The *Trichoplax* Genome and the Nature of Placozoans

Mansi Srivastava* [1], Emina Begovic[1,2], Jarrod Chapman[2], Nicholas H. Putnam[2], Uffe Hellsten[2], Takeshi Kawashima [1,3], Alan Kuo[2], Therese Mitros[1], Asaf Salamov[2], Meredith L. Carpenter[1], Ana Y. Signorovitch[4,7,9], Maria A. Moreno[4], Kai Kamm[5], Jane Grimwood[6], Jeremy Schmutz[6], Harris Shapiro[2], Igor V. Grigoriev[2], Leo W. Buss[7,8], Bernd Schierwater[4,5], Stephen L. Dellaporta[4], Daniel S. Rokhsar* [1,2]

(*) Corresponding authors

1. Center for Integrative Genomics and Department of Molecular and Cell Biology, University of California, Berkeley, CA 94720, USA.
2. Department of Energy Joint Genome Institute, Walnut Creek, CA 94598, USA.
3. Department of Zoology, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan
4. Department of Molecular, Cellular, and Developmental Biology, Yale University, New Haven, CT 06520, USA.
5. Division of Ecology and Evolution, Institut für Tierökologie und Zellbiologie, Tieraerztliche Hochschule Hannover, Bünteweg 17d, D-30559 Hannover, Germany
6. Stanford Human Genome Center, Stanford University, Stanford, CA
7. Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT 06520, USA.
8. Department of Geology and Geophysics, Yale University, New Haven, CT 06520, USA.
9. Current address: Department of Molecular and Cell Biology, Harvard University, Cambridge, MA 03128, USA.

Placozoans are arguably the simplest free-living animals, possibly evoking an early stage in metazoan evolution, yet their biology is poorly understood. Here we report the sequencing and analysis of the ~98 million base pair nuclear genome of the placozoan *Trichoplax adhaerens*. Whole genome phylogenetic analysis suggests that placozoans belong to a "eumetazoan" clade that includes cnidarians and bilaterians, with sponges as the earliest diverging animals. The compact genome exhibits conserved gene content, gene structure, and synteny relative to the human and other complex eumetazoan genomes. Despite the apparent cellular and organismal simplicity of *Trichoplax*, its genome encodes a rich array of transcription factor and signaling pathway genes that are typically associated with diverse cell types and developmental processes in eumetazoans, motivating further searches for cryptic cellular complexity and/or as yet unobserved life history stages.

Introduction

Placozoans (literally, "flat animals") are small (1-2 mm), disc-shaped creatures that were initially discovered[1] on the walls of a saltwater aquarium in the late 1800's. They were largely neglected until rediscovered[2] in the 1970's and subsequently found throughout tropical and subtropical oceans in nearshore habitats, particularly mangrove communities.[3,4] Placozoans are readily collected in the wild and can be maintained in the laboratory on diverse food sources. While placozoans found in diverse locations are morphologically indistinguishable, they exhibit surprising diversity at the DNA level, suggesting that cryptic species may exist.[5-7] The only named species in the phylum is *Trichoplax adhaerens* F.E. Schulze.[1]

Trichoplax appears as a flat disc of cells consisting of two epithelial layers, which sandwich a layer of multi-nucleate fiber cells (Figure 1a). Only four cell types have been described (Figure 1b);[8,9] nerves, sensory cells, and muscle cells are apparently absent. To feed, *Trichoplax* climbs atop its food using the bottom surface as a temporary extraorganismal gastric cavity; digestion is both extracellular and by phagocytosis.[10,11] When not feeding, the animals move by cilia on the bottom surface and by the fiber cell layer.[10] Placozoans have no evident axis of symmetry other than top versus bottom and periphery versus interior polarity, and show no regular directionality in their movement, although both positive and negative phototaxis have been observed (personal communication)[12].

In culture, *Trichoplax* reproduces by binary fission, whereby one half of the animal moves away from the other half until their connection is broken (Figure 1c-e). Sexual reproduction has never been observed in culture but putative oocyte formation in degenerating animals is routinely seen.[13] These large cells have been observed to undergo cleavage (Figure 1f-g) up to a 256-cell stage before degenerating (personal communication[14]). While Grell also described sperm,[15] they have not been seen by other investigators. Population genetic analyses, however, reveal allelic variation and evidence for genetic recombination in animals in the wild that is consistent with sex.[16]

The phylogenetic relationship of placozoans to other metazoans remains controversial. While early studies based on a small number of genes suggested that placozoans could

be secondarily simplified cnidarians,[17] other analyses refuted this position.[18,19] Small subunit rRNA analysis suggested placozoans as eumetazoans, either as sister to bilaterians,[18] or the earliest eumetazoan branch[20] (although addition of the large subunit rRNA sequences reduced the resolution of the tree). Analyses of the complete mitochondrial genomes of *Trichoplax adhaerens*[21] and other placozoans,[22] however, led to the proposal that Placozoa could be the earliest branching basal metazoan phylum (but see Ruiz-Trillo *et al*., 2008)[23].

Here we report the draft nuclear genome sequence of *Trichoplax adhaerens*, and use it to begin to address the nature of placozoans. Our phylogenetic analysis supports the identification of placozoans as a basal eumetazoan lineage that diverged prior to the separation of cnidarians and bilaterians but after the divergence of sponges from other animals. The compact genome shows remarkable complexity, including conserved gene content, gene structure, and synteny relative to the human and other eumetazoan genomes. Despite the absence of any known developmental program and only a modest number of cell types, the *Trichoplax* genome encodes a rich array of transcription factors and signaling genes that are typically associated with embryogenesis and cell fate specification in eumetazoans, as well as other genes that are consistent with cryptic patterning of cells or unobserved life history stages and/or complex execution of biological processes such as fission and embryonic development in these enigmatic creatures.

The *Trichoplax* genome

We produced a high quality draft sequence of the ~98 million base pair (Mb) *Trichoplax* genome using whole genome shotgun methods[24] with ~8-fold redundant sequence coverage. Since there are currently no genetic or physical maps of *Trichoplax,* we could not reconstruct entire chromosomes, but the completeness of the draft assembly (98% of the 14,571 expressed sequence tags aligned) and its long-range linkage (19 scaffolds longer than 1 Mb represent 80% of the assembly) make it an excellent substrate for annotation and comparative analysis (Supplementary Note S2).

As expected from genomic sequences derived from an asexually reproducing laboratory culture of diploid animals, only two alleles are observed at each locus. The single

nucleotide polymorphism averages 1% and is distributed as expected for two haplotypes selected from a panmictic sexual population (Supplementary Note S3). We observed 35 extended regions of unusually low polymorphism (<0.25% over more than 40 kb), indicating recent shared ancestry of the two haplotypes (via sex or gene conversion) and/or the influence of selective sweeps. Sampling of more than two haplotypes is required to distinguish these two possibilities.

The *Trichoplax* gene complement and conserved gene structures

We estimate that the *Trichoplax* genome contains 11,514 protein coding genes, based on a combination of homology-based and *ab initio* methods (Supplementary Note S4). Nearly 87% of these predicted genes have detectable similarity to proteins known from other animals, and most (83%) of the ~7,800 gene families that are conserved between the sea anemone and bilaterians[25] have homologs in *Trichoplax* as detected by BLAST. *Trichoplax* genes have an intron density (7.6 per kb) comparable to that found in vertebrates (8.5 per kb) and the starlet sea anemone (6.7 per kb).[25]

Analysis of the exon-intron structure of orthologous genes demonstrates a high degree of conservation in *Trichoplax* relative to other eumetazoans, extending the antiquity of many animal introns (Supplementary Note S5).[25] For example, in conserved regions, 82% of human introns have orthologous counterparts with the same position and phase in *Trichoplax*. The retention of ancient introns in *Trichopla*x is in contrast to other animals with small genomes that show extensive intron loss (e.g., fruit fly, soil nematode, and sea squirts) that presumably accompanied their reduction in genome size.[25]

Relationship of *Trichoplax* to other animals

With the complete nuclear genome in hand, we reassessed the phylogenetic position of placozoans relative to other metazoans using Bayesian, maximum likelihood, and parsimony analyses of a concatenation of 104 slowly evolving single copy nuclear genes (6,783 aligned amino acid positions) drawn from nine diverse fully sequenced genomes (Figure 2 and Supplementary Note S7). With 100% Bayesian support and 92% likelihood bootstrap support, placozoans are found to be a sister group to the other

eumetazoans (as represented by two cnidarians and a sampling of diverse bilaterians), with demosponge sequences diverging prior to the *Trichoplax*-cnidarian-bilaterian clade. This topology is further supported by parsimony analysis, albeit with weaker support. There is no support for *Trichoplax* as a derived or basal cnidarian or bilaterian, and these hypotheses are rejected by statistical phylogenetic tests (see Supplementary Note S7.6). Although there is strong likelihood bootstrap and Bayesian support for the topology in Figure 2, our analysis can only reject the placement of *Trichoplax* basal to other animals at the p=0.07 level.

Although our result disagrees with results from mitochondrial trees,[21,22,26] these analyses are complicated by the long branch lengths (*i.e.*, unusually high levels of amino acid divergence) found in bilaterian mitochondrial peptides relative to their basal metazoan orthologs.[26,27] Figure 2 shows that peptides encoded by the nuclear genome show no notable differences in amino acid substitution levels between basal metazoans and bilaterians, suggesting that our proposed phylogeny based on nuclear genes is less susceptible to long-branch attraction artifacts.

Conserved synteny with other eumetazoans

Although the placozoan lineage diverged from that of other animal phyla in the Precambrian, we find evidence for limited conserved local gene order as well as substantial blocks of longer-range conserved linkage (synteny) in the *Trichoplax* genome relative to the larger vertebrate and the starlet sea anemone genomes (Supplement Table S8.1). This is in sharp contrast to the relatively small genomes of flies and nematodes, which show no such conservation. Quantitative analysis of gene neighborhoods in *Trichoplax*, human, and *Nematostella* genomes shows that the *Trichoplax* genome has the lowest level of local rearrangement relative to the common placozoan-cnidarian-bilaterian ancestor (Supplementary Note S8.1).

The *Trichoplax* genome also exhibits larger blocks of conserved synteny (*i.e.*, conserved linkage without requiring colinearity[28]) relative to the human genome. Each of the 21 longest gene-rich *Trichoplax* scaffolds contains segments with a significant concentration of orthologs on one or more human chromosome segments (Supplementary Note S8.2). These segments are clearly visible in a dot-plot of the

Trichoplax scaffolds versus the seventeen ancestral chordate linkage groups[29] (Figure 3), which shows that (1) many of these linkages have been preserved in the *Trichoplax* genome, and (2) most of the chordate linkage groups date back to the placozoan-vertebrate last common ancestor. Neither flies nor nematodes show such conservation. For example, chordate linkage group #10 (comprising portions of human chromosomes 1q, 6p, and 9q) appears in its entirety as a relatively compact 3.2Mb segment of *Trichoplax* scaffold 2, with substantial gene order changes (Figure 3 and Supplemental Figure S8.1). The observation of blocks of conserved synteny is consistent with a relatively low rate of local rearrangement in *Trichoplax*

Genes that regulate development in Eumetazoa: transcription factors

With only four (possibly five[30]) morphologically identifiable somatic cell types, one might naively expect *Trichoplax* to possess few transcription factors associated with the complex regulation of cell fate, patterning, and differentiation found in other eumetazoans (Table 1). Nevertheless, targeted studies of homeoboxes in *Trichoplax* have identified a modest complement of these essential eumetazoan patterning genes, including paired box genes[31] and members of the Antp class.[32] Two of these (Trox2 and Not) are expressed around the rim of the animal, defining the only known molecular patterning of its body plan.[30,33]

Analysis of the draft genome shows a rich repertoire of transcription factors (Table 1 and Supplementary Note S9) commonly associated with patterning and regionalization during eumetazoan development, including additional homeobox containing genes from the ANTP, Paired, POU, and Six homeobox subfamilies.[34] *Trichoplax* also has members of many subfamilies of the animal-specific SOX (Sry-related HMG-box) family involved in the regulation of embryonic development, the T-box family including *brachyury* whose expression defines the blastopore in eumetazoan gastrulation[35], and the opisthokont (animal and fungi)-specific FOX (forkhead/winged-helix) family.

Transcription factors that regulate cell type specification and differentiation in bilaterians are also abundant in *Trichoplax*, including multiple LIM-homeobox genes typically associated with sub-type specification in neurons, multiple basic helix-loop-helix (bHLH) family genes associated with neural and muscle cell fates, a (linked) pair

of POU/homeodomain family genes implicated in neuroendocrine development in bilaterians, and a pair of GATA-family zinc finger transcription factors that participate in the specification of endodermal, cardiac, and blood cell fates in bilaterians. Thus the *Trichoplax* genome encodes a variety of nominally cell-type specific markers despite having only a few recognizable cell types, suggesting either massive redundancy of function, alternative functions not directly analogous to those in other animals, or cryptic cellular or developmental complexity.

Genes that regulate development in Eumetazoa: signaling pathways

Trichoplax lacks consistent directionality in its movements, and possesses upper-lower and center-rim axes that show no evident homology to bilaterian antero-posterior or dorso-ventral axes. Yet components of a complete Wnt/beta-catenin signaling pathway -- used for axial patterning in bilaterians and cnidarians[36] and in demosponge larvae[37] -- are present in *Trichoplax* (Figure 4a). All essential components of the BMP/TGFβ signaling pathway are also present in the *Trichoplax* genome (Figure 4b). In bilaterians BMP signaling is responsible for the establishment of the embryonic dorso-ventral axis during bilaterian development with a similar axis-defining role proposed in cnidarian[38] and demosponge larvae.[37]

We did not find evidence for a functioning hedgehog pathway, as there is no evident hedgehog ligand, Patched or Smoothened receptors, or GLI-like transcription factor.[39] Components of other signaling pathways such as the Notch and JAK/STAT pathways are present, but these pathways appear to be incomplete in that they lack molecular components critical to signal transduction (for example a Notch-like gene with a true Notch domain in the Notch pathway, or a Janus kinase in the JAK/STAT pathway) (Table S9.1). Four nuclear receptor family transcription factors are present, suggesting signaling by lipid-soluble ligands (Table 1). Elements of the animal stress response NF-kappaB pathway are found in *Trichoplax*, along with a nearly complete set of orthologs of genes typically involved in eumetazoan apoptosis (except for TNFR and Fas-receptor) (Supplementary Table S9.1).

The absence of components of some animal signaling pathways in *Trichoplax* relative to

their completeness inferred in the cnidarian-bilaterian ancestor[25] suggests that *Trichoplax* branched off from an ancestor that either did not possess all animal signaling pathways or that these genes were lost in the placozoan lineage. This latter interpretation is consistent with the presence of some of these 'missing' components in sponges (Hedgeling, EGFR, Notch).[39-41]

Molecular elements associated with neuroendocrine function

Although *Trichoplax* has no nervous system, it exhibits behavioral responses to environmental stimuli, and sensitivity to the neuropeptide RFamide has been reported.[42] In the *Trichoplax* genome we find various ion channels that are implicated in neural signaling in animals. For example, different members of the Kv family of voltage-dependent potassium channel alpha subunits (the electrically active Shaker and Shaw and the electrically inactive Kv9) and beta subunits (KCNAB) are present in the *Trichoplax* genome, along with inward rectifier K-channels and homologs of voltage-gated sodium channels and voltage-gated L-type calcium channel alpha1 subunits and their regulatory beta subunit (Figure 4c).

Components of neurotransmitter biosynthesis and vesicle transport systems, as well as a putative neuroendocrine-like secretory apparatus, are also found in the genome (Figure 4c). DOPA decarboxylase and DBH-like monooxygenase (involved in dopamine, norepinephrine and epinephrine synthesis in adrenergic cells), and putative vesicular amine transporters (used for uptake of neurotransmitters) are present. The *Trichoplax* genome encodes members of the synaptic core complex (SNAP-25, synaptobrevin and syntaxin). While synaptobrevin and syntaxin are found in diverse eukaryotic groups and are generally involved in vesicular trafficking, *Trichoplax* SNAP-25 has a distinct domain found only in animal versions of this protein and not in other eukaryotic members of this family (*e.g.*, Sec9 in yeast).[43] The genome also contains homologs of the animal neurosecretory vesicle membrane-bound proteins synaptophysin and synaptotagmin, which aid in calcium-dependent vesicle docking and fusion by interacting with SNAP-25.

Putative neurotransmitter and neuropeptide receptors are also present, including abundant seven transmembrane G-protein coupled receptors (GPCRs) that could be

candidate sensory transducers. Four putative opsin genes possess a crucial lysine residue in the 7th transmembrane domain and thus are suspected to serve a function in light reception are present. 85 members of the class 3 GPCR family (unrelated to other GPCR families by sequence), including putative metabotropic glutamate receptors, are also found. Transmembrane proteins important in nerve conduction (multiple candidate ionotropic glutamate receptors) and in neurotransmitter release and uptake (e.g. sodium neurotransmitter symporter) are encoded by the genome (Figure 4c). Synapse formation proteins (such as neurexin and neuroligin) and structural elements of the post-synaptic scaffold known to be present in sponges[44] (such as discs large) are also found in *Trichoplax*, including the receptors and channels missing from the sponge genome. Genes associated with neural migration and axon guidance in bilaterians (Slit, netrin, NCAM, semaphorin/CD100 and spondin) are also present. While all of these nominally neural elements may be functioning in non-neural roles, the *Trichoplax* genome encodes the basic machinery required for neurotransmitter synthesis, release and uptake, synapse formation, and conduction of electrical impulses and photoreception.

Extracellular matrix and cell adhesion

Trichoplax has regular cell-cell junctions between epithelial cells but is reported to lack an underlying basal lamina, or indeed, any described extracellular matrix (ECM).[2] Its genome, however, contains a diverse set of putative ECM proteins (Figure 4d). These include collagen IV, laminin-alpha, -beta, and -gamma, and nidogen, although fibronectin, fibrin, elastin, and vitronectin are apparently absent. Heparan sulfate proteoglycans (including two glypicans) and a matrilin-2 like gene are also found. Since many of these genes were also represented in the ESTs from cultured animals, an ECM may be present in a way that evades traditional histological stains.

The *Trichoplax* genome also encodes cell surface adhesion proteins (alpha and beta integrins, cadherins, selectins, and immunoglobulin superfamily members) that interact with each other and the ECM, as well as cytoskeletal linker proteins (paxillin, vinculin, talin, alpha and beta-catenin) that help organize the actin cytoskeleton and/or transduce signals in other eumetazoans (Figure 4d). Similarly, the genome encodes the protein components (focal adhesion kinase, paxillin, talin) that would permit dual functions of beta-catenin and integrin receptors in adhesion and signal transduction (through Wnt

and FAK signaling respectively). Enzyme families known to modify ECM components and/or signalling molecules in the matrix, such as lysyl oxidases, the ADAM metalloproteases (including the TACE family), and the TIMP metalloprotease inhibitor are present.

Sex and Germ Cells

Given the ancient eukaryotic origins of meiosis,[45] the production of putative oocytes by *Trichoplax*,[2] and inference of recombination in wild populations[16] it is perhaps not surprising that meiosis-associated genes are found in the genome (Table S9.1). *Trichoplax* has an ortholog of the zinc finger protein Nanos and a member of the vasa/PL10 family of DEAD box helicases, as well as homologs of mago nashi, par-1, pumilio and tudor, all implicated in primary germ cell development in eumetazoans. While these results suggest that *Trichoplax* has the same genetic tools that cnidarians and bilaterians use to segregate the germline, studies to document the expression and functions of these genes are needed to verify germline formation in placozoans.

Conclusions

Our whole genome analyses are consistent with placozoans being the earliest diverging eumetazoan phylum, *i.e.*, the sister group to the cnidarian-bilaterian clade. Although we cannot formally exclude a more basal position, our analysis rejects the derivation of placozoans from within cnidarians or bilaterians. Further studies (including additional sequences, ideally from whole genomes) will be needed to test this phylogenetic hypothesis.

Although *Trichoplax* has a compact genome relative to vertebrates and many other animals, we find that it has not experienced the same degree of intron loss and genomic rearrangement as have other small (~100 Mb) metazoan genomes (*e.g.*, the sequences of flies and soil nematodes). This suggests that many structural aspects (introns, local gene order, larger-scale linkages) of the small *Trichoplax* genome could be primitive eumetazoan characteristics.

Trichoplax's apparent genomic primitiveness, however, is separate from the question of whether placozoan morphology or life history is a relict of the eumetazoan ancestor. For example, the flat form and gutless feeding could be an "primitive" ancestral feature, with the cnidarian-bilaterian gut arising secondarily by the invention of a developmental process for producing an internal body cavity, as in Butschli's "plakula" theory,[46,47] or it could be a "derived" uniquely placozoan feature that resulted from the loss of an ancestral eumetazoan gut. Unfortunately the genome sequence alone cannot answer these questions, although it provides a platform for further studies.

While the *Trichoplax* body plan is simple, with only four or five morphologically defined cell types, its genome encodes a rich array of transcription factors and signaling pathways that are typically associated with eumetazoan developmental patterning and cell type specification. What roles do these genes play in placozoans? Cellular morphology may be deceptive, and complex gene expression patterns may define functionally distinct but morphologically cryptic cellular subtypes.[30,33,35] This would be consistent with scenarios in which transcription factors associated with gene expressions for specific differentiated cell functions in the eumetazoan ancestor were co-opted in cnidarians and bilaterians for patterning roles.[48] We speculate that signaling and transcription factor genes may be involved in complex regulatory events required for the known processes of growth, fission, and/or swarming, or the as yet undescribed processes of sexual reproduction and embryonic development (Figure S9.1).

It has been suggested that *Trichoplax* is a “living fossil” relict of an early stage of animal evolution.[9,46] At least from a genomic perspective, *Trichoplax* retains many ancestral features of its last common ancestor with cnidarians and bilaterians, which lived in the Precambrian. The extent to which the physiology, behavior, and life history of placozoans retains primitive features remains unclear. With the genome in hand, renewed interest in this "simple" animal with a complex genome will likely add to our appreciation of animal diversity and perhaps yield fundamental insights into early animal evolution.

Methods Summary

Detailed methods are described in Supplementary Information. The genome assembly, gene model sequences, predicted proteins and EST clusters and sequences can be downloaded from the JGI website http://www.jgi.doe.gov/trichoplax. Browser display of the genome sequence, including gene predictions and EST and homologous protein alignments are also available at this site. These data have been deposited with DDBJ/EMBL/GenBank as project accession ABGP00000000.

Table 1. Developmental Transcription Factors in the *Trichoplax* Genome

Transcription factor family	**Gene number in *Trichoplax***	**Subfamilies represented**
Homeobox	**35**	
ANTP-class	14	Trox-2 (Hox/ParaHox-like)[36]; Not[38]; Dlx [37]; Mnx [37]; Hmx [37]; Hex; Dbx; 7 others
paired-class	9	PaxB[42]; Pitx, Otp, Goosecoid; 5 others
POU-class (POU domain and homeobox)	2	POU class 4 (Brn-3); 1 other
LIM (lim domain and homeobox)-class	4	Islet; Apterous; Lhx1/5; 1 other
SIX (sine oculis homeobox)-class	2	Six3/6; 1 other
TALE-class	3	Pbx/Exd; Irx; MEIS
HNF-class	1	HNF
Helix loop helix	**27**	
Group A	6	Ptf; 5 others
Group B	10	Srebp; Myc; Max; Bigmax; Usf; AP4; 4 others
Group C	4	Ahr; Arnt; 2 Hif/Sim
Group D	2	2 Hes/Hey
Atonal Group	5	5 unclassified
Zinc Finger		
GATA	2	GATA-1/2/3; GATA-4/5/6
nuclear receptor	4	HNF4; retinoid X receptor; NR2; 1 other
C2H2	50	Zic; 3 SP family; 5 Klf family; Snail; Scratch; Ovo; Egr; Dpf; Gfi; MizF; Fez; Zfp277; Zfp143; Wt1; AE binding protein; 29 others
SOX (SRY-related HMG-box)	**6**	Sox8/10/E; Sox2/3; 3 other Sox; Tcf/Lef
FOX (forkhead/winged-helix)	**18**	FoxA, B, D, F, J, K, O, Q; 2 FoxN; 2 FoxG; 6 others
T-box	**5**	Brachyury[48]; Tbx2/3; 3 others
bZIP	**15**	Atf2; Atf6; Creb; Crem; Jun; Hlf; MafB; nfIL3; 7 others
ETS	**7**	Ets; Pea3; 5 others

The subfamily memberships of genes listed here were determined by phylogenetic analyses using neighbor-joining and parsimony with bootstrap methods for all families of transcription factors except bZIP and C2H2 zinc fingers. These two groups were characterized by BLAST.

Figure Legends

Figure 1. Placozoan body plan and reproduction. (a) Adult *Trichoplax* in laboratory culture (scale bar 200um). **(b)** Schematic rendering of a transverse section through *Trichoplax*: UE, upper epithelium; LE, lower epithelium; FC, contractile fiber cell; GC, gland cell; SS, shiny sphere; B, bacterium in endoplasmic cisterna; Mc, mitochondrial complex (taken from Syed and Schierwater, 2002[47]). **(c-e)** *Trichoplax* progressing through asexual reproduction by fission (scale bars 200um) **(f)** A cleaving *Trichoplax* "embryo" at the four-cell stage (scale bar 20um). **(g)** A cleaving *Trichoplax* "embryo" at the sixteen-cell stage (scale bar 20um).

Figure 2. Metazoan phylogeny and *Trichoplax*. Bayesian phylogeny of metazoans places *Trichoplax* as the sister group to cnidarians and bilaterians. Maximum parsimony applied to the same alignment results in a single tree with the same topology shown here. Posterior probabilities are reported above each branch, and likelihood bootstrap support values are reported below.

Figure 3. Conserved genomic features between the *Trichoplax* and human genomes. Blue dots represent orthologs in the 21 most gene rich *Trichoplax* scaffolds and the human genome. Human chromosomal segments have been grouped by ancestral chordate chromosome[29] (Supplementary Note S8.2). Horizontal lines divide groups of human segments descended from each of the 17 ancestral chordate chromosomes; vertical lines divide *Trichoplax* scaffolds; alternating bars outside the dotplot represent individual human segments (vertical) or *Trichoplax* scaffolds (horizontal). Red bars and dotted lines highlight the genomic regions compared in Supplemental Figure S8.1.

Figure 4. Metazoan signaling pathway and biological process genes in the *Trichoplax* genome. Known signaling and biological processes in bilaterians are shown in schematic form, with the colors of the protein names indicating their presence

(green) or absence (red) in the *Trichoplax* genome. **(a)** Wnt/b-catenin signaling, **(b)** TGF-beta signaling, **(c)** Synapse formation and conduction of nerve impulse **(d)** Cell adhesion and extracellular matrix components.

References

1. Schulze, F. E. *Trichoplax adhaerens*, nov. gen., nov. spec. *Zool Anz* **6**, 92-97 (1883).

2. Grell, K. G., Ruthmann, A. in *Placozoa, Porifera, Cnidaria and Ctenophora* (ed. Harrison, F. W., Westfall, J.A.) 13-27 (Wiley-Liss, New York, 1991).

3. Pearse, V. B. Growth and behavior of *Trichoplax adhaerens*: first recod fo the phylum Placozoa in Hawaii. *Pacific Science* **43**, 117-121 (1989).

4. Maruyama, Y. K. Occurrence in the field of a long-term, year-round, stable population of placozoans. *Biol Bull* **206**, 55-60 (2004).

5. Voigt, O. et al. Placozoa -- no longer a phylum of one. *Curr Biol* **14**, R944-5 (2004).

6. Pearse, V. B. & Voigt, O. Field biology of placozoans (*Trichoplax*): distribution, diversity, biotic interactions. *Integr. Comp. Biol.* **47**, 677-692 (2007).

7. Signorovitch, A. Y., Dellaporta, S. L. & Buss, L. W. Caribbean placozoan phylogeography. *Biol Bull* **211**, 149-56 (2006).

8. Grell, K. G. *Trichoplax adhaerens*, F.E. Schulze und die Entstehung der Metazoen. *Naturw Rundschau* **24** (1971).

9. Schierwater, B. My favorite animal, *Trichoplax adhaerens*. *Bioessays* **27**, 1294-302 (2005).

10. Ruthmann, A., Behrendt, G., Wahl, R. The ventral epithelium of *Trichoplax adhaerens* (Placozoa): Cytoskeletal structures, cell contacts and endocytosis. *Zoomorphology* **106**, 115-112 (1986).

11. Wenderoth, H. Transepithelial cytophagy by *Trichoplax adhaerens* F.E. Schulze (Placozoa) feeding on yeast. *Z Naturforsch* **41C**, 343-347 (1986).

12. Chevalerie, v. d. (2007).

13. Grell, K. G. Eibildung und Furchung von *Trichoplax adhaerens* F.E. Schulze (Placozoa). *Z Morph Tiere* **73** (1972).

14. Eitel, M. (2007).

15. Grell, K. G., Benwitz, G. Ergänzende Untersuchungen zur Ultrastruktur von *Trichoplax adhaerens* F.E. Schulze (Placozoa). *Zoomorphology* **98**, 47-67 (1981).

16. Signorovitch, A. Y., Dellaporta, S. L. & Buss, L. W. Molecular signatures for sex in the Placozoa. *Proc Natl Acad Sci U S A* **102**, 15518-22 (2005).

17. Bridge, D., Cunningham, C. W., DeSalle, R. & Buss, L. W. Class-level relationships in the phylum Cnidaria: molecular and morphological evidence. *Mol Biol Evol* **12**, 679-689 (1995).

18. Collins, A. G. Evaluating multiple alternative hypotheses for the origin of Bilateria: an analysis of 18S rRNA molecular evidence. *Proc Natl Acad Sci U S A* **95**, 15458-63 (1998).

19. Ender, A. & Schierwater, B. Placozoa are not derived cnidarians: evidence from molecular morphology. *Mol Biol Evol* **20**, 130-4 (2003).

20. da Silva, F. B., Muschner, V. C. & Bonatto, S. L. Phylogenetic position of Placozoa based on large subunit (LSU) and small subunit (SSU) rRNA genes. *Genetics Mol. Biol.* **30**, 127-132 (2007).

21. Dellaporta, S. L. et al. Mitochondrial genome of *Trichoplax adhaerens* supports placozoa as the basal lower metazoan phylum. *Proc Natl Acad Sci U S A* **103**, 8751-6 (2006).

22. Signorovitch, A. Y., Buss, L. W. & Dellaporta, S. L. Comparative genomics of large mitochondria in placozoans. *PLoS Genet* **3**, e13 (2007).

23. Ruiz-Trillo, I., Roger, A. J., Burger, G., Gray, M. W. & Lang, B. F. A phylogenomic investigation into the origin of Metazoa. *Mol Biol Evol* (2008).

24. Weber, J. L. & Myers, E. W. Human whole-genome shotgun sequencing. *Genome Res* **7**, 401-9 (1997).

25. Putnam, N. H. et al. Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science* **317**, 86-94 (2007).

26. Haen, K. M., Lang, B. F., Pomponi, S. A. & Lavrov, D. V. Glass sponges and bilaterian animals share derived mitochondrial genomic features: a common ancestry or parallel evolution? *Mol Biol Evol* **24**, 1518-27 (2007).

27. Lavrov, D. V., Forget, L., Kelly, M. & Lang, B. F. Mitochondrial genomes of two demosponges provide insights into an early stage of animal evolution. *Mol Biol Evol* **22**, 1231-9 (2005).

28. Renwick, J. H. The mapping of human chromosomes. *Annu Rev Genet* **5**, 81-120 (1971).

29. Putnam, N. H. et al. The *Amphioxus* genome and the evolution of the chordate karyotope (in press). *Nature* (2008).

30. Jakob, W. et al. The Trox-2 Hox/ParaHox gene of *Trichoplax* (Placozoa) marks an epithelial boundary. *Dev Genes Evol* **214**, 170-5 (2004).

31. Hadrys, T., DeSalle, R., Sagasser, S., Fischer, N. & Schierwater, B. The *Trichoplax* PaxB gene: a putative Proto-PaxA/B/C gene predating the origin of nerve and sensory cells. *Mol Biol Evol* **22**, 1569-78 (2005).

32. Monteiro, A. S., Schierwater, B., Dellaporta, S. L. & Holland, P. W. A low diversity of ANTP class homeobox genes in Placozoa. *Evol Dev* **8**, 174-82 (2006).

33. Martinelli, C. & Spring, J. Expression pattern of the homeobox gene Not in the basal metazoan *Trichoplax adhaerens*. *Gene Expr Patterns* **4**, 443-7 (2004).

34. Schierwater, B. et al. The ancestral Antp gene repertoire: Insights from the placozoan genome. (in press). *PLos ONE* (2008).

35. Martinelli, C. & Spring, J. Distinct expression patterns of the two T-box homologues Brachyury and Tbx2/3 in the placozoan *Trichoplax adhaerens*. *Dev Genes Evol* **213**, 492-9 (2003).

36. Lee, P. N., Kumburegama, S., Marlow, H. Q., Martindale, M. Q. & Wikramanayake, A. H. Asymmetric developmental potential along the animal-vegetal axis in the anthozoan cnidarian, Nematostella vectensis, is mediated by Dishevelled. *Dev Biol* **310**, 169-86 (2007).

37. Adamska, M. et al. Wnt and TGF-beta Expression in the Sponge *Amphimedon queenslandica* and the Origin of Metazoan Embryonic Patterning. *PLoS ONE* **2**, e1031 (2007).

38. Matus, D. Q., Thomsen, G. H. & Martindale, M. Q. Dorso/ventral genes are asymmetrically expressed and involved in germ-layer demarcation during cnidarian gastrulation. *Curr Biol* **16**, 499-505 (2006).

39. Adamska, M. et al. The evolutionary origin of Hedgehog proteins. *Curr Biol* **17**, R836-7 (2007).

40. Nichols, S. A., Dirks, W., Pearse, J. S. & King, N. Early evolution of animal cell signaling and adhesion genes. *Proc Natl Acad Sci U S A* **103**, 12451-6 (2006).

41. Muller, W. E. & Schacke, H. Characterization of the receptor protein-tyrosine kinase gene from the marine sponge Geodia cydonium. *Prog Mol Subcell Biol* **17**, 183-208 (1996).

42. Schuchert, P. *Trichoplax adhaerens* (phylum Placozoa) has cells that react with antibodies against the neuropeptide RFamide. *Acta Zoologica* **74**, 115-117 (1993).

43. Risinger, C. et al. Evolutionary conservation of synaptosome-associated protein 25 kDa (SNAP-25) shown by *Drosophila* and *Torpedo* cDNA clones. *J Biol Chem* **268**, 24408-14 (1993).

44. Sakarya, O. et al. A post-synaptic scaffold at the origin of the animal kingdom. *PLoS ONE* **2**, e506 (2007).

45. Ramesh, M. A., Malik, S. B. & Logsdon, J. M., Jr. A phylogenomic inventory of meiotic genes; evidence for sex in *Giardia* and an early eukaryotic origin of meiosis. *Curr Biol* **15**, 185-91 (2005).

46. Butschli, O. Bemerkungen zur Gastraeatheorie. *Morph Jb*. **9** (1884).

47. Syed, T., Schierwater, B. The evolution of the Placozoa: a new morphological model. *Senckenbergiana lethaea* **82**, 259-270 (2002).

48. Erwin, D. H. & Davidson, E. H. The last common bilaterian ancestor. *Development* **129**, 3021-32 (2002).

Supplementary Information is linked to the online version of the paper on **www.nature.com/nature**.

Acknowledgements This work was performed under the auspices of the US Department of Energy's Office of Science, Biological and Environmental Research Program, and by the University of California, Lawrence Livermore National Laboratory, Lawrence Berkeley National Laboratory, and Los Alamos National Laboratory. The Center for Integrative Genomics is supported by a grant from the Gordon and Betty Moore Foundation. D.S.R. acknowledges support from R.A. Melmon. L.B., S.D. ,and M.M. were

supported by the National Science Foundation. B.S. acknowledges support from the German Science Foundation.

Author Information Correspondence and requests for materials should be addressed to M.S. (e-mail: msrivast@berkeley.edu) or D.S.R. (e-mail: dsrokhsar@lbl.gov).

Online Methods

A detailed description of methods used in this study can be found in the Supplementary Information available online.

Genome Sequencing

Genomic DNA was sheared and cloned into plasmid and fosmid vectors for whole genome shotgun sequencing as described.[49] The data were assembled using release 2.10.6 of Jazz, a WGS assembler.[49] The *Trichoplax* 8X assembly and the preliminary data analysis can be downloaded from http://www.jgi.doe.gov/trichoplax and has been deposited at DDBJ/EMBL/GenBank as project accession ABGP00000000.

Gene Prediction and Annotation

The JGI annotation pipeline took scaffolds, repeats, and ESTs as inputs and produced gene models and other features that are stored in a relational database. The data can be publicly accessed through the JGI genome portal at http://www.jgi.doe.gov/trichoplax. Protein-coding gene predictions are deposited in DDBJ/EMBL/GenBank as accession ABGP00000000.

Phylogenetic Methods

104 single-copy orthologous genes from nine genomes were aligned using default parameters using both CLUSTALW[50] and MUSCLE,[51] and poorly aligned regions were excluded using Gblocks, yielding 6,783 aligned amino acid positions. Phylogenetic analyses were conducted using Bayesian inference, maximum likelihood with bootstrap and maximum parsimony with bootstrap using mrbayes,[52] PHYML,[53] and PAUP[54] respectively. Ribosomal sequences (18S, 5.8S and 28S) were added to the nuclear dataset and analyzed independently using maximum parsimony with bootstrap. Alternative likelihood topologies were tested using TREEPUZZLE[55] and CONSEL.[56]

Identification of *Trichoplax* Orthologs of Specific Bilaterian Genes

A list of *Trichoplax* gene models annotated with PANTHER hidden Markov models[57] or PFAM domains[58] was analyzed for genes involved in various biological processes in bilaterians. In many cases, *Trichoplax* orthologs of bilaterian genes were identified by BLAST against the *Trichoplax* assembly. The resulting genes were analyzed by BLAST against nr, PFAM domain composition and phylogenetic trees to determine orthology to the vertebrate query.

49. Aparicio, S. et al. Whole-genome shotgun assembly and analysis of the genome of *Fugu rubripes*. *Science* **297**, 1301-10 (2002).

50. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting,

position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**, 4673-80 (1994).

51. Edgar, R. C. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinformatics* **5**, 113 (2004).

52. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-4 (2003).

53. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**, 696-704 (2003).

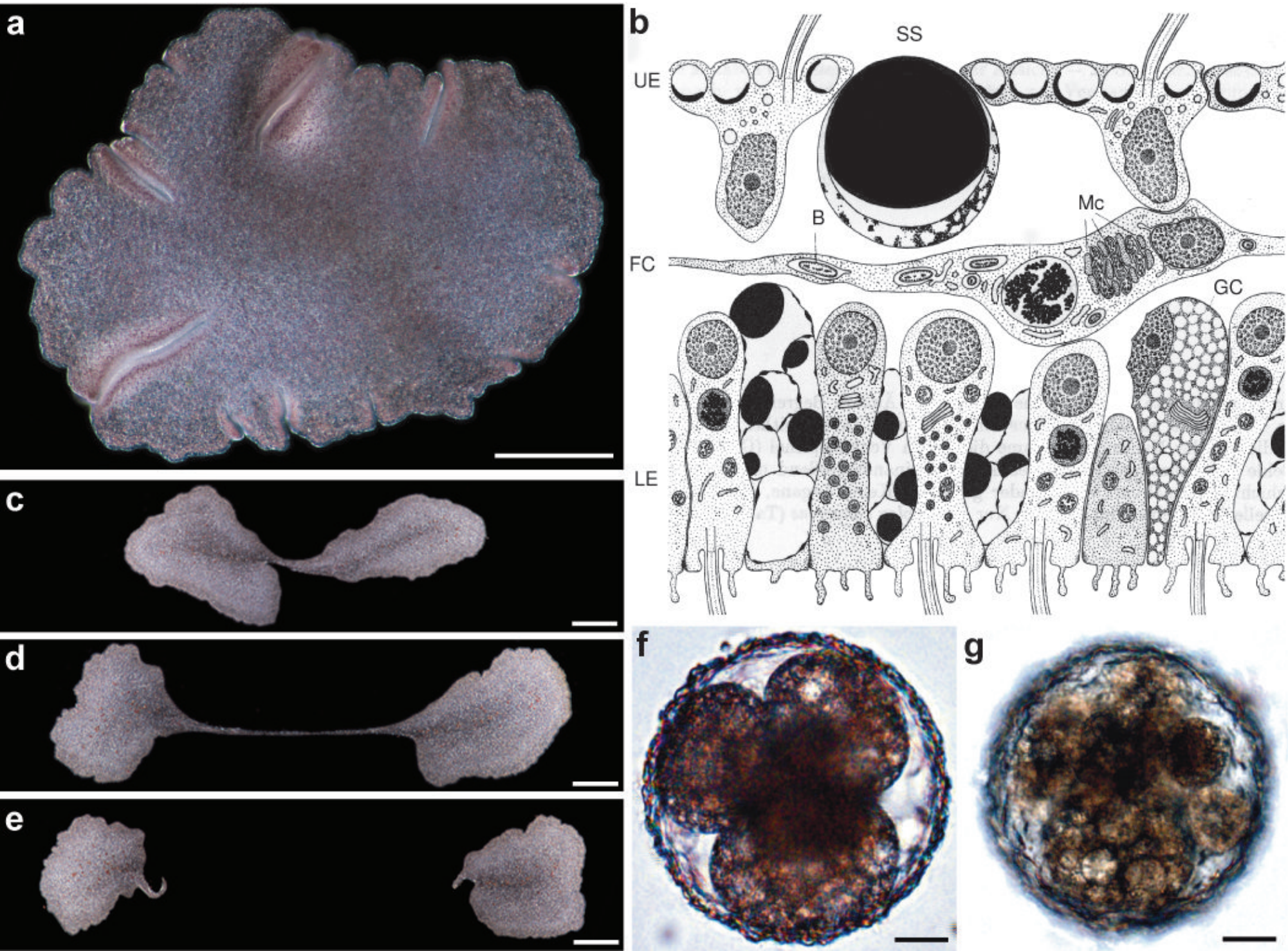
54. Swofford, D. L. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods), version 4.0*. (Sinauer Associates, Sunderland, Massachusetts, 1999).

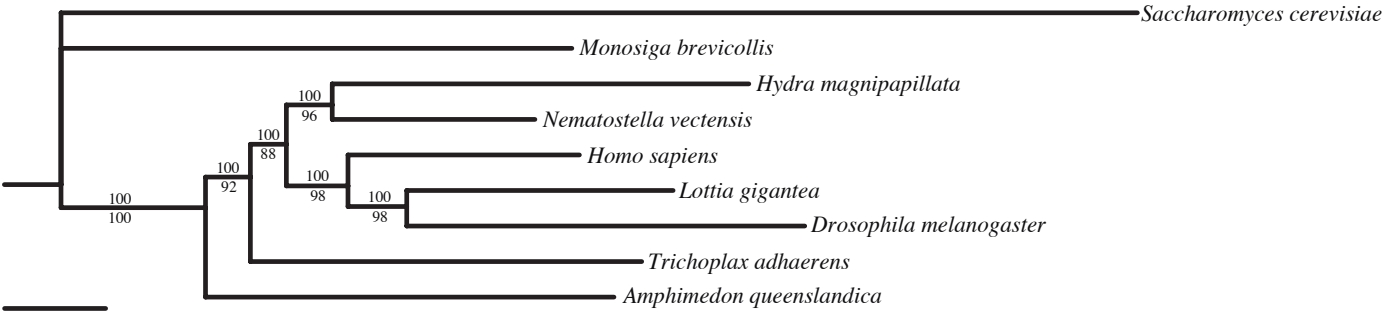
55. Schmidt, H. A., Strimmer, K., Vingron, M. & von Haeseler, A. TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* **18**, 502-4 (2002).

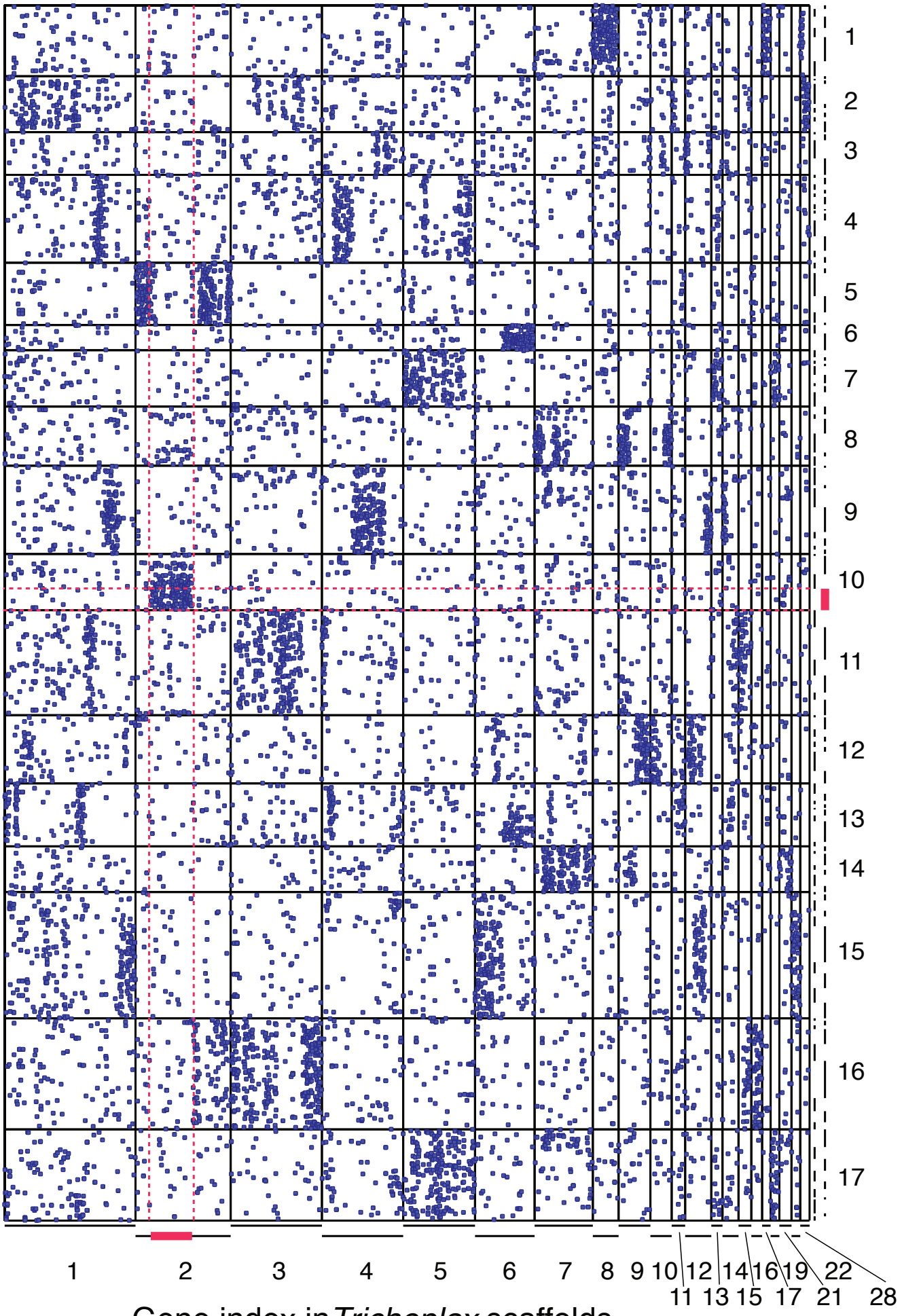
56. Shimodaira, H. & Hasegawa, M. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* **17**, 1246-7 (2001).

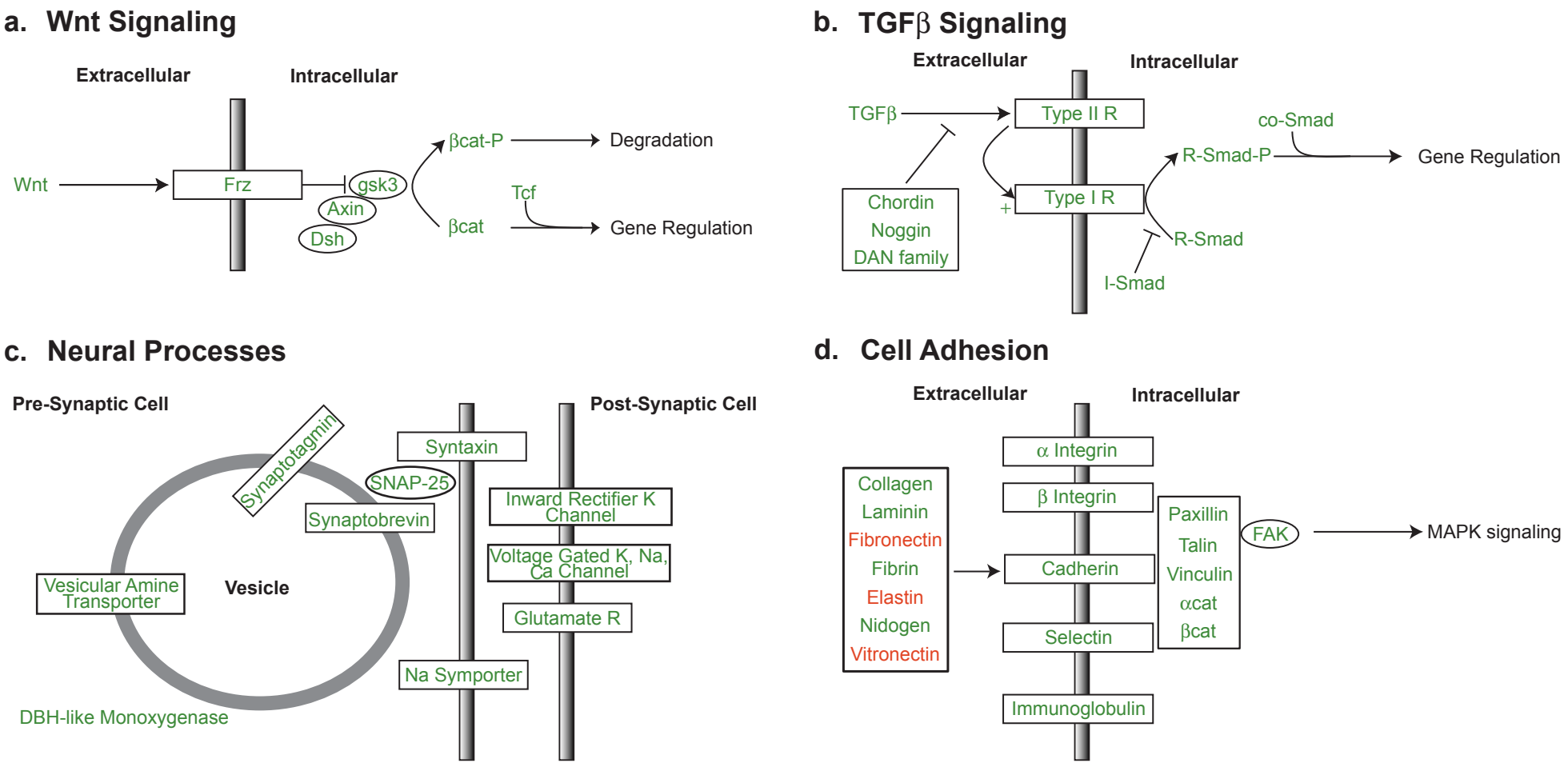
57. Thomas, P. D. et al. PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res* **13**, 2129-41 (2003).

58. Bateman, A. et al. The Pfam protein families database. *Nucleic Acids Res* **32**, D138-41 (2004).


a
b
SS
UE
B
Mc
FC
GC
LE
c
d
e
f
g


Saccharomyces cerevisiae
Monosiga brevicollis
Hydra magnipapillata
Nematostella vectensis
Homo sapiens
Lottia gigantea
Drosophila melanogaster
Trichoplax adhaerens
Amphimedon queenslandica
100
96
100
88
100
92
100
98
100
98
100
100
0.05 changes per site


1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
Gene index in human chromosome segments
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
19
21
22
28
Gene index in *Trichoplax* scaffolds


a. Wnt Signaling
Extracellular
Intracellular
Wnt
Frz
gsk3
Axin
Dsh
βcat-P
Degradation
βcat
Tcf
Gene Regulation
b. TGFβ Signaling
Extracellular
Intracellular
TGFβ
Type II R
co-Smad
R-Smad-P
Gene Regulation
Chordin
Noggin
DAN family
+
Type I R
R-Smad
I-Smad
c. Neural Processes
Pre-Synaptic Cell
Post-Synaptic Cell
Synaptotagmin
Syntaxin
SNAP-25
Synaptobrevin
Inward Rectifier K Channel
Voltage Gated K, Na, Ca Channel
Vesicular Amine Transporter
Vesicle
Glutamate R
Na Symporter
DBH-like Monoxygenase
d. Cell Adhesion
Extracellular
Intracellular
α Integrin
β Integrin
Collagen
Laminin
Fibronectin
Fibrin
Elastin
Nidogen
Vitronectin
Cadherin
Selectin
Immunoglobulin
Paxillin
Talin
Vinculin
αcat
βcat
FAK
MAPK signaling