Dere     796 VIKCPALKCKST-EQLLQREGECCPICVSKK---------EAADYSSQ------------  833
sog-PA Dpers    823 PQRS----------------------SVSAS---------ASASASS-------------  838
sog-PA Dwill    702 IIKCPALKCKS-------------PNEQLIQ---------REGECCPS------------  727
sog-PA Dmoj     795 LIKCPALKCKAG-EQLLQREGECCPSCVAK----PADWAQAQGQLSSS------------  837
sog-PB Dvril    798 LIKCPALKCKTGSEQLLQRDGECCPSCVARR---DADWHTAGSSSSSSSSSSTGSGSGS  854
sog-PA Dgrim    785 PIKCPALKCKPLTEQLLQREGECCPTCVARRGENGADWHMAAGATTSSTS---------I  835
                      ..                                        .

sog-PA Dmel     816 -SSPATNATDLLQQ--RRGCRLGEQFHPAGASWHPFLPPNGFDTCTTCSCDPLTLEIRCP  872
sog-PB Dsim     816 -SSPATNATDLLQQ--RRGCRLGEQFHPAGASWHPFLPPNGFDTCTTCSCDPLTLEIRCP  872
sog-PA Dsec     816 -SSPATNATDLLQQ--RRGCRLGEQFHPAGASWHPFLPPNGFDTCTTCSCDPLTLEIRCP  872
sog-PA Dyak     831 -SSPATNASDLLQQ--RRGCRLGEQFHPAGASWHPFLPPNGFDTCTTCSCDPLTLEIRCP  887
sog-PA Dere     834 -SSPATNATDLLQQ--RRGCRLGEQFHLAGASWHPFLPPNGFDTCTTCSCDPLTLEIRCP  890
sog-PA Dpers    839 -SLPAVNGSDLLLQQQRRGCRLGDQFHAAGASWHPFLPPNGFDTCTTCSCDPLTLEIRCP  897
sog-PA Dwill    728 -CMPKREPTALASES-GRGCRLGDQFHPAGASWHPFLPPNGFDTCTTCSCDPLTLEVRCP  785
sog-PA Dmoj     838 ---PAINASDLLQQ--RRGCRLGEQFHAAGASWHPFLPPNGFDTCTTCSCDAQTLEVRCP  892
sog-PB Dvril    855 GSSPALNASDVLQQ--RRGCRLGEQFHAAGASWHPFLPPNGFDTCTTCSCDALTLEVRCP  912
sog-PA Dgrim    836 AAHSNNGSSDLLQP--HRGCRLGDQFHAAGASWHPFLPPNGFDTCTTCSCDILTLEVRCP  893
                            .  .        ******.*** **********************  ***.***
```

Development • Supplementary information

```
sog-PA Dmel    873 RLVCPPLQCSEKLAYRPDKKACCKICPEGKQSSS--NGHKTTPNNPNVLQDQAMQRSPSH  930
sog-PB Dsim    873 RLVCPPLQCSEKLAYRPDKKACCKICPEGKQSSS--NGHKTAPNNPNVLQDQAMQRSPSH  930
sog-PA Dsec    873 RLVCPPLQCSEKLAYRPDKKACCKICPEVKQSSS--NGHKTAPNNPNVLQDQAMQRSPSH  930
sog-PA Dyak    888 RLVCPPLQCSEKLAYRPDKKACCKICPEGKQSSS--NGHKAAPNNPNVLQDQAMQRSPSH  945
sog-PA Dere    891 RLVCPPLQCSEKLAYRPDKKACCKICPEGKQSSS--NGHKAAPNNPNVLQDQAMQRSPSH  948
sog-PA Dpers   898 RLVCPPLQCSEKLAYRPDKKACCKVCPEGKQSSP--AGGKAAPNNPNVLQDQAMQRSPQH  955
sog-PA Dwill   786 RMVCPPLQCSEKLAYRPDKKACCKVCPEGKQSSS--NSDKSMPNNPNVLHDQAMPRTPGH  843
sog-PA Dmoj    893 RMVCPPLQCSEKLAYRPDKKACCKVCPEGKQSNHGASSKHQAPANHNELHDQAVQRTAAH  952
sog-PB Dvril   913 RMVCPPLQCSEKLAYRPDKKACCKVCPEGKQSNHGQSGKH-TPSNPNVLLDQAVQRTAAH  971
sog-PA Dgrim   894 RMVCPPLQCIEKLAFRPDKKACCKVCPEGKQSTHGQSSKH-TPSNPNVLHDQAMQHSAAH  952
                   *.******* ****.*********.*** ***       .  * * * * ***. ..  *

sog-PA Dmel    931 SAEEVLANGGCKVVNKVYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ  990
sog-PB Dsim    931 SAEEVLANGGCKVVNKVYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ  990
sog-PA Dsec    931 SAEEVLANGGCKVVNKVYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ  990
sog-PA Dyak    946 SAEEVLANGGCKVVNKVYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ 1005
sog-PA Dere    949 SAEEVLANGGCKVVNKVYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ 1008
sog-PA Dpers   956 SAEDVLAAGGCKVVNKVYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ 1015
sog-PA Dwill   844 TAEEILAAGGCKVVNKIYENGREWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ  903
sog-PA Dmoj    953 NAEEVLAAGGCKVVNKVYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ 1012
sog-PB Dvril   972 NAEEVLAAGGCKVVNKIYENGQEWHPILMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ 1031
sog-PA Dgrim   953 TAEEVLNAGGCKVVNKIYENGQEWHPVLMSHGEQKCIKCRCKDSKVNCDRKRCSRSTCQQ 1012
                    **..*  ********.****.****.*********************************

sog-PA Dmel    991 Q---TRVTSKRRLFEKP---------DAAAPAIDECCSTQCRRSRRHHKRQPHHQQRSSS 1038
sog-PB Dsim    991 Q---TRVTSKRRLFEKP---------DAAAPAIDECCSTQCRRSRRHHKRQPHHQQRSSS 1038
sog-PA Dsec    991 Q---TRVTSKRRLFEKP---------DAAAPAIDECCSTQCRRSRRHHKRQPHHQQRSSS 1038
sog-PA Dyak   1006 Q---TRVTSKRRLFEKP---------DAAAPAIDECCSTQCRRSRRHHKRQPHHQQRSSS 1053
sog-PA Dere   1009 Q---TRVTSKRRLFEKP---------DAAAPAIDECCSTQCRRSRRHHKRQPHHQQRSSS 1056
sog-PA Dpers  1016 Q---TRVSSKRRIFEKP---------DS-APTLDECCSTQCRRSRRHHKRQPHHQQRSNA 1062
sog-PA Dwill   904 ----TRVSSKRRPFEKPGN-----GNDGSGAPIDECCSTQCRRSRRHHKRQPHHQ-----  949
sog-PA Dmoj   1013 ----TRVSSKRRLFDKQTGVAEL--VGGVAQPIDECCSTQCRRSRRHHKRQPHHQQRATS 1066
sog-PB Dvril  1032 QQQQTRVSGKRRLFDKPTAGSELGGVAGAAAPIDECCSTQCRRSRRHHKRQPHHQQRAAS 1091
sog-PA Dgrim  1013 ----TRVSGKRRLFDKPGSEAATG-AGGGPQPIDECCSTQCRRSRRHHKRQPHHQQRAAS 1067
                       ***. *** *.*                .**********************

sog-PA Dmel   1039 ----- 1038
sog-PB Dsim   1039 ----- 1038
sog-PA Dsec   1039 ----- 1038
sog-PA Dyak   1054 ----- 1053
sog-PA Dere   1057 ----- 1056
sog-PA Dpers  1063 AASSS 1067
sog-PA Dwill   950 -----  949
sog-PA Dmoj   1067 ----- 1066
sog-PB Dvril  1092 S---- 1092
sog-PA Dgrim  1068 S---- 1068
```

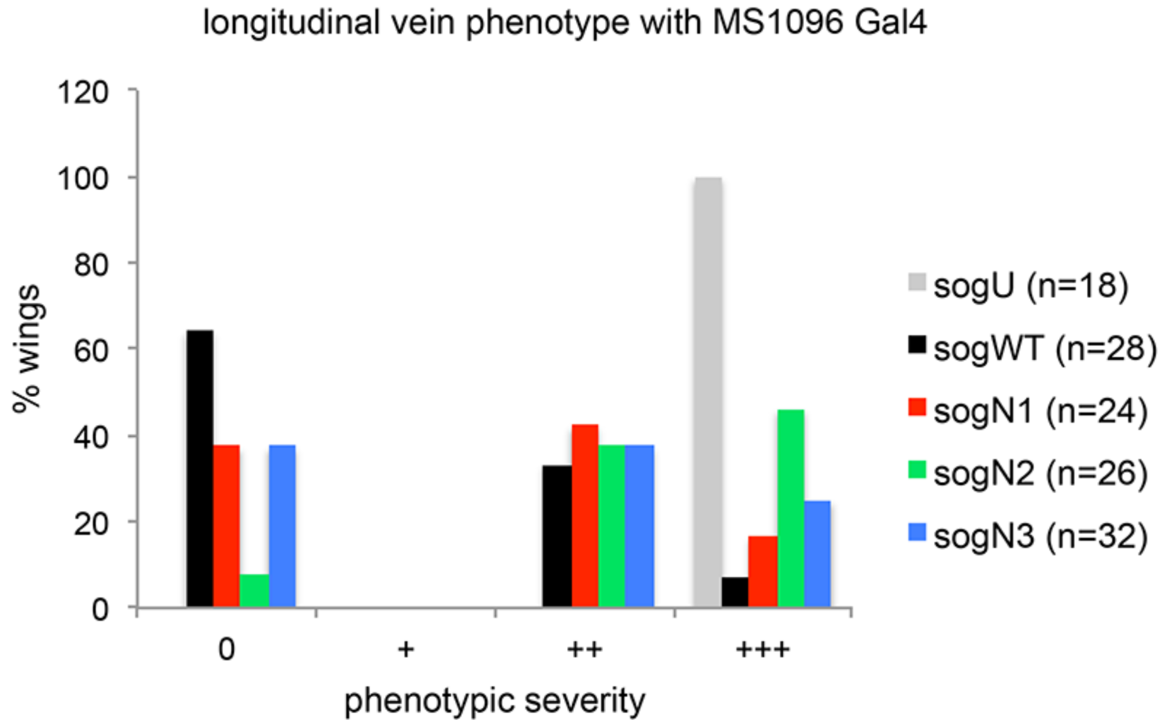

Figure S3. Ubiquitous expression of *sogN* mutants leads to vein truncation. Longitudinal vein truncation phenotypes resulting from ubiquitous overexpression of wild type *sog* and *sogN* mutants in the wing primordia, driven by MS1096-Gal4. The uncleavable *sog-u* mutant (Peluso et al, 2011) was also included in the analysis.

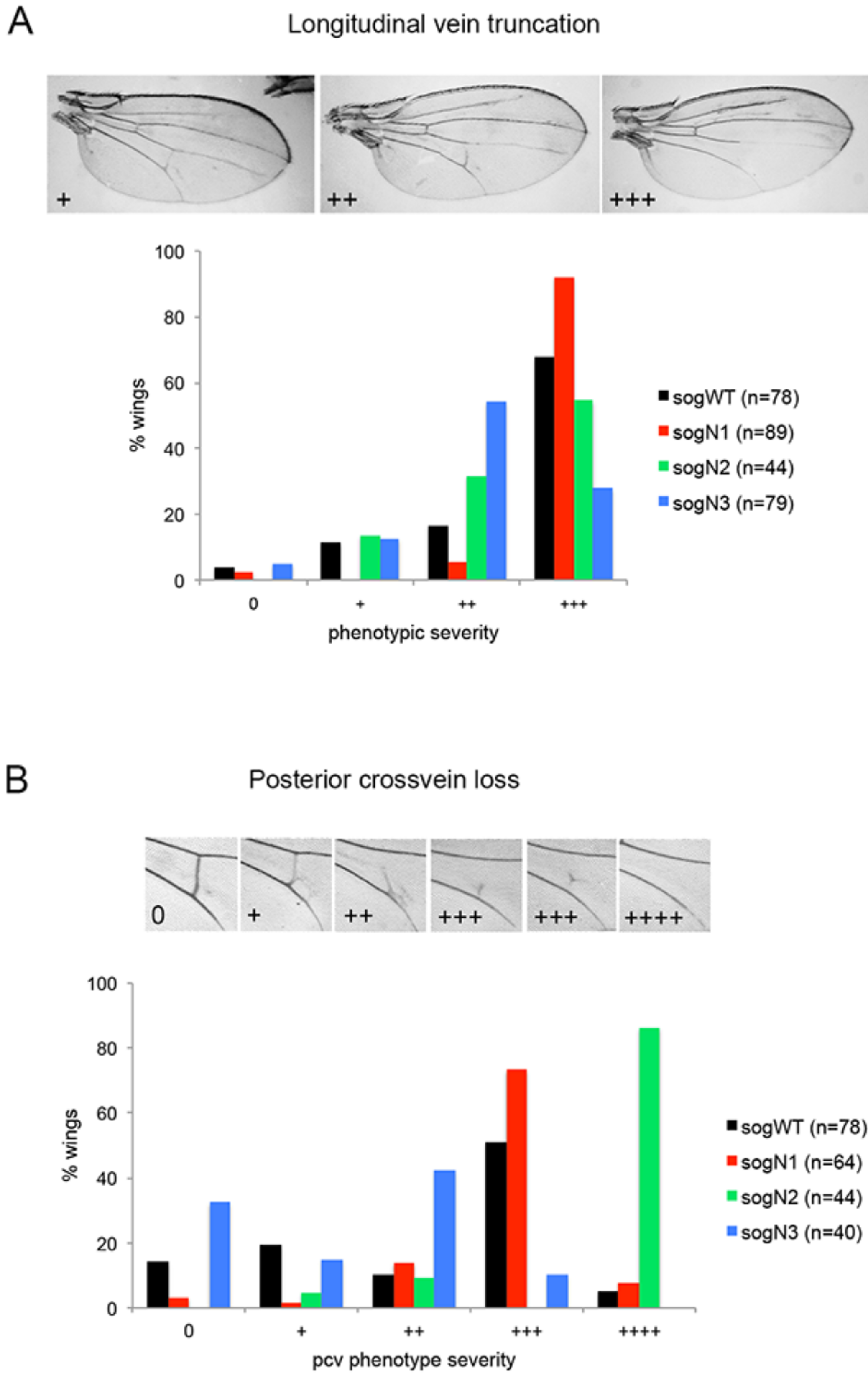

Fig. S4. Longitudinal vein expression of *sogN* mutants increases a *dpp* loss-of-function phenotype. Longitudinal vein truncation phenotypes (A) or PCV phenotypes (B) resulting from

Development • Supplementary information

longtudinal vein overexpression of wild-type *sog* and *sogN* mutants in the wing primordia, driven by shv-Gal4. Expression was driven in a *dpp[shv]* heterozygous background, revealing that all constructs induce a strong vein loss phenotype. 100% of *dpp[shv]/*+ wings have a wild type (0) wing venation pattern.

Development • Supplementary information

Supplementary Material and Methods

Sequence accession numbers were: NP-476736 (D. *melanogaster*), FBpp0320361 (XP_016039421.1*, D. simulans*), FBpp0137934 (XP_001978067, *D. erecta*), FBpp0261045 (XP_0002100689, *D. yakuba*), FBpp0393205 (*D. virilis*), FBpp0170731 (XP_002010774*, D. mojavensis*), FBgn 0131639 (*D. grimshawi*), FBpp0254766 (XP_002071108, *D. willistoni*), FBp0180550 (XP_002022952, *D. persinilis*), FBpp0204020 (XP_002044153 (*D. sechellia*). Chordin sequences were downloaded from NCBI for *Mus musculus (NP_001264970 isoform 2 precursor and NP_034023 isoform 1 precursor), Homo sapiens (NP_001291401 isoform 2 precursor and NP_003732 isoform 1 precursor), Gallus gallus (NP_990311), Danio rerio (NP_571048) and Xenopus laevis (NP_001081778)*.

Production of S2 cell and transgenic fly constructs

Wild type *sog* or mutant *sog* were amplified from Retrogen mutated pBluescript II KS (+), using SogKpnI and SogNotI primers and transferred to pMT-V5/HisC (Invitrogen), by digestion with KpnI plus NotI.
For the production of double mutant pBS-sogN2,3 a EcoRI + NotI fragment was produced from pBS-sogN3 and ligated into EcoRI + NotI digested pBS-sogN2. For production of the triple mutant pBS-sogN1,2,3 a SmaI + NotI pBS-sogN2,3 fragment was ligated into SmaI + NotI digested pBS-sogN1. Full-length wild type *sog* or single mutant *sog* were obtained from pBluescript II KS (+) by KpnI + XbaI digestion and ligated into pTiger.
All constructs were sequenced before use. To specifically check for the presence of single nucleotide mutations and the presence of the V5 tag, the primers below were used in addition to T7 and T3 universal primers.

Table S1.

primer	
SogKpnI (F)	TAGGTACCACATGGCCAACAAGCTGAGGAAATC
SogNotI (R)	ATGCGGCCGCGCTGGAGGATCGCTGCTGATGATG
primers for mutant and tag sequencing	
SogN1	AATTGGGGTCCACCTGGTAT
SogN2	GAATCCTCGACCAAGAGCAG
SogN3	TTGCGAGGTGATCAAGTGTC
V5/His	GCAGCTGCTCCGGCCATC

Table S2. Primers used in this study for double-stranded RNA production and RNA interference assays (Fig. 6):

Gene	forward primer	reverse primer
tkvH	TAATACGACTCACTATAGGGAGAGAGAA GCTGCGCAAGC	TAATACGACTCACTATAGGGAGAGCTGGTTT GCCAGGGGT
tkvM	TAATACGACTCACTATAGGGAGAGAACC ATTGCCAAGCAGATTCAGAT	TAATACGACTCACTATAGGGAGATGAATGA CATCCAGTTCCGAGTTGT
saxH	TAATACGACTCACTATAGGGAGAGTGAA TGTGGTCTGCTGTG	TAATACGACTCACTATAGGGAGACTCCCG CTTCCAGGACT
saxM	TAATACGACTCACTATAGGGAGACGCGA TGCCGATGGTCAGGTGCAGGAG	TAATACGACTCACTATAGGGAGACCTCGT CCAATGCACTCGATCAGGG
mysH	TAATACGACTCACTATAGGGAGACCTCT TCGGTGGAGATGAA	TAATACGACTCACTATAGGGAGAGGATTTG GTCGCTTGTGG
mysM	TAATACGACTCACTATAGGGAGAATGAA GGACAACGCCACTGGAGATG	TAATACGACTCACTATAGGGAGACTCAATG GGATTTGGTCGCTTGTGG
mewH	TAATACGACTCACTATAGGGAGAGCAAG GAGCTGCATAACAT	TAATACGACTCACTATAGGGAGACGGCTC CACCCGGATAA
mewM	TAATACGACTCACTATAGGGAGAAACAC GGACTTGGATGATAACTCCT	TAATACGACTCACTATAGGGAGACAATATC GGATTGAGGCGTACCAAA
ifH	TAATACGACTCACTATAGGGAGACTGTT AATGCCGAAACCAG	TAATACGACTCACTATAGGGAGAATCTGG CTCGCAAATATTGT
ifM	TAATACGACTCACTATAGGGAGATGCGT GGACACAGCACACAGGGTCA	TAATACGACTCACTATAGGGAGACCTCATT CTCGCGGCCATAGTAGGG

Development • Supplementary information