Indirect Sexual Selection Drives Rapid Sperm Protein Evolution

Damien B. Wilburn[1*], Lisa M. Tuttle[2], Rachel E. Klevit[2], and Willie J. Swanson[1]

Affiliations:

[1]Department of Genome Sciences, University of Washington, Seattle, WA

[2]Department of Biochemistry, University of Washington, Seattle, WA

*Correspondence to: dwilburn@u.washington.edu

Abstract:

Sperm-specific genes often evolve faster than the rest of a genome, but it is unclear why. The marine mollusk abalone is a classic reproductive model where sperm secrete enormous quantities of two rapidly evolving fertilization proteins that are stored at extraordinary concentrations. Another rapidly evolving sperm protein enables packaging by forming Fuzzy Interacting Transient Zwitterion (FITZ) complexes. This protein – the FITZ Anionic Partner (FITZAP) – is intrinsically disordered and selectively forms FITZ complexes by exploiting differences in the intracellular and oceanic ionic environments. As fertilization proteins adapt to maintain interactions with egg receptors via direct sexual selection, FITZAP experiences this selection indirectly by recognizing the same molecular interface. This tethered molecular arms race provides a framework for understanding the widespread pattern of rapid sperm protein evolution.

Genes facilitating fertilization are usually among the fastest evolving in any genome (*1*). Given the necessity of fertilization for genetic propagation of species with sexual reproduction, these evolutionary dynamics may be perceived as paradoxical. For most somatic traits evolving via natural selection, benefits to reproduction and fitness are usually indirectly conferred by increasing the organism's survivability and improving the likelihood of that individual and its offspring to procreate (*2*). However, for traits and phenotypes associated directly with reproduction and fertilization, sexual selection favors traits that improve mate competitiveness and increase reproduction rates. As parental genomes recombine, the attractants and their preferences become genetically coupled and produce a Fisherian runaway process of increasing elaboration, even for highly costly traits such as male ornamentation (*3, 4*). Similarly, sexual selection and rapid coevolution occurs between interacting male and female proteins (*5, 6*). The first discovered pair of interacting reproductive proteins, sperm lysin and the egg vitelline envelope receptor of lysin (VERL) in the marine mollusk abalone, provide a textbook example of this phenomenon.

VERL is a major component of the vitelline envelope (VE), an extracellular barrier that protects the egg from the external environment and restricts the entry of sperm. Lysin dissolves the VE by binding to ZP-N domains within VERL (*7, 8*). As mutations have fixed in VERL (possibly due to genetic drift and/or mitigating polyspermy risk), this created positive sexual selection on lysin to maintain binding affinity and produced a coevolutionary chase. Consequently, extant lysin proteins dissolve conspecific VEs more efficiently than those of closely related taxa, providing one mechanism of species-specific fertilization (*8, 9*). As a nonenzymatic, stoichiometrically-dependent process, the rate of VE dissolution is positively correlated with lysin concentration, and selection has favored sperm that express enormous quantities of lysin (*10*). A single male abalone can contain ≥1 gram of lysin, reflecting > 0.1% of its total body weight. Inside sperm cells, lysin is stored in a specialized secretory granule termed the acrosome. Based on electron microscopy imaging (*11, 12*), we estimate that the acrosomal concentration of lysin is ~0.1 – 1.0 M, in stark contrast to reported saturation concentrations of ~0.001 M under *in vitro* conditions (*13*). We investigated what special conditions permit such extraordinarily high lysin concentrations inside the sperm acrosome.

Lysin is not the only highly abundant, rapidly evolving protein within abalone sperm acrosome. Another is sp18, a paralog of lysin that is fusogenic and is thought to mediate fusion between the egg and sperm plasma membranes (*14*). The receptor of sp18 is unknown, but given its interaction with the abalone egg, its accelerated evolution is likely due to sexual selection (*15*). Sp18 is nearly as abundant as lysin, so it must also be packaged at high concentrations yet its fusogenic properties make it even less soluble than lysin (*16*). Recently, a new family of small acrosomal proteins termed sperm protein 6 kDa (sp6) was discovered by shotgun transcriptomics and proteomics (*17*). While also rapidly evolving and hypothesized to evolve via sexual selection, initial efforts to identify an egg binding partner for sp6 were unsuccessful (*18*). However, while lysin and sp18 are highly positively charged proteins (+12 to +24), isoforms of sp6 are highly anionic (-6 to -16) and include an N-terminal poly-aspartate region of variable length (1 – 11 residues). Given this charge complementarity, we hypothesized that sp6 may facilitate packaging of lysin and sp18 inside the sperm acrosome. We demonstrate that the rapid evolution of sp6 is due to intra-sperm protein coevolution with lysin and sp18 to allow for their dense storage in the acrosome via novel Fuzzy Interacting Transient Zwitterion (FITZ) complexes. Heterodimers of lysin-sp6 or sp18-sp6 form through hydrophobic interactions, and these heterodimers polymerize into large particles (diameter > 100 nm) through ionic interactions of the complementary positive and negative charges. Upon secretion of the acrosomal contents into highly ionic sea water, FITZ complexes are disrupted, with lysin and sp18 dispersed to facilitate fertilization. In light of its newly identified function, we have named sp6 the FITZ Anionic Partner (FITZAP).

Different species of abalone express different numbers of FITZAP isoforms named for the length of the N-terminal poly-aspartate region (*18*). In red abalone (*Haliotis rufescens*), two isoforms (FITZAP-4D and FITZAP-8D) result from alternative splicing of two different versions of exon 1 (that contains a signal peptide and the N-terminus with the poly-aspartate region) with a common exon 2 (that contains the majority of the C-terminus) (Fig S1). Through a combination of strong anion exchange (SAX) chromatography and reverse-phase high performance liquid chromatography (RP-HPLC), each isoform was purified to homogeneity. Mass spectral analysis revealed that both isoforms were smaller than their cDNA opening reading frame predicted (~3-4 kDa vs ~6 kDa). The observed masses are consistent with proteolytic processing of FITZAP in the Golgi complex by the proteases furin and carboxypeptidase B (Fig S2). Despite their high net positive charges, both lysin and sp18 co-eluted with FITZAP at high salt concentrations by SAX chromatography (Fig S3). Particularly striking is that sp18 (+22) eluted at higher salt concentrations than lysin (+12), which mirrors the elution profiles of FITZAP-4D (-10) versus FITZAP-8D (-8), suggesting isoform-specific interactions with lysin and sp18. While purified lysin showed no affinity for anion exchange resin, its elution was retarded when mixed with either FITZAP-4D or FITZAP-8D *in vitro*, consistent with the co-elution of lysin/sp18 and FITZAP from sperm lysate resulting from *in vivo* interactions.

Nuclear magnetic resonance (NMR) spectroscopy was used to investigate how FITZAP interacts with the cationic fertilization proteins. We focused on lysin and FITZAP-8D because 1) they showed the strongest coelution by SAX chromatography (Fig S3B), 2) lysin is more soluble than sp18 (*14, 15*), 3) interactions with its egg receptor have been characterized (*7, 13*), and 4) a solution structure has been determined (*13*). Chemical shift analysis of FITZAP-8D revealed that it is an intrinsically disordered protein; even when bound to lysin, FITZAP-8D remained highly dynamic and adopted no regular secondary structure (Fig S4). Thus, lysin and FITZAP form a fuzzy complex: protein complexes that exist in an ensemble of different interchanging configurations. Formation of the fuzzy complex is primarily due to packing between hydrophobic amino acids in residues 21-29 of FITZAP (Fig S5) and an exposed hydrophobic face on lysin near the nexus of the N- and C-termini. Covering this hydrophobic region with FITZAP imparts a high net negative charge to this face of lysin, likely explaining the affinity of lysin to anion exchange resins when FITZAP is present. Significantly, the FITZAP-binding region of lysin is also the same surface that recognizes its egg receptor VERL (Fig 1). We postulate that this exposed hydrophobic surface is responsible for purified lysin having a much lower *in vitro* solubility limit compared to the acrosome where FITZAP is also highly abundant.

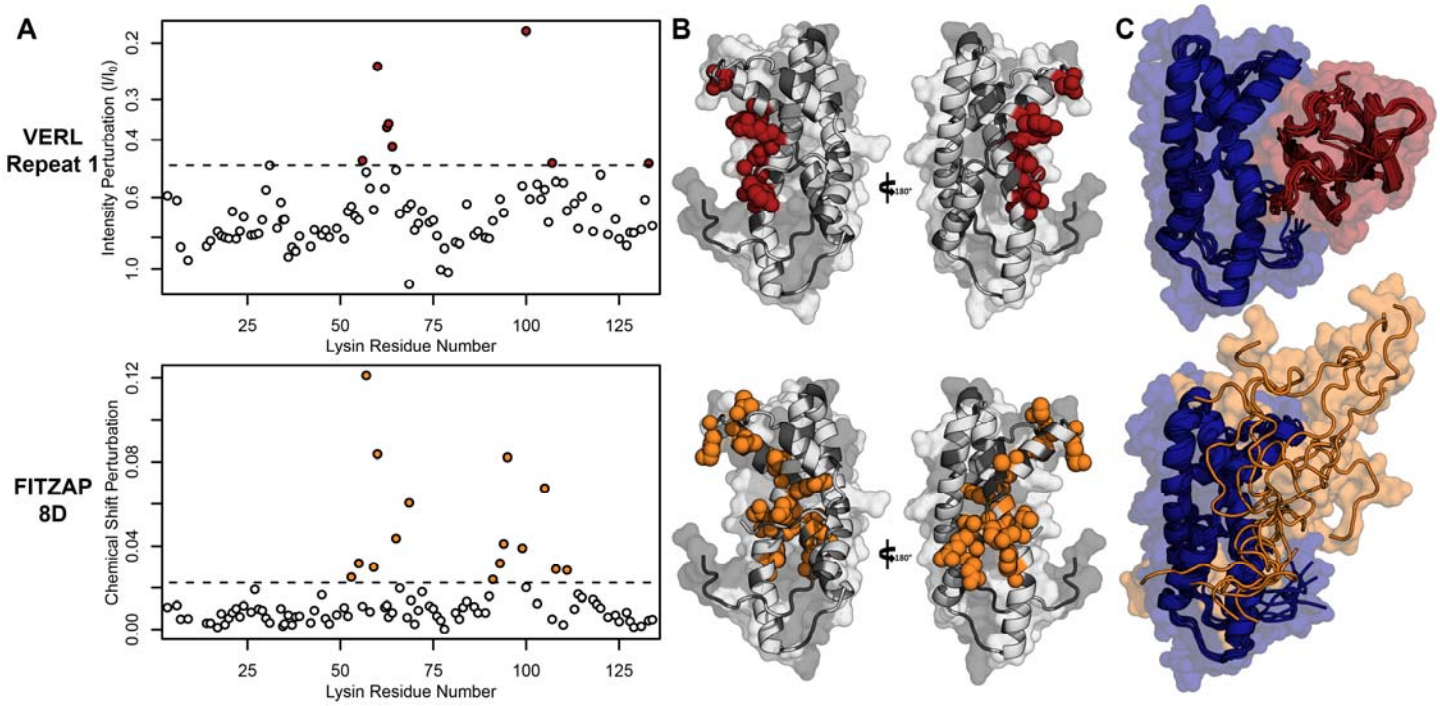

Fig. 1. Lysin recognizes VERL and FITZAP through a common binding interface. (A) NMR perturbation of red lysin upon conspecific binding of either VERL repeat 1 (data from Wilburn et al. (*13*)) or FITZAP-8D. Two regions near residue 60 and residue 100 of lysin are perturbed in both cases (dashed line: median + 1.5 * interquartile range). (B) Mapping of perturbation onto a lysin solution structure (PDB 5utg) shows spatial clustering of these two regions to a single binding surface of lysin. (C) Docking of VERL repeat 1 (based on PDB 5mr3 by Raj et al. (*7*)) and FITZAP-8D (based on restraints from paramagnetic relaxation enhancement) to lysin supports that FITZAP is an intrinsically disordered protein that shields the VERL binding interface.

Given a single interaction surface for two different identified binding partners, lysin must separately recognize FITZAP inside the acrosome and then VERL once it is secreted upon contact with the VE. As abalone is a marine mollusk, we postulated that local inorganic salt concentrations may play an important role. The total salt concentration of sea water is ~500 mM, yet the intracellular environment of marine animals is less concentrated. While fish and marine mammals principally rely on active transport of water and ions to maintain proper osmolality, abalone are osmoconformers: osmolality (but not isotonicity) is established by high intracellular concentrations of free amino acids, betaines, and other highly soluble metabolites (*19*). While the exact intracellular concentration of inorganic ions in abalone sperm is unknown, it is likely much lower than sea water based on data from sea urchin gametes (*20, 21*) and other abalone tissues (*22*). The inorganic salt concentration may be even lower in the acrosome where – given their extraordinary concentrations and net charges – the acrosomal proteins may themselves serve as osmolytes. To compare how lysin-FITZAP interactions may be influenced by differing environmental contexts, we performed biophysical experiments under salt concentrations approximating either intracellular (150 mM NaCl) or sea water (500 mM NaCl) conditions. Under both conditions, the hydrophobic patch of FITZAP is found to interact with lysin, yet NMR intensity perturbation of the poly-aspartate region was only observed under more intracellular salt concentrations where they may be forming intermolecular salt bridges (Fig S5). Consistent with additional types of molecular interactions, differences in NMR perturbation establish that lysin and FITZAP undergo slower subunit exchange under intracellular salt concentrations compared to sea water (Fig S6). Equimolar mixtures of lysin and FITZAP-8D form extremely large oligomers with an average diameter of ~400 nm (compared to a mean lysin diameter of ~6 nm) under intracellular salt conditions; these large particles are not present under sea

water conditions (Fig 2). Altogether, our data establish that lysin and FITZAP associate hydrophobically to form fast-exchanging, fuzzy heterodimers that, although highly charged, are essentially zwitterionic. These heterodimers that we call Fuzzy Interacting Transient Zwitterions (FITZs) are able to form intermolecular salt bridges that allow tight packaging under intracellular conditions. Upon secretion into sea water when sperm contact an egg, the FITZ complexes are disrupted and subunit exchange rate increases, allowing lysin to be rapidly liberated from FITZAP, permitting interactions with VERL via a common binding interface. The formation and dissolution of FITZ complexes based on the environmental context provides a mechanism for tight packaging of lysin inside the sperm acrosome as well as its rapid dispersal when fertilization may be imminent.

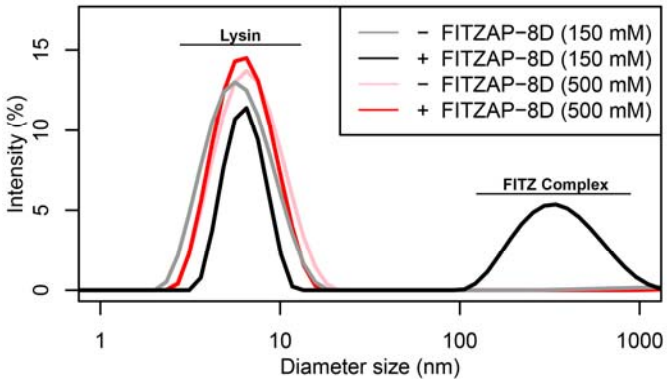

Fig. 2. FITZ complex formation is dependent on both FITZAP and the ionic environment. Dynamic light scattering measurements of lysin with and without equimolar FITZAP-8D at intracellular (150 mM) and sea water (500 mM) salt conditions. Lysin and FITZAP associate into FITZ complexes with mean diameter of ~ 400 nm only under intracellular salt conditions.

Lysin and VERL are rapidly coevolving and exhibit species-specific interactions (5), so given the common binding interface, we postulated that there may be similar coevolution between lysin and FITZAP. While we did not detect coevolution between lysin and FITZAP using sequence-based approaches (Fig S7), the small size and intrinsic disorder of FITZAP can reduce statistical power of such analyses. A functional consequence of lysin and FITZAP coevolution would be species-specific interaction, as observed between lysin and VERL. Therefore, lysin and FITZAP from the same species should have higher binding affinity compared to heterospecific pairs. Using fluorescence polarization, binding affinities were measured between lysin and all FITZAP isoforms for three abalone species (red, disk, and green). Like red abalone, green abalone has two FITZAP isoforms but with shorter poly-aspartate regions (1D and 4D), while disk abalone has a single FITZAP isoform with an even longer poly-aspartate array (11D). Lysin had greater affinity for the high-D FITZAP isoforms compared to low-D forms. For all three species, lysin bound the conspecific high-D isoforms of FITZAP with equilibrium dissociation constants of ~1-2 μM under intracellular salt conditions. Except for disk lysin and red FITZAP-8D, all cases of heterospecific binding were significantly weaker, consistent with lysin-FITZAP coevolution (Fig 3A). Under sea water conditions, there was a ≥ 20-fold decrease in conspecific binding affinity (Fig 3B), consistent with the liberation of lysin from FITZ complexes after its release from the acrosome.

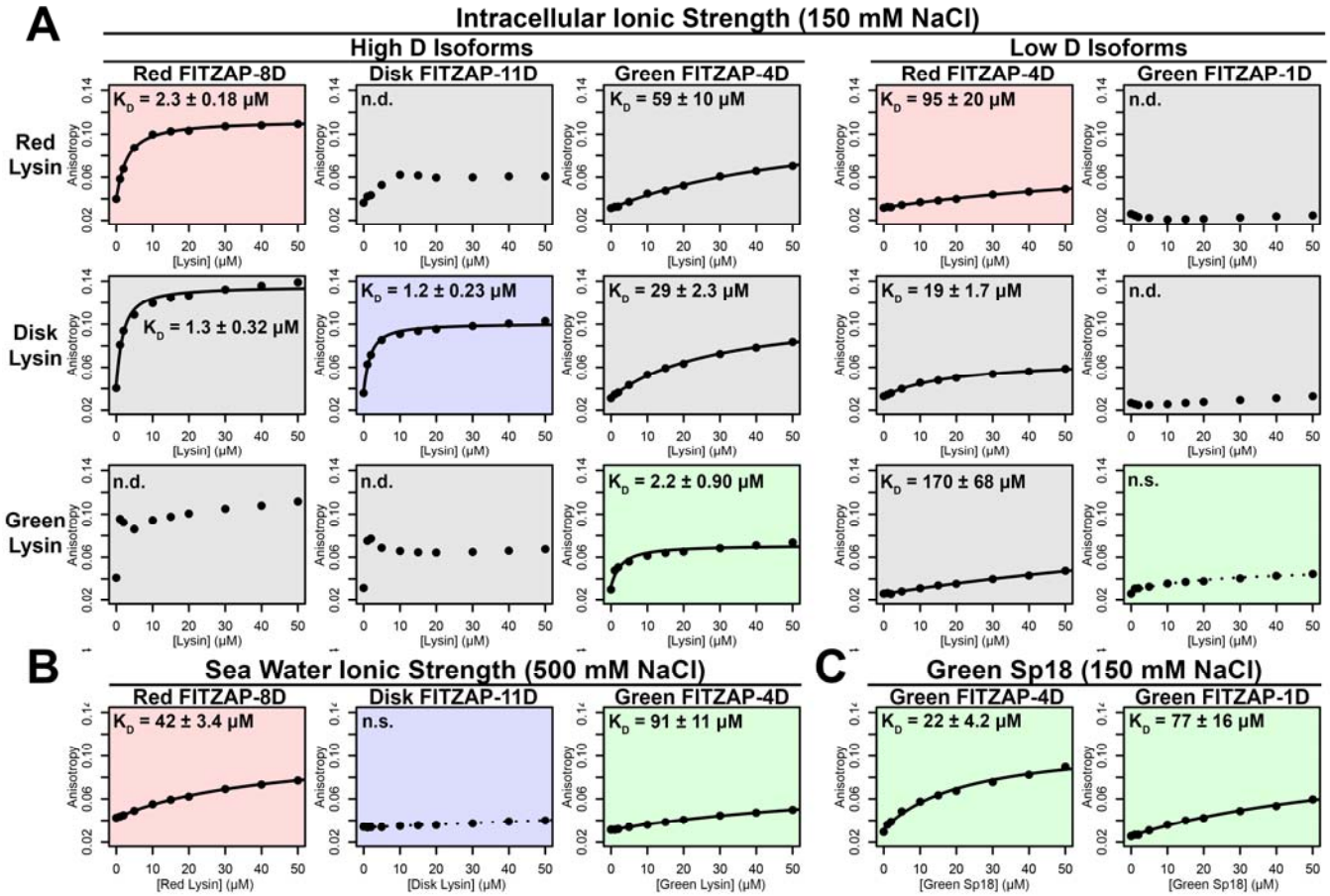

Fig. 3. FITZAP interactions with cationic acrosomal proteins within and between species. (A) Lysin-FITZAP interactions were measured by fluorescence polarization under approximately intracellular salt conditions for all combinations within and between species of red, disk, and green abalone (conspecific interactions are shaded as red, blue, and green, respectively). Except in the case of disk lysin with red FITZAP-8D, low micromolar binding affinities were observed for conspecific interactions of lysin with high-D FITZAP isoforms. (B) Binding affinities between conspecific lysin and FITZAP high-D isoforms are substantially reduced under extracellular sea water conditions. (C) Green sp18 shows tighter binding to conspecific FITZAP-4D compared to FITZAP-1D; however, the relative affinity of sp18 for 4D over 1D (77/22= 3.5 X) is relatively greater than that of lysin (> 100 X), suggesting that sp18 has relatively higher preference for low D isoforms compared to lysin. (n.s. = not significant at $p < 0.05$; n.d. = not determined if anisotropy was not monotonically positive consistent with single-state binding).

We propose a system of tethered coevolution between VERL, lysin, and FITZAP where VERL imposes direct sexual selection on lysin and indirect sexual selection on FITZAP (summarized in Fig 4). As with any coevolving system, there is likely reciprocity with all binding partners imposing some form of selection on one another; however, we choose to focus on the unidirectional case of VERL influencing lysin influencing FITZAP for several reasons. First, sexual selection theory has classically emphasized "female choice," because the higher energetic cost of oocytes in most species should favor greater mate selectivity (*3*). These assumptions are most apt when discussing genes involved in fertilization. Second, lysin evolves ~5X faster than its coevolving regions of VERL (*23*), suggesting that it is experiencing greater directional selection than VERL. Lysin, but not VERL, is also monomorphic within some abalone populations and experienced recent selective sweeps (*24*). Third, in sea urchins (another organism with broadcast spawning), longitudinal measurements of allele frequencies for gamete recognition proteins support that female proteins will shift to lower affinity interactions

in response to increased polyspermy risk (positive natural selection), which is then followed by male proteins adapting to restore high affinity binding (positive sexual selection) (*25*). Together, this supports that evolutionary dynamics of lysin are more responsive to VERL than vice versa. Third, as FITZAP is an IDP whose function is to facilitate the storage and dispersal of cationic acrosomal proteins, we anticipate minimal selection on its primary sequence beyond directional selection imposed by its binding partners.

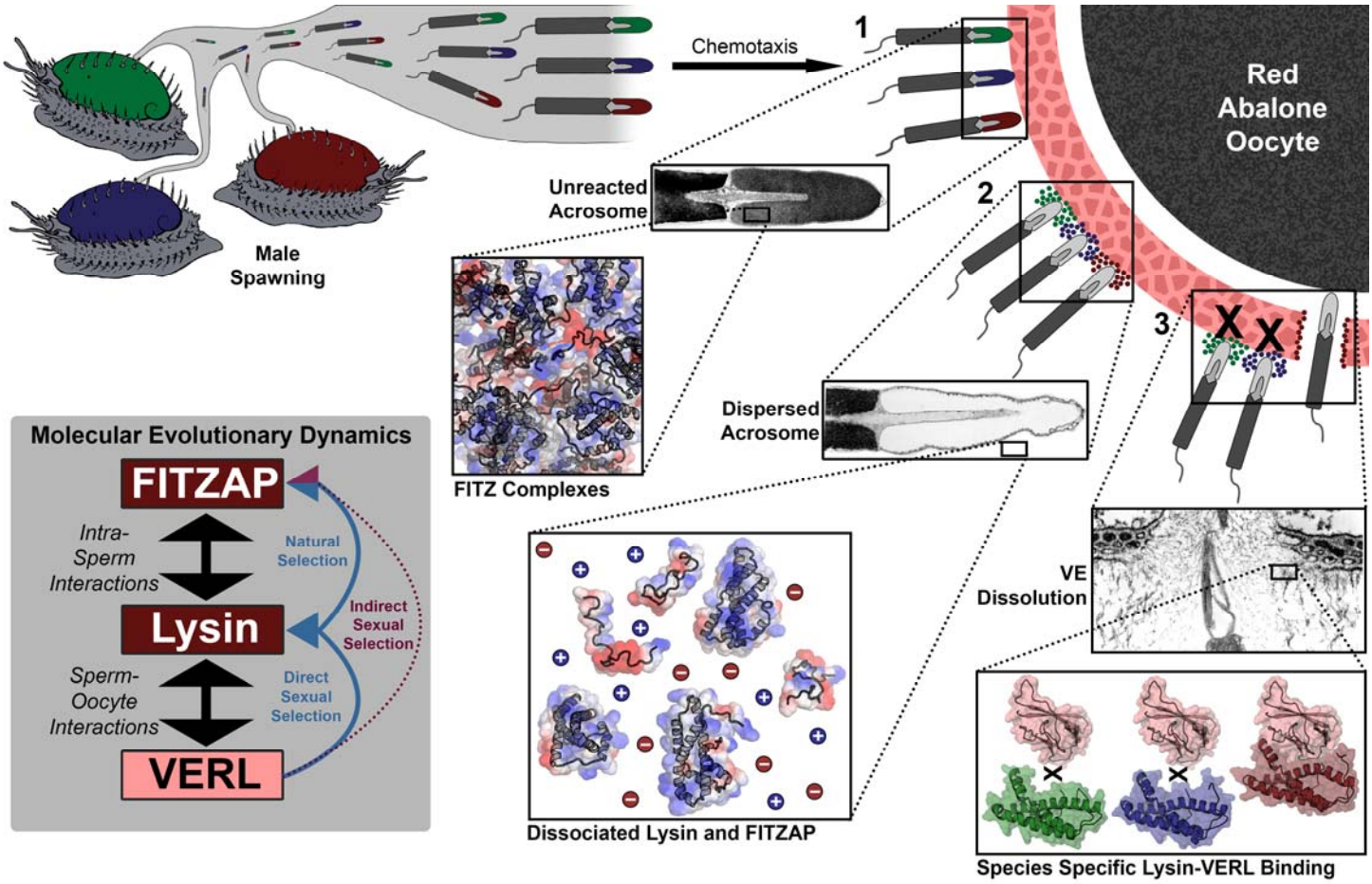

Fig. 4. Rapid evolution of FITZAP is due to indirect sexual selection from VERL. Male abalone have overlapping habitat ranges and spawning periods such that there is opportunity for hybridization to occur. Abalone oocytes release L-tryptophan as a chemoattractant for sperm (*26*), and after permeating through the egg jelly (not shown), sperm bind to the vitelline envelope **(1)**. At this stage, the acrosome is still intact (*10*) with lysin and FITZAP tightly packaged through electrostatic FITZ complex interactions. Binding to the VE causes the sperm to acrosome react **(2)**, releasing its contents into highly ionic sea water which disperses lysin and FITZAP. Liberated lysin can now dissolve the VE by binding VERL ZP-N domains **(3)**; these interactions are species-specific and provide one barrier to hybridization (*8*). Because the rapid evolution of lysin is driven by direct sexual selection to maintain sperm-egg interactions, and FITZAP is co-evolving to maintain affinity to the same interface that is evolving via direct sexual selection, VERL consequently imposes indirect sexual selection onto FITZAP.

Because IDPs lack a single favored conformation, they generally experience less purifying selection to maintain a tertiary fold and have lower sequence conservation (*27*). Beyond relaxed purifying selection and greater rates of genetic drift, IDPs also more commonly experience greater positive selection compared to

structural domains (*28*), presumably in response to coevolution with their binding partners. As such, we expect FITZAP to respond more to selection from lysin than vice versa. It has been suggested in the literature that intrinsic disorder may be a mechanism of adaptation to shifts in environmental conditions; for example, host-changing parasites have higher levels genome-wide of predicted protein disorder compared to obligate intracellular parasites and endosymbiotes (*29*). The change in ionic strength experienced by secreted proteins of most marine animals is likely an extreme example of such shifts between chemical environments. Thus, the intrinsic disorder of FITZAP may be crucial to its apparent structural versatility and high evolvability in response to a rapidly coevolving partner.

While we focused on lysin-FITZAP interactions because of their biochemical tractability, we were also interested in the interactions of FITZAP with sp18. As sp18 likely facilitates the fusion of egg and sperm plasma membranes (*14*), it is even more hydrophobic than lysin and may therefore be even more reliant on a partner for storage and dispersal. Lysin and sp18 are paralogs with similar tertiary structures that have subfunctionalized following gene duplication (*14, 16*), so the multiple FITZAP isoforms may also have subfunctionalized to serve these different paralogs. As lysin showed much tighter binding for high-D FITZAP isoforms (Fig 3A), we hypothesized that sp18 may be coevolving with low-D isoforms. Across Pacific abalone, we indeed observe correlated rates of molecular evolution between sp18 and low-D FITZAP isoforms, consistent with coevolution (Fig S7). While the highly fusogenic sp18 is mostly insoluble when purified (*14, 15*), we were able to measure conspecific affinities between sp18 and FITZAP isoforms from green abalone. While green sp18 also bound green FITZAP-4D more tightly than FITZAP-1D (Fig 3C), its relative affinity for 4D over 1D (~3.5X) is substantially less compared to the relative affinity of 4D to 1D for lysin (>100X). Additionally, for our anion exchange experiments in red abalone, we observed tighter co-elution between lysin/FITZAP-8D and sp18/FITZAP-4D (Fig S3). Therefore, both evolutionary and biochemical analyses support that FITZAP isoforms have subfunctionalized to respond to the divergent evolutionary trajectories of lysin and sp18. Furthermore, sp18 may be the major driver of FITZAP duplication and subfunctionalization. For abalone species with more than one FITZAP isoform, we observe faster rates of sp18 on the branches leading to these species (mean $d_N/d_S \pm$ SD: 6.89 ± 4.65) compared to those with one isoform (1.43 ± 0.66). We observe no differences in the evolutionary rate of lysin for species with one isoform (3.34 ± 0.72) or multiple isoforms (3.26 ± 1.73). Therefore, on branches where sp18 is evolving relatively slowly, it may be possible for a single FITZAP isoform to adapt and service both cationic clients; however, when sp18 is experiencing particularly strong selection, indirect sexual selection favors duplication and subfunctionalization of an additional isoform.

It is well documented in animals, plants, and microbes that genes associated with reproduction evolve at accelerated rates (*1, 5, 30*), and in mammals specifically, spermatozoa-specific genes show the greatest divergence between species (*31*). Here we demonstrated how indirect sexual selection drives rapid evolution of a sperm protein not directly associated with fertilization. We postulate that such indirect selection may further propagate throughout the sperm proteome and be a general mechanism to explain accelerated gametic protein evolution.

References and Notes:

1. W. J. Swanson, V. D. Vacquier, The rapid evolution of reproductive proteins. *Nature Review Genetics* **3**, 137-144 (2002).
2. H. Kokko, M. D. Jennions, R. Brooks, Unifying and Testing Models of Sexual Selection. *Annual Review of Ecology, Evolution, and Systematics* **37**, 43-66 (2006).
3. S. J. Arnold, L. D. Houck, Can the Fisher-Lande process account for birds of paradise and other sexual radiations? *The American Naturalist* **187**, 717-735 (2016).
4. R. Lande, Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Sciences* **78**, 3721-3725 (1981).

5. D. B. Wilburn, W. J. Swanson, From molecules to mating: Rapid evolution and biochemical studies of reproductive proteins. *Journal of Proteomics* **135**, 12-25 (2016).
6. D. B. Wilburn, S. J. Arnold, L. D. Houck, P. W. Feldhoff, R. C. Feldhoff, Gene duplication, co-option, structural evolution, and phenotypic tango in the courthsip pheromones of plethodontid salamanders. *Herpetologica* **73**, 206-219 (2017).
7. I. Raj *et al.*, Structural basis of egg coat-sperm recognition at fertilization. *Cell* **169**, 1315-1326. e1317 (2017).
8. W. J. Swanson, V. D. Vacquier, The abalone egg vitelline envelope receptor for sperm lysin is a giant multivalent molecule. *Proceedings of the National Academy of Sciences* **94**, 6724-6729 (1997).
9. V. D. Vacquier, Y.-H. Lee, Abalone sperm lysin: unusual mode of evolution of a gamete recognition protein. *Zygote* **1**, 181-196 (1993).
10. C. A. Lewis, C. F. Talbot, V. D. Vacquier, A protein from abalone sperm dissolves the egg vitelline layer by a nonenzymatic mechanism. *Developmental Biology* **92**, 227-239 (1982).
11. K. Haino-Fukushima, N. Usui, Purification and immunocytochemical localization of the vitelline coat lysin of abalone spermatozoa. *Developmental Biology* **115**, 27-34 (1986).
12. C. A. Lewis, D. L. Leighton, V. D. Vacquier, Morphology of abalone spermatozoa before and after the acrosome reaction. *Journal of Ultrastrcture Research* **72**, 39-46 (1980).
13. D. B. Wilburn, L. M. Tuttle, R. E. Klevit, W. J. Swanson, Solution structure of sperm lysin yields novel insights into molecular dynamics of rapid protein evolution. *Proceedings of the National Academy of Science USA* **115**, 1310-1315 (2018).
14. W. J. Swanson, V. D. Vacquier, Liposome fusion induced by a Mr 18 000 protein localized to the acrosomal region of acrosome-reacted abalone spermatozoa. *Biochemistry* **34**, 14202-14208 (1995).
15. J. E. Aagaard, V. D. Vacquier, M. J. MacCoss, W. J. Swanson, ZP domain proteins in the abalone egg coat include a paralog of VERL under positive selection that binds lysin and 18-kDa sperm proteins. *Molecular biology and evolution* **27**, 193-203 (2010).
16. N. Kresge, V. D. Vacquier, C. D. Stout, The crystal structure of a fusagenic sperm protein reveals extreme surface properties. *Biochemistry* **40**, 5407-5413 (2001).
17. M. R. Palmer *et al.*, Mass spectrometry and next-generation sequencing reveal an abundant and rapidly evolving abalone sperm protein. *Molecular Reproduction and Development* **80**, 460-465 (2013).
18. M. R. Palmer, University of Washington, Seattle, WA (2013).
19. L. Venter *et al.*, The cross-tissue metabolic response of abalone (Haliotis midae) to functional hypoxia. *Biology Open* **7**, bio031070 (2018).
20. L. Horthschild, H. Barnes, The inorganic constituents of the sea-urchin egg. *Journal of Experimental Biology* **30**, 534-544 (1953).
21. E. Rodriguez, A. Darszon, Intracellular sodium changes during the speract response and the acrosome reaction in sea urchin sperm. *Journal of Physiology* **546**, 89-100 (2003).
22. Y. Jia, X. Liu, Expression of Na+/K+-ATPase was affected by salinity change in Pacific abalone *Haliotis discus hannai*. *Frontiers in Physiology* **9**, (2018).
23. B. E. Gallindo, V. D. Vacquier, W. J. Swanson, Positive selection in the egg receptor for abalone sperm lysin. *Proceedings of the National Academy of Science USA* **100**, 4639-4643 (2003).
24. N. L. Clark *et al.*, Coevolution of interacting fertilization proteins. *PLOS Genetics* **5**, e1000570 (2009).
25. D. R. Levitan, R. Buchwalter, Y. Hao, The evolution of gametic compatibility and compatibility groups in the sea urchin Mesocentrotus franciscanus: An avenue for speciation in the sea. *Evolution* **0**.
26. J. A. Riffell, P. J. Krug, R. K. Zimmer, Fertilization in the sea. *the chemical identity of an abalone sperm attractant* **205**, 1439-1450 (2002).
27. C. J. Brown, A. K. Johnson, A. K. Dunker, G. W. Daughdrill, Evolution and disorder. *Current Opinion in Structural Biology* **21**, 441-446 (2011).

28. A. Afanasyeva, M. Bockwoldt, C. Cooney, I. Heiland, T. I. Gossmann, Human long intrinsically disordered protein regions are frequent targets of positive selection. *Genome Research*, (2018).
29. R. Pancsa, P. Tompa, Structural disorder in eukaryotes. *PLOS One* **7**, e34687 (2012).
30. N. L. Clark, J. E. Aagaard, W. J. Swanson, Evolution of reproductive proteins from animals and plants. *Reproduction* **131**, 11-22 (2006).
31. D. G. Torgerson, R. J. Kulathinal, R. S. Singh, Mammalian sperm proteins are rapidly evolving: evidence of positive selection in functionally diverse genes. *Molecular Biology and Evolution* **19**, 1973-1980 (2002).
32. H. Schägger, G. von Jagow, Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Analytical Biochemistry* **166**, 368-379 (1987).
33. B.-H. Nam *et al.*, Genome sequence of pacific abalone (Haliotis discus hannai): the first draft genome in family Haliotidae. *GigaScience* **6**, 1-8 (2017).
34. R. E. Masonbrink *et al.*, An annotated genome for *Haliotis rufescens* (red abalone) and resequenced green, pink, pinto, black, and white abalone species. *Genome Biology and Evolution* **11**, 431-438 (2019).
35. C. Camacho *et al.*, BLAST+: architecture and applications. *BMC Bioinformatics* **10**, 421 (2009).
36. R. K. Bradley *et al.*, Fast Statistical Alignment. *PLOS Computational Biology* **5**, e1000392 (2009).
37. A. Stamatakis, RAxML Version 8: A tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* **30**, 1312-1313 (2014).
38. F. Sievers *et al.*, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* **7**, 539-539 (2011).
39. Z. Yang, PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution* **24**, 1586-1591 (2007).
40. R. B. Kapust, J. Tözsér, T. D. Copeland, D. S. Waugh, The P1′ specificity of tobacco etch virus protease. *Biochemical and Biophysical Research Communications* **294**, 949-955 (2002).
41. L. W. Donaldson *et al.*, Structural characterization of proteins with an attached ATCUN motif by paramagnetic relaxation enhancement NMR spectroscopy. *Journal of American Chemical Society* **123**, 9843-9847 (2001).
42. F. Delaglio *et al.*, NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *Journal of Biomolecular NMR* **6**, 277-293 (1995).
43. W. Lee, M. Tonelli, J. L. Markley, NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. *Bioinformatics* **31**, Epub 2014 Dec 2012 (2015).
44. Y. Shen, A. Bax, Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *Journal of Biomolecular NMR* **56**, 227-241 (2013).
45. C. D. Schwieters, J. J. Kuszewski, N. Tjandra, G. M. Clore, The Xplor-NIH NMR Molecular Structure Determination Package. *Journal of Magnetic Resonance* **160**, 66-74 (2003).
46. C. D. Schwieters, J. J. Kuszewski, G. M. Clore, Using Xplor-NIH for NMR molecular structure determination. *Progress in Nuclear Magnetic Resonance Spectroscopy* **48**, 47-62 (2006).
47. G. M. Clore, C. D. Schwieters, Docking of protein-protein complexes on the basis of highly ambiguous intermolecular distance restraints derived from 1HN/15N chemical shift mapping and backbone 15N-1H residual dipolar couplings using conjoined rigid body/torsion angle dynamics. *American Chemical Society* **125**, 2902-2912 (2003).
48. E. Jurrus *et al.*, Improvements to the APBS biomolecular solvation software suite. *Protein Science* **27**, 112-128 (2018).
49. T. J. Dolinsky, J. E. Nielsen, J. A. McCammon, N. A. Baker, PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic Acids Research* **32**, W665-W667 (2004).

Acknowledgements

The authors would like to thank members of the Swanson lab for feedback on the research and/or manuscript. **Funding:** This work was supported by National Institutes of Health grants R01-HD076862 to WJS and K99-HD090201 to DBW. **Author Contributions:** The study was conceived and designed by DBW and WJS. All experiments and data analyses were performed by DBW and LMT. The manuscript drafted by DBW and edited by LMT, REK, and WJS. **Competing interests:** The authors have no competing interests. **Data and materials availability:** Sequence data has been deposited into Genbank (Accession # MN102340-MN102343) and NMR assignments have been deposited to the BMRB (27962). All other data is available in the main text or supplementary materials.

Supplementary Materials:

Materials and Methods

Figures S1-S7