Construction of a Large Extracellular Protein Interaction Network and Its Resolution by Spatiotemporal Expression Profiling*⑤

Stephen Martin‡, Christian Söllner‡§, Varodom Charoensawan¶, Boris Adryan¶‖, Bernard Thisse**, Christine Thisse**, Sarah Teichmann¶, and Gavin J. Wright‡‡‡

Extracellular interactions involving both secreted and membrane-tethered receptor proteins are essential to initiate signaling pathways that orchestrate cellular behaviors within biological systems. Because of the biochemical properties of these proteins and their interactions, identifying novel extracellular interactions remains experimentally challenging. To address this, we have recently developed an assay, AVEXIS (avidity-based extracellular interaction screen) to detect low affinity extracellular interactions on a large scale and have begun to construct interaction networks between zebrafish receptors belonging to the immunoglobulin and leucine-rich repeat protein families to identify novel signaling pathways important for early development. Here, we expanded our zebrafish protein library to include other domain families and many more secreted proteins and performed our largest screen to date totaling 16,544 potential unique interactions. We report 111 interactions of which 96 are novel and include the first documented extracellular ligands for 15 proteins. By including 77 interactions from previous screens, we assembled an expanded network of 188 extracellular interactions between 92 proteins and used it to show that secreted proteins have twice as many interaction partners as membrane-tethered receptors and that the connectivity of the extracellular network behaves as a power law. To try to understand the functional role of these interactions, we determined new expression patterns for 164 genes within our clone library by using whole embryo *in situ* hybridization at five key stages of zebrafish embryonic development. These expression data were integrated with the binding network to reveal where each interaction was likely to function within the embryo and were used to resolve the static interaction network into dynamic tissue- and stage-specific subnetworks within the developing zebrafish embryo. All these data were organized into a freely accessible on-line database called ARNIE (AVEXIS Receptor Network with Integrated Ex- pression; www.sanger.ac.uk/arnie) and provide a valuable resource of new extracellular signaling interactions for developmental biology.   *Molecular & Cellular Proteomics 9:2654–2665, 2010.***

The individual cells within a multicellular organism communicate with each other to provide coordinated and appropriate cellular responses that ensure the normal development and maintenance of the organism as a whole. Frequently, these intercellular dialogues are initiated by specific binding events mediated by cell surface receptor glycoproteins, the molecular bridges through which cells receive information from their immediate environment and subsequently relay it to cytoplasmic signaling networks. Despite their importance in many different biological contexts, identifying novel extracellular protein interactions remains technically challenging because membrane proteins are difficult to biochemically manipulate, and their interactions are typified by extremely low interaction strengths (1, 2). Several scalable methods to identify this class of interactions have been developed that account for some or all of these challenges: they include approaches based on protein complementation (3), phage display (4), mass spectrometry (5), surface plasmon resonance (6–9), and detection of direct binding between multivalent recombinant proteins (Refs. 10–12; for a review, see Ref. 1). Our own approach, called a̲vidity-based e̲xtracellular i̲nteraction s̲creen (AVEXIS)[1] (see Fig. 1*A*), involves expression by mammalian cells of soluble recombinant ectodomain regions of cell surface receptor and secreted proteins as either a monobiotinylated bait or a pentamerized β-lactamase-tagged prey. The pentamerization is achieved through a peptide derived from the cartilage oligomeric matrix protein (COMP) (13) and is necessary to increase the local concentration of the ectodomain fragments to effect gains in the overall binding avidity so that even very transient interactions can be detected. We have determined the parameters of the assay and shown that it can

From the ‡Cell Surface Signalling Laboratory, Wellcome Trust Sanger Institute, Cambridge CB10 1HH, United Kingdom, ¶Computational Biology Group, Medical Research Council Laboratory of Molecular Biology, Cambridge CB2 0QH, United Kingdom, and **Department of Cell Biology, University of Virginia, Charlottesville, Virginia 22908

[1] The abbreviations used are: AVEXIS, avidity-based extracellular interaction screen; ARNIE, AVEXIS Receptor Network with Integrated Expression; IgSF, immunoglobulin superfamily; LRR, leucine-rich repeat; hpf, h postfertilization; BLAST, basic local alignment search tool; ZFIN, Zebrafish Information Network; MCL, Markov clustering algorithm.

reliably detect interactions that are very transient, having monomeric half-lives of ≤0.1 s when subsequently measured in their monomeric state by surface plasmon resonance using purified proteins (10). This assay now permits the systematic screening of thousands of binary interactions and can be used to identify novel receptor-ligand pairs that were previously difficult to detect.

To identify novel extracellular receptor-ligand interactions that initiate signaling pathways important for vertebrate development, we have selected two protein families that constitute a significant proportion of the extracellular proteome: the immunoglobulin superfamily (IgSF) (14) and leucine-rich repeat (LRR) (15) families. These two families, which both contain membrane-tethered receptors and secreted proteins, are disproportionately expanded in vertebrates relative to invertebrates and are therefore likely to be involved in initiating vertebrate-specific signaling processes (16). IgSF proteins are used as cell surface recognition molecules in a diverse array of tissues, including the immune and nervous systems (17). They function by forming specific receptor-ligand pairs with other IgSF receptors but also other protein domain families, including the LRR (18). Extracellular LRR domain-containing proteins can be separated into two main families: the small leucine-rich proteoglycans, which are secreted and form a major component of the extracellular matrix (19), and cell surface receptors that are involved in the development and maintenance of the nervous system (20). To elucidate the functional role of novel interactions in vertebrate development, we have chosen to screen for novel receptor-ligand pairs from the zebrafish. This popular model organism has many experimental advantages to study early developmental processes, including the amenability to forward (21–23) and reverse (24, 25) genetics approaches. Importantly, the ready accessibility of large numbers of externally developing translucent embryos and the ability to determine large scale gene expression patterns at different stages of embryonic development make whole organism spatiotemporal gene expression profiling possible (26).

Using this approach, we first reported a network of 43 interactions within the zebrafish IgSF proteins (10) and subsequently a neural network of 34 interactions between membrane-tethered receptors from both the IgSF and LRR families (27). Here, we expanded our zebrafish recombinant library by 87 proteins to include many more secreted factors and proteins from other families and discovered 111 new interactions. By combining them with interactions identified in our previous screens, we compiled an expanded static network containing 188 interactions that we analyzed to identify properties of extracellular interaction networks. To begin the functional validation of identified interactions, we determined the developmental expression patterns of all genes encoding proteins within our library and describe here 164 new gene expression patterns, many of which are highly dynamic and tissue-restricted. By integrating the spatiotemporal expression pat-

terns with the interaction network, we were able to determine in which stages and tissues each interaction was likely to function and resolved the aggregate interaction network into stage- and tissue-specific subnetworks. The work described in this study is summarized schematically in Fig. 1*B*. To facilitate navigation of these data, they were organized into a freely accessible on-line database that provides a valuable and accessible resource of new extracellular signaling interactions for developmental biology.

EXPERIMENTAL PROCEDURES

Large Scale Extracellular Interaction Screening and Analysis—Full details for expression construct generation, protein production, and interaction screening are described elsewhere (10). Briefly, genes encoding IgSF or LRR domains were identified in the zebrafish genome sequence, amplified by RT-PCR, and cloned using the topoisomerase cloning system (Invitrogen) or, where available, were purchased as clones from the Integrated Molecular Analysis of Genomes (IMAGE) consortium. All clones were fully sequenced and submitted to GenBank™ (see supplemental Table 1). The predicted ectodomain regions, including the native signal peptide of cloned genes and of those provided by other researchers, were amplified by PCR using oligonucleotides that contain flanking NotI and AscI rare cutting restriction enzymes. The ectodomain fragments were fully sequenced and subcloned into both bait and prey expression vectors to produce rat Cd4d3+4-tagged soluble proteins. Baits contain a C-terminal BirA substrate and are monobiotinylated during expression by cotransfecting with a modified BirA expression plasmid. Preys contain a C-terminal COMP pentamerization domain N-terminal to the β-lactamase enzyme. All baits and preys were expressed by transient transfection of HEK293E cells, harvested after 6 days, filtered, dialyzed (for the baits only), and stored at 4 °C after addition of 10 mM NaN$_3$. Bait and prey proteins were normalized to threshold levels as described previously (10, 27). Protein pairs producing positive binding results from the first round of screening were produced as fresh samples and retested in an independent validation screen. The protein interactions from this publication have been submitted to the International Molecular Exchange Consortium (IMEx) (http://imex.sf.net) through IntAct (28) and assigned the identifier IM-11669; they are fully detailed in Table I. The extracellular protein interaction network was visualized using Cytoscape (29). Networks were analyzed and compared using igraph (30) within the statistical application R (31).

High Throughput Spatiotemporal Expression and Analysis—Genes within our clone library (now containing 249 different proteins) for which there was no representative image in the ZFIN expression database were identified by using BLAST against this database; in total, 164 genes had no previously documented expression pattern. Antisense probe templates were made by amplifying the entire extracellular region used in the interaction screen. Probe synthesis, *in situ* hybridization, and image capture were performed as described (26). All images were annotated using Open Biomedical Ontology-compliant terms from the "zebrafish anatomy and development" ontology (Biomedical Ontologies Foundry; http://www.obofoundry.org) as of February 2008 and have been deposited in the ZFIN database (www.zfin.org). The anatomical gene expression pattern descriptions for the 92 genes encoding proteins within the interaction network were retrieved from ZFIN, yielding 1,433 partially redundant spatiotemporal descriptions. Expression was summarized by classifying gene expression into five periods (gastrula (5.25–10 h postfertilization (hpf)), segmentation (10–24 hpf), pharyngula (24–48 hpf), hatching (48–72 hpf), and larval (72–120 hpf)), and spatially into 10 major systems (nervous and sensory, endocrine, digestive, liver, renal, skeletal, mus-

culature, cardiovascular, hematopoietic, and immune). Because expression was never observed in any reproductive or respiratory tissue, these terms were removed from further analysis; expression networks were visualized using Cytoscape (29). Zebrafish were handled in strict accordance with good animal practice as defined by the relevant national and local animal welfare bodies.

Protein Homology—Two or more proteins were grouped into the same paralogous family if identified by Ensembl Compara (32) or if they have sequence similarity greater than 50%. A neighbor-joining tree was built for the 92 proteins in the network using ClustalW2 (33) to illustrate the phylogenetic relationship between protein sequences. Orthologous proteins in other metazoan species were obtained from Ensembl Compara (species name, Ensembl release): *Caenorhabditis elegans*, 37.10; *Drosophila melanogaster*, 37.4; *Danio rerio*, 47.7; *Oryzias latipes*, 41.1; *Mus musculus*, 37.34; and *Homo sapiens*, 43.36. Additional orthologs were detected by running OrthoMCL 1.4 (34) using the default BLAST E-value cutoff of 10^{-5}, but the MCL inflation index was raised to 5 to ensure that only tight orthologous clusters were obtained. A table describing orthologous gene clusters in other species and the best homologous matches for each of the 92 proteins is available in the AVEXIS Receptor Network with Integrated Expression (ARNIE) database (www.sanger.ac.uk/arnie).

Connectivity Bias Assessment—The expected number of interactions between the proteins from the same and different families was calculated from 1,000 random family-shuffling experiments of the interaction network while maintaining the network node connectivity. The significance of the enrichment of heterophilic interactions occurring within the same paralogous cluster and the fraction of interactions that arose via a gene duplication event were assessed by randomizing the assigned paralogous clusters of the nodes in the network.

RESULTS

Identification of Novel Signaling Pathways Using AVEXIS— To identify novel signaling pathways that are important for vertebrate development, we expanded a library of recombinant zebrafish receptor and secreted proteins by 87 to a total of 249. We estimate that our protein library now contains ~40% of all non-rearranging IgSF and ~85% of the LRR families in the zebrafish genome, which accounts for around a tenth of all type I/glycosylphosphatidylinositol-linked cell surface receptors in this organism. The expression constructs for the additional proteins came from three main sources: IMAGE clones, genes amplified by RT-PCR using gene predictions from the zebrafish genome sequence, or expression clones kindly provided by other researchers (supplemental Table 1). These additional proteins were mainly members of the IgSF, but other protein families such as the small leucine-rich proteoglycans (12 proteins), Ephrins (three proteins), amyloid plaque proteins (two proteins), Netrin (one protein), and a secreted frizzled-related protein (one protein) were also included. To more closely mimic *in vivo* conditions of the extracellular environment and ensure that proteins were glycosylated and that disulfide bonds were formed, all proteins were expressed as soluble ectodomain fragments (including their own signal peptide) by transient transfection of a mammalian cell line. Proteins were produced as both biotinylated monomeric baits and enzyme-tagged pentameric preys and screened for direct, binary interactions using AVEXIS as de-

scribed (10). Importantly, because both baits and preys were expressed at levels spanning several orders of magnitude, we normalized them to a previously determined stringent activity threshold using a "gold standard" set of quantified positive and published negative interactions within the human CD2 family that we have previously shown result in low false positive rates (10). Biotinylated baits were arrayed on streptavidin-coated 96-well microtiter plates including positive and negative controls and screened against the prey library; examples of typical screening plates are shown in Fig. 1*C*. The expanded protein library described here enabled a significant increase in the size of the screen to 16,544 unique interactions tested (that is, reciprocal interactions are not counted twice), more than doubling the number of interactions tested in previous studies (6,105 (10) and 7,592 (27)). Protein pairs that showed positive interactions from the first round of screening were then retested in both bait-prey orientations using independently produced and normalized samples.

Screening of the 87 new proteins for interactions using the AVEXIS assay against both themselves and our existing protein library identified a total of 111 interactions, corresponding to an interaction detection frequency of 0.7% within these proteins; 56% of heterophilic interactions were detected in both bait-prey orientations (Table I). Fifteen interactions were expected from previous studies of orthologous proteins, demonstrating that the recombinant ectodomain fragments were correctly folded and retained their binding properties. These included interactions between secreted ligands and receptors, for example, the Robo-Slit (18) and Netrin-Neogenin interactions (35); homophilic interactions of the Kirrel family members (36); and heterophilic interactions between membrane-tethered receptors belonging to the Negr and Cadm families (37, 38). The large majority (86%) of detected interactions, however, were entirely novel, including 15 proteins that had no previously documented ligand. These interactions are listed in supplemental Table 2 and include Lrrc17, a negative regulator of osteoclast differentiation (39), and Ncam3 (40), which interacted with Neuroplastin, a receptor previously only known to interact homophilically and involved in hippocampal long term potentiation (41). Our systematic approach also identified novel ligands for well known receptors. For example, one striking feature of the network was the large number of interactions (17) with the vascular receptor Kdrl (VEGFR-2/Flk-1), which was one of the most highly connected nodes of the network. We found that the extracellular region of Kdrl could reproducibly interact in both bait-prey orientations with the two Robo3 splice isoforms present in the screen but not with Robo1, -2, or -4. This network represents a resource of novel extracellular receptor-ligand interactions, which are difficult to identify using other approaches.

Properties of Extracellular Protein Interaction Networks—To gain a more global understanding of extracellular protein interaction networks and their role in signaling, we constructed a larger network by including 77 interactions identified using
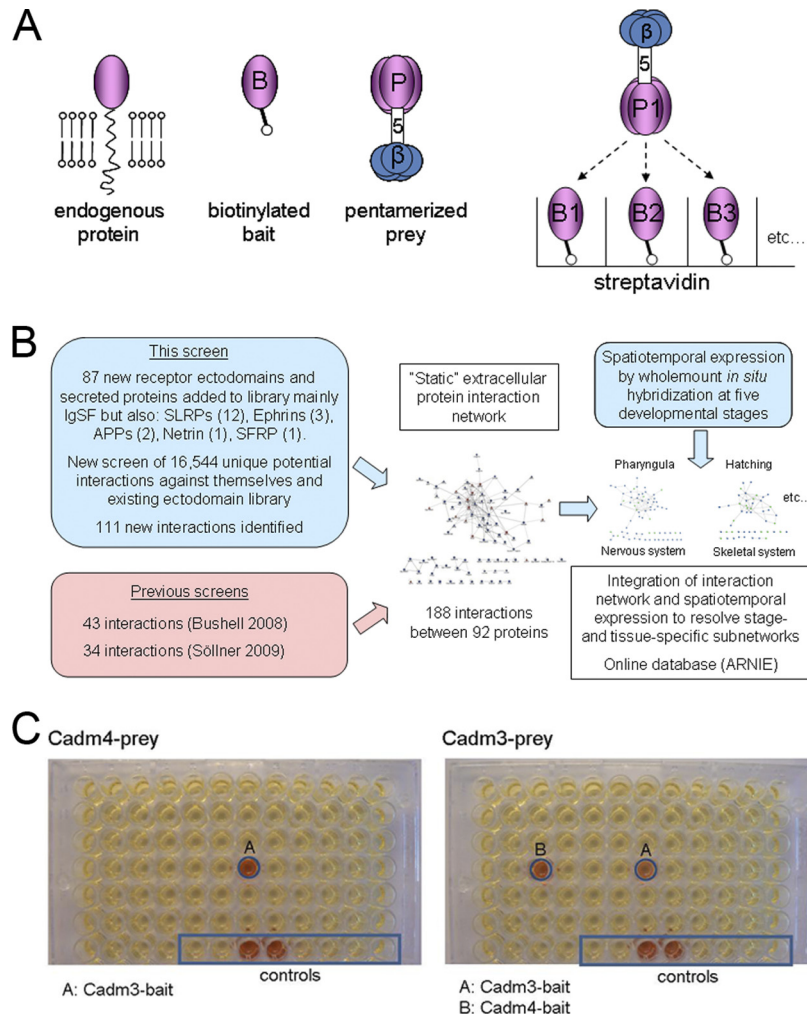
FIG. 1. **Summary of AVEXIS method and overall approach to construct and resolve extracellular protein interaction networks.** *A*, the entire ectodomains of endogenous cell surface receptors (*pink*) are expressed as both monomeric biotinylated baits (*B*) and pentamerized β-lactamase-tagged (*β*) preys (*P*). Bait proteins are immobilized in individual wells of streptavidin-coated microtiter plates and probed with a normalized prey, which, if the two proteins physically interact, is captured within the well. *B*, a flowchart presenting an overall summary of the work described here, including how interactions from two previous screens (Refs. 10 (*Bushell 2008*) and 27 (*Söllner 2009*)) were integrated into the larger network of interactions. *C*, positive interactions are detected by adding a colorimetric β-lactamase substrate, nitrocefin, which is converted from a *yellow* to *red* product. Two typical screening plates are shown illustrating a heterophilic interaction between Cadm3 and Cadm4 that is detected in both bait-prey orientations and a homophilic interaction involving Cadm3. The controls for each prey included a negative bait, the rat Cd4d3+4 protein tag alone (well G6), and a biotinylated anti-rat Cd4 monoclonal antibody (OX68), which captured the Cd4-tagged preys (well G7). Each plate also contained positive control interactions: the rat Cd200R prey was probed against rat Cd200 baits immobilized at the normalized screening threshold and at 1:500 and 1:1,000 dilutions (wells G8–G10, respectively) and against the negative Cd4d3+4 bait (well G12). An additional negative bait (Fgfr1b) was included for both preys (wells G5 and G11). *SLRPs*, small leucine-rich proteoglycans.

the same method from previous non-overlapping interaction screens that screened different bait and prey combinations (10, 27). This expanded network contained 188 interactions among 92 proteins: 157 heterophilic and 31 homophilic interactions involving 59 proteins with IgSF domains, 22 with LRR domains, seven containing both IgSF and LRR, and four "other" proteins that contained other domains including EGF, laminin, A4 extra, and β amyloid precursor protein (app) domains (Fig. 2).

We first considered whether secreted ligands or membrane-tethered receptors had different connectivity behaviors within the network and found that the average number of interaction partners for secreted proteins was 5.9 compared with 3.3 for receptors ($p < 0.01$, Wilcoxon two-sample test). The large majority (80%) of interactions made by secreted proteins involved a membrane-bound receptor, showing that most secreted ligands have multiple receptors.

The expanded network enabled us to determine the characteristics of extracellular signaling networks for the first time; for example, we found that the connectivity distribution within the network behaved as a power law (supplemental Fig. S1)

TABLE I

Summary of all interactions detected in this study

Interactions are numbered and separated into homophilic and heterophilic interactions. They are listed with their official ZFIN protein nomenclature and IntAct accession numbers for both bait-prey orientations where applicable. Interactions were considered to be of high confidence if they were positive in the primary screen and could be detected in both bait-prey orientations in either the primary or validation screens (classes A, C, D, E, and F). Interactions were classified according to Bushell *et al.* (10).

Interaction number	Interacting glycoproteins Protein 1	Protein 2	AVEXIS class	IntAct accession number Bait(1)-prey(2)	Bait(2)-prey(1)
Homophilic					
1	Cadm3		A	EBI-2267903	
2	Sc:d0647		A	EBI-2267920	
3	NadI1.2		A	EBI-2267934	
Heterophilic					
4	Cadm3	Cadm4	C	EBI-2267903	EBI-2268425
5	Pvrl2l	Zgc:158817	C	EBI-2267961	EBI-2268434
6	Kdrl	Robo3	C	EBI-2267977	EBI-2268443
7	Slit2	Robo3	D	EBI-2268001	EBI-2268452
8	Bgn	Lrrtm4l2	D	EBI-2268010	EBI-2268461
9	Zgc:92901	Ntm	E	EBI-2268026	EBI-2268470
10	Zgc:92901	Negr1	E	EBI-2268039	EBI-2268479
11	Pvrl2l	Pvrl3l	E	EBI-2268055	EBI-2268488
12	Nptn	Ncam3	E	EBI-2268064	EBI-2268497
13	Optc	Ntn1b	E	EBI-2268088	EBI-2268506
14	Optc	Kdrl	E	EBI-2268097	EBI-2268515
15	Kdrl	Epyc	E	EBI-2268106	EBI-2268524
16	Kdrl	Robo3	E	EBI-2268119	EBI-2268533
17	Kdrl	Igfbp7	E	EBI-2268128	EBI-2268542
18	Kdrl	Sc:d805	E	EBI-2268141	EBI-2268551
19	Kdrl	Lrrtm4l1	E	EBI-2268150	EBI-2268560
20	Kdrl	Boc	E	EBI-2268164	EBI-2268569
21	Kdrl	Lrrtm1	E	EBI-2268173	EBI-2268578
22	Ntn1b	Neo1	E	EBI-2268187	EBI-2268587
23	Fgfr2	Sc:d0805	E	EBI-2268200	EBI-2268596
24	Fgfr2	Boc	E	EBI-2268209	EBI-2268605
25	Fgfr2	Fgfrl1b	E	EBI-2268222	EBI-2268614
26	Fgfr3	Fgfrl1b	E	EBI-2268231	EBI-2268623
27	Fgfr3	Sc:d0805	E	EBI-2268244	EBI-2268632
28	Bcan	Sc:d0805	E	EBI-2268258	EBI-2268641
29	Bcan	Boc	E	EBI-2268267	EBI-2268650
30	Lrrtm1	Igfbp7	E	EBI-2268294	EBI-2268659
31	Lrrtm1	Sid4	E	EBI-2268276	EBI-2268668
32	Lrrtm4l1	Sid4	E	EBI-2268285	EBI-2268677
33	Lrrc24	Igfbp7	E	EBI-2268303	EBI-2268686
34	Lrrtm4l2	Sid4	E	EBI-2268316	EBI-2268695
35	Slit2	Optc	F	EBI-2268325	EBI-2268704
36	Kdrl	Zgc:77222	F	EBI-2268337	EBI-2268713
37	Kdrl	Pdgfrl	F	EBI-2268350	EBI-2268722
38	Nitr14b	Epyc	F	EBI-2268359	EBI-2268731
39	Nitr14b	Mag	F	EBI-2268372	EBI-2268740
40	Fgfr2	Lrrtm4l1	F	EBI-2268381	EBI-2268749
41	Appb	Fgfrl1a	F	EBI-2268394	EBI-2268758
42	Fgfr3	Fgfrl1a	F	EBI-2268407	EBI-2268767
43	Bgn	Pdgfrl	F	EBI-2268416	EBI-2268776
Homophilic					
44	Bcan		B	EBI-2268785	
45	Kirrel3		B	EBI-2268794	
46	Lgi1a		B	EBI-2268807	
47	Sc:d0254		B	EBI-2268820	
48	Zgc:92647		B	EBI-2268837	
49	Lrrc17		B	EBI-2268850	
50	Flrt3		B	EBI-2608410	
51	Lrrc4a		B	EBI-2608413	

Interaction number	Interacting glycoproteins Protein 1	Protein 2	AVEXIS class	IntAct accession number Bait(1)-prey(2)	Bait(2)-prey(1)
Heterophilic					
52	Cadm3	Pdgfra	G	EBI-2268863	
53	Cadm3	Igsf4d	G	EBI-2268872	
54	Cadm4	Pdgfra	G	EBI-2268885	
55	Cadm4	Igsf4d	G	EBI-2268894	
56	Pvrl2l	Zgc:123297	G	EBI-2268906	
57	Robo1	Silt1b	G	EBI-2268920	
58	Optc	Zgc:162468	G	EBI-2268934	
59	Ntn1b	Fgfr4	G	EBI-2268946	
60	Cntn1a	NadI1.2	G	EBI-2268960	
61	Appb	Fgfrl1b	G	EBI-2268969	
62	Appb	Cntfr	G	EBI-2268982	
63	Appb	Robo3	G	EBI-2268991	
64	Fgfr3	Mxra8a	G	EBI-2269004	
65	Mxra8a	Appa	G	EBI-2269013	
66	Robo3	Silt1b	G	EBI-2269026	
67	Robo3	Appa	G	EBI-2269035	
68	Zgc:56201	Mag	G	EBI-2269047	
69	Bgn	Mag	G	EBI-2269056	
70	Chad	Lrrtm4l1	G	EBI-2269069	
71	Chad	Lrrtm4l2	G	EBI-2269078	
72	Lrrtm1	Epyc	H	EBI-2269087	EBI-2269469
73	Kdrl	Fgfr2	H	EBI-2269096	EBI-2269478
74	Kdrl	Bcan	H	EBI-2269105	EBI-2269487
75	Kdrl	Fgfr3	H	EBI-2269114	EBI-2269496
76	Lingo1b	Igfbp7	H	EBI-2269123	EBI-2269505
77	Lrrtm1	Hapln2	H	EBI-2269132	EBI-2269514
78	Lrrtm1	Bgn	H	EBI-2269155	EBI-2269523
79	Slit2	Robo2	I	EBI-2269141	
80	Slit2	Robo1	I	EBI-2269164	
81	Robo1	Slit3	I	EBI-2269173	
82	Optc	Mxra8a	I	EBI-2269187	
83	Kdrl	Fgfr4	I	EBI-2269196	
84	Kdrl	Fgfr1b	I	EBI-2269209	
85	Kdrl	Flrt3	I	EBI-2269218	
86	Dcn	Boc	I	EBI-2269227	
87	Dcn	Sid4	I	EBI-2269240	
88	Bgn	Sid4	I	EBI-2269249	
89	Optc	Igfbp7	I	EBI-2269258	
90	Optc	Sid4	I	EBI-2269267	
91	Optc	Pdgfrl	I	EBI-2269276	
92	Epyc	Boc	I	EBI-2269285	
93	Epyc	Fgfrl1a	I	EBI-2269294	
94	Epyc	Sid4	I	EBI-2269303	
95	Epyc	Pdgfrl	I	EBI-2269312	
96	Chad	Sid4	I	EBI-2269321	
97	Chad	Pdgfrl	I	EBI-2269330	
98	Elfn2	Sid4	I	EBI-2269339	
99	Lrrtm2	Sid4	I	EBI-2269352	
100	Flrt3	Igfbp7	I	EBI-2269361	
101	Flrt3	Sid4	I	EBI-2269370	
102	Flrt3	Pdgfrl	I	EBI-2269379	
103	Chad	Igfbp7	J	EBI-2269388	EBI-2269532
104	Dcn	Pdgfrl	J	EBI-2269397	EBI-2269541
105	Nitr14b	Optc	J	EBI-2269406	EBI-2269550
106	Fgfr2	Lrrtm1	J	EBI-2269415	EBI-2269559
107	Fgfr2	Fgfrl1a	J	EBI-2269424	EBI-2269568
108	Lrrtm1	Pdgfrl	J	EBI-2269433	EBI-2269577
109	Lrrtm4l2	Igfbp7	J	EBI-2269442	EBI-2269586
110	Lrrtm4l1	Hapln2	L	EBI-2269451	EBI-2269595
111	Lrrtm4l2	Hapln2	L	EBI-2269460	EBI-2269604

(42). We were also able to strengthen our previous observations that extracellular protein interaction networks are enriched for interactions between paralogous proteins because 9.6% (15 of 157) of heterophilic interactions occurred within the same paralogous cluster compared with 1.09 $\pm$ 0.8% (S.D.; $p < 0.001$) observed within randomized paralogous clusters. We also determined that over half (51.6%) of the interactions within the network could be explained by gene
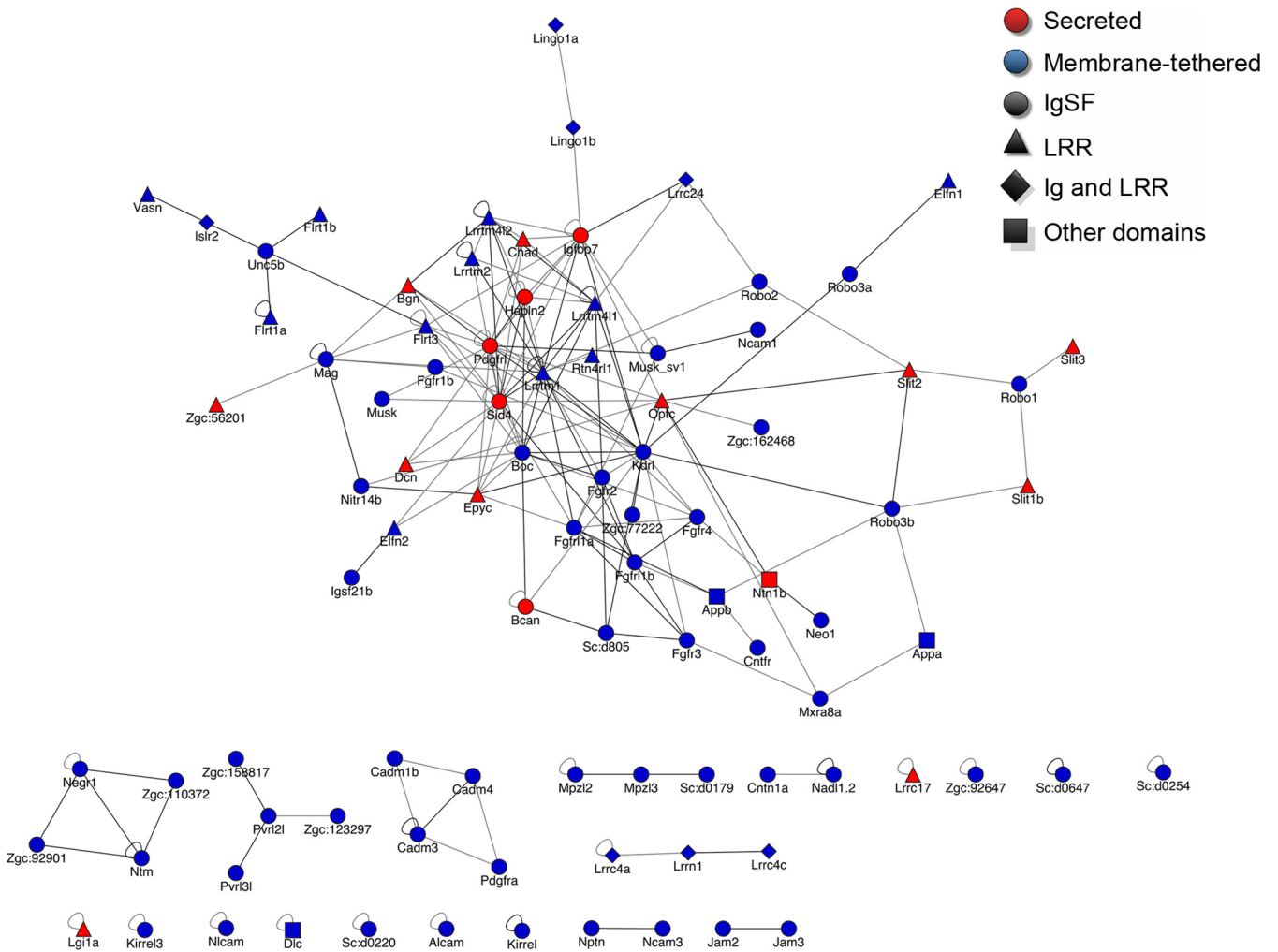
FIG. 2. **Large extracellular protein interaction network systematically determined by AVEXIS.** The interaction network consists of 188 interactions between 92 proteins. The family to which each protein belongs and its predicted subcellular localization are indicated: *red*, secreted; *blue*, membrane-tethered; *circle*, IgSF; *triangle*, LRR; *diamond*, IgSF + LRR; *square*, other domains.

duplication of one or both of the interacting partners, a figure that is much higher than expected by chance alone (7.7 $\pm$ 2.2% (S.D.; $p <$ 0.001)), and is a known property of other protein interaction networks (43).

Spatiotemporal Gene Expression Profiling Reveals Dynamic Expression of IgSF and LRR Genes during Early Vertebrate Development—To obtain functional information for the interactions within the network, we determined the expression profiles of all the genes within our protein library, for which no expression pattern had been previously determined, using whole-mount *in situ* hybridization from gastrula to larval periods of zebrafish embryonic development. In total, we determined 164 new expression patterns, which have been deposited in the publically accessible ZFIN database (44). A complete description of the principal events within these developmental periods is detailed in Kimmel *et al.* (45): they cover the initiation and completion of primary organogenesis (segmentation to hatching) and the development of circulation

(pharyngula) and simple behaviors (larval) (Fig. 3*A*). To enable a summarized, systematic analysis of the expression data, the gene expression patterns were annotated using a restricted vocabulary and subsequently categorized into 10 general biological systems using the zebrafish anatomical ontology.

We first compared the number of genes expressed in each tissue across all periods of development. This revealed that the large majority of the genes in the network were expressed within the nervous and sensory system (Fig. 3*B*) with almost twice as many genes expressed in this tissue compared with the next highest, the skeletal system. We then compared the expression of the genes at different developmental periods, noting whether the expression of that gene was ubiquitous or had a tissue-restricted expression pattern. We observed that the number of expressed genes gradually increased during early development, peaking at the pharyngula period (Fig. 3*C*), and with the exception of the gastrula stage, their expression patterns were largely restricted to particular tissues. We also
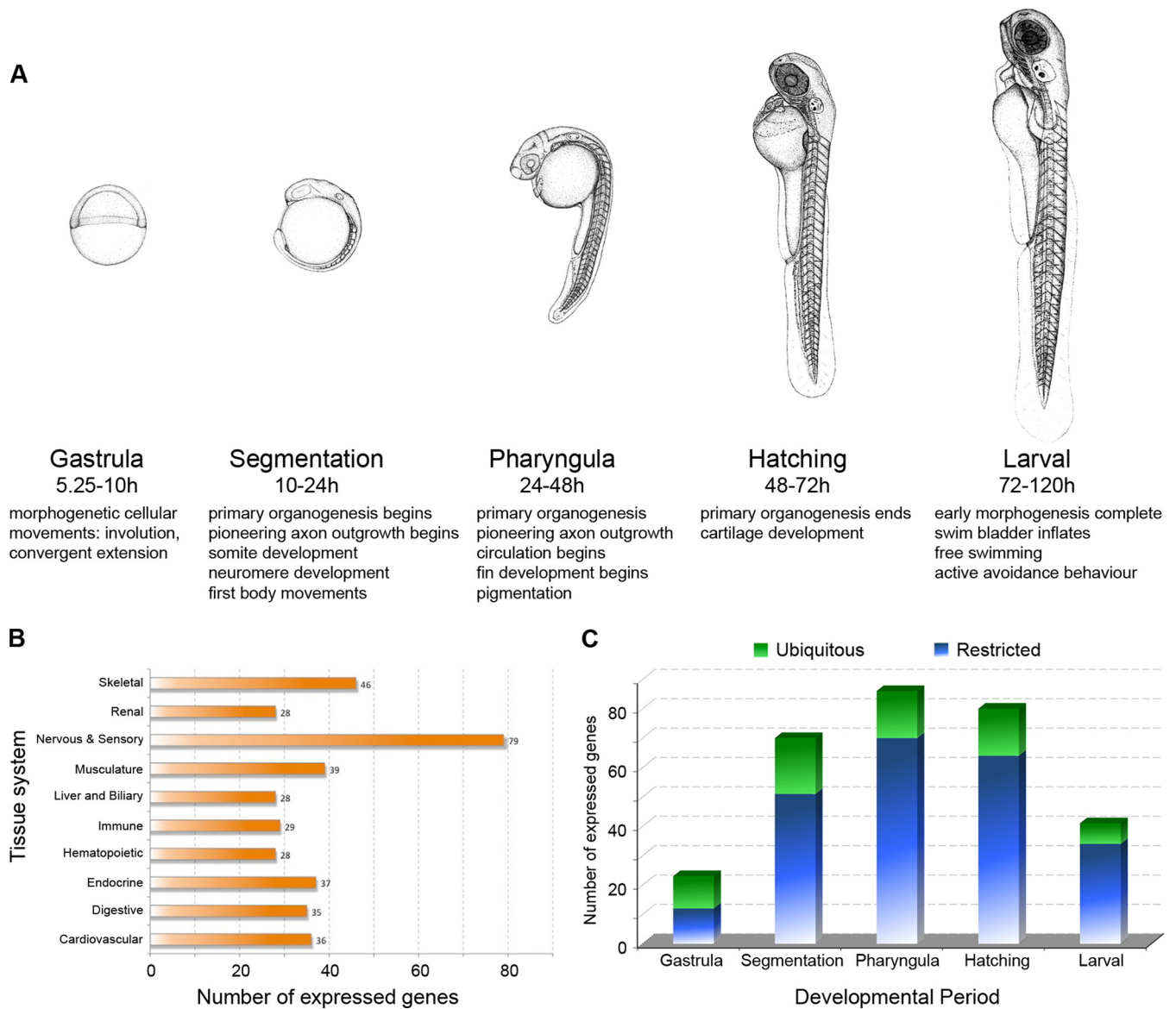
FIG. 3. **Genes encoding proteins within extracellular interaction network are expressed in tissue-restricted manner during early development.** *A*, the development of the zebrafish embryo. Drawings of representative stages within each of the five main periods of zebrafish development are shown *above* a brief description of the main morphogenetic landmarks within each period. Drawings are taken with permission from Kimmel *et al.* (45). *B*, a summary of the tissue expression patterns of the genes encoding proteins from the interaction network. The expression patterns were annotated and placed into all the appropriate non-exclusive tissue categories. *C*, a summary of the temporal expression patterns of the genes encoding proteins from the interaction network. The number of genes expressed at each period of development is plotted, indicating those genes whose expression was ubiquitous (*green*) or restricted to particular tissues (*blue*).

observed that closely related paralogs usually had very different expression patterns, suggesting that the control of gene expression can evolve rapidly after gene duplication. This analysis showed an increased need for IgSF and LRR receptor-mediated cell surface recognition functions in the nervous and sensory system relative to other tissues as development proceeds. Overall, the analysis showed that the IgSF and LRR gene families were dynamically expressed in all tissues during embryonic development: only one gene lacked detectable expression (*nitr14b*), and just three genes (*sc:d0647*, *lgi1a*,

and *zgc:77222*) were expressed constitutively (all tissues at all tested periods).

Integration of the Gene Expression Patterns with the Interaction Network Identifies Tissue Specificity of Interactions—A prerequisite for initiating signals through membrane receptor proteins is that their genes must be compatibly expressed in overlapping or adjacent cells at the same time. To identify for each interaction the tissues and developmental stages for which this criterion was met, we integrated the expression data with the interaction network by highlighting where each

gene pair encoding interacting proteins was spatially and temporally co-expressed (Fig. 4). Of the 157 heterophilic interactions identified, 76% were compatible based on spatiotemporal co-expression. Interactions were clustered into those that functioned in specific tissues or throughout the whole organism and revealed that over half (63%) of the interactions were restricted to a particular system. In the vast majority (81%) of these cases, productive interactions were localized within the nervous and sensory systems, highlighting the role of these receptor families in cellular recognition events during the development of these tissues. Thirteen interactions involved genes expressed throughout the entire organism (two were constitutive, and 11 were stage-specific) and are likely to have general roles in adhesion rather than regulated intercellular communication (Fig. 4). This analysis also revealed that 30 interactions had spatially and/or temporally incompatible gene expression patterns during embryonic development. Remarkably, in all cases where the genes were spatially incompatible, at least one member of the pair was predicted to be a secreted protein for which one might not expect to see precise spatial colocalization by *in situ* hybridization. Likely reasons for observing incompatible gene expression include the incomplete nature of the protein library, the perdurance of a protein product where the corresponding mRNA has since degraded, and simply non-functional "leaky" expression.

We next asked whether genes encoding interacting proteins were spatiotemporally co-expressed more often than one would expect by chance alone. To do this, we excluded all homophilic interactions (which always have compatible expression) and those involving at least one secreted protein (where one might not expect to see compatible spatial expression), leaving 81 heterophilic interactions of which only three have incompatible gene expression patterns. The expression for each gene within the 81 heterophilic interactions was randomized within the annotated stages and tissues. We observed that within the randomized networks the average number of compatible interactions was 55.9 ± 3.3 (S.D.; empirical *p* value <0.001), significantly less than the observed 78 compatible interactions, demonstrating a significant enrichment in compatible gene expression.

Spatiotemporal Expression Profiling Resolves Stage- and Tissue-specific Receptor Signaling Networks—The interaction network shown in Fig. 2 is a static aggregate of all possible IgSF- and LRR-mediated signaling networks that might occur *in vivo* at different developmental stages and in different tis-
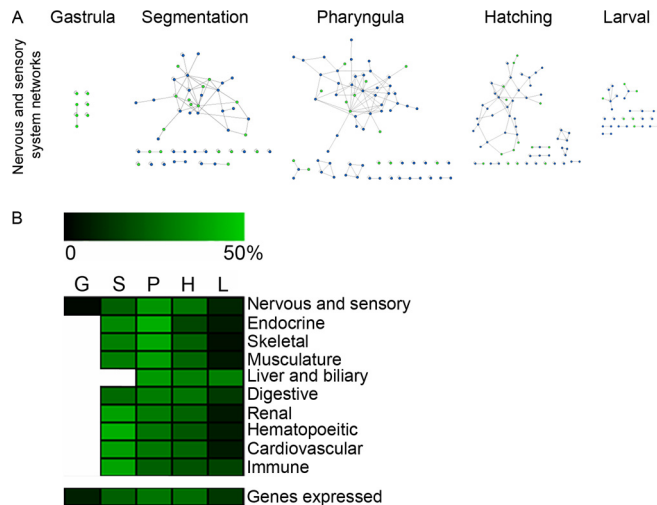

Fig. 5. **Resolution of extracellular interaction network into stage- and tissue-specific signaling networks.** *A*, graphs show time-resolved extracellular interaction networks at each period of development within the nervous and sensory system. The node key is as follows: *blue*, gene whose expression is restricted to nervous system; *green*, ubiquitously expressed gene at the stage considered. *B*, heat map showing the percentage of extracellular interaction network edges at a given developmental period relative to the others within each tissue. For comparison, the percentage of genes expressed at each period is also shown. *G*, gastrula; *S*, segmentation; *P*, pharyngula; *H*, hatching; *L*, larval.

sues. To show how the extracellular network changed in different tissues during development, we used the expression data to deconvolute the static composite network into more physiologically relevant stage- and tissue-specific signaling subnetworks. As an example, we show how the receptor interaction network changes at different stages during the development of the nervous and sensory system (Fig. 5*A*). To provide a simplified overview of how the interaction networks change in each tissue system during development, we calculated the percentage of all interactions within each developmental period and compared them by plotting a heat map (Fig. 5*B*). This showed that within half the tissues, including the nervous and sensory, skeletal, and endocrine systems, the complexity of the network paralleled the relative number of genes expressed at that stage. In the remaining tissues, however, including the renal, hematopoietic, and immune systems, the interaction networks were relatively more complex at the segmentation stage.

ARNIE, an On-line Database to Navigate Extracellular Interaction Networks and Expression Patterns—By comparing the

Fig. 4. **Integration of spatiotemporal gene expression profiles with extracellular protein interaction network.** All 188 gene pairs that encode interacting proteins are listed *vertically*; those encoding a secreted protein are highlighted in *green*. Each gene of a pair is arbitrarily assigned 1 (*left column*, *purple*) or 2 (*right column*, *gold*), and their spatiotemporal expression patterns are indicated by *shading* an appropriate *box* in the matrix, which is organized into 10 organ systems (nervous and sensory to immune), each of which is subdivided in up to five developmental stages as appropriate (*G*, gastrula; *S*, segmentation; *P*, pharyngula; *H*, hatching; *L*, larval). Where both genes are compatibly expressed, the *box* is shaded *red*. Interactions are first organized into those whose gene pairs are compatibly or incompatibly expressed and then divided further into functional subcategories as indicated.

detailed spatiotemporal expression patterns of genes that encode interacting cell surface receptor and secreted proteins, hypotheses regarding the function of the interaction can be formulated and tested. To facilitate this comparison and provide easy access to the interactions and expression patterns identified in the network, we created a publically accessible on-line database called ARNIE (http://www.sanger.ac.uk/arnie). This database allows users to perform rapid side-by-side comparisons of the detailed whole mount *in situ* expression images, which are organized into stage- and orientation-matched pairs. ARNIE also enables users to interactively resolve the overall network into individual stage- and/or tissue-specific signaling subnetworks using drop-down menus. To extend the application of the protein interaction network to other organisms, this database also includes orthology mapping of all the proteins within the network to orthologous proteins in other popular model organisms and humans.

DISCUSSION

Interactions made by membrane-embedded receptor signaling proteins have important roles in both genetic and infectious diseases and yet remain technically challenging to identify on a large scale due to transient (monomeric half-lives <1 s) interaction strengths and the requirement for structurally critical posttranslational modifications. By using a specialized assay, AVEXIS, and a protein library containing the ectodomain fragments of receptors and secreted proteins mainly from the zebrafish IgSF and LRR families, we have systematically constructed and analyzed a network of extracellular interactions. This extends our previous studies by screening at a much larger scale and including many more secreted proteins and other protein families. By combining these networks, our aim was to identify novel signaling pathways that are important for vertebrate developmental processes. To take the first steps toward elucidating the function of these interactions, we have determined the spatiotemporal expression pattern of each gene within our protein library using whole-mount *in situ* hybridization during early development. By integrating the expression with the interaction network, we were able to transform the static network into a dynamic signaling map and resolve the interactions into tissue- and stage-specific functions. In the vast majority (81%) of these cases, productive interactions were localized within the nervous and sensory systems, highlighting the role of these receptor families in cellular recognition events during the development of these tissues. Future work from our laboratory will be focused toward identifying the functional role of the interactions identified from this and our previous screens using the genetic advantages of the zebrafish. The integrated interaction network and expression patterns have been organized into an on-line database and together are a useful resource of novel signaling interactions for developmental biology.

All our screens have been performed at a stringent prey threshold that was initially determined using a set of quantified positive and published negative interactions within the human CD2 family (10). This stringency threshold enabled the detection of 15 known orthologous heterophilic interactions. However, the screen did not identify some known homophilic interactions, including that involving Ncam and one known orthologous heterophilic interaction (Netrin1b-Unc5b). We have previously reported that homophilic interactions constitute a class of false negatives using the AVEXIS technique most likely due to prey-prey associations (10). The inability to detect the Unc5b-Netrin1b interaction is unlikely to be due to incorrect folding of the recombinant proteins because we detect known Unc5b interactions with Flrt receptors (27, 46) and the Netrin1b-Neogenin interaction (35). The zebrafish genome contains two *netrin1* paralogues, *1a* and *1b*, and it remains possible that because of this redundancy Netrin1b no longer interacts with Unc5b. Because Netrin1a was not sufficiently expressed, we were unable to test whether Netrin1a interacted with Unc5b (supplemental Table 1).

In comparison with cytoplasmic protein interaction networks, extracellular networks are less well characterized. Largely, this is due to the fact that current popular high throughput protein interaction mapping assays such as the yeast two-hybrid and tandem affinity purification tagging approaches, which have been very successful at identifying cytoplasmic protein interactions, are generally unsuitable to detect transient extracellular interactions (47). This study, together with our previous work, represents one of the few attempts to systematically build extracellular protein interaction networks and shows that our approach can be used to screen for interactions within recombinant protein libraries containing hundreds of different proteins. Related approaches from others have focused on identifying homophilic interactions between different alternative splice isoforms of the *Drosophila* Dscam neural recognition molecule (12) or between extracellular matrix proteins and their receptors centered around endostatin (6).

The integration of expression data with protein interaction networks has been successfully used in unicellular organisms to reveal novel context-dependent functional complexes (48, 49) and to gain insight into how gene co-expression and network topologies relate to one another (50, 51). In multicellular organisms, high throughput spatiotemporal promoter activity profiling has been used to provide additional evidence that two proteins could indeed interact *in vivo* or conversely show that they were unlikely to interact (52), or to explain how ubiquitously expressed proteins can have tissue-specific functions (53). The integrated interaction and expression map, which has been organized into an on-line database, can also be used to resolve which interactions could function at a particular time and place during embryonic development. In the cases where genes encoding interacting proteins were not compatibly expressed, the interaction could have a function

at later stages of development, or the perdurance of one of the protein products expressed earlier might enable a productive interaction to occur. The interaction and expression data can be used to form testable hypotheses regarding the function of interactions that we have identified.

During the course of our analysis, we noticed that the expression patterns of even closely related paralogous proteins were often very different. By using our expression and interaction data sets, we were able to find support for a model where the evolution of signaling pathways after gene duplication is more likely to occur by changes in the spatiotemporal expression of receptors and then by divergence in cytoplasmic signaling sequences, whereas extracellular sequences and interactions are more conserved. This analysis is reported separately (54).

We have demonstrated here that by using the AVEXIS technique it is now possible to begin constructing large systematic extracellular protein interaction networks involving membrane-embedded receptor and secreted proteins, which have so far been a challenge to identify experimentally on this scale. By integrating these extracellular networks with gene distribution data, maps of the intercellular recognition processes used by multicellular organisms to coordinate their cellular behaviors can be built. Critically, these extracellular interactions will provide the means to link up the currently more extensive intracellular protein interaction networks (55) and build models of how neighboring cells within a multicellular organism respond to external stimuli.

Ⓢ This article contains supplemental Fig. S1 and Tables 1 and 2.

§ Present address: Max Planck Inst. for Developmental Biology, Dept. 3 (Genetics), Spemannstraße 35, 72076 Tübingen, Germany.

‖ Present address: Cambridge Systems Biology Centre, University of Cambridge, Cambridge CB2 1QR, UK.

‡‡ To whom correspondence should be addressed. Tel.: 44-1223-496852; Fax: 44-1223-496802; E-mail: gw2@sanger.ac.uk.

REFERENCES

1. Wright, G. J. (2009) Signal initiation in biological systems: the properties and detection of transient extracellular protein interactions. *Mol. Biosyst.* **5,** 1405–1412
2. van der Merwe, P. A., and Barclay, A. N. (1994) Transient intercellular adhesion: the importance of weak protein-protein interactions. *Trends Biochem. Sci.* **19,** 354–358
3. Urech, D. M., Lichtlen, P., and Barberis, A. (2003) Cell growth selection system to detect extracellular and transmembrane protein interactions. *Biochim. Biophys. Acta* **1622,** 117–127
4. de Wildt, R. M., Tomlinson, I. M., Ong, J. L., and Holliger, P. (2002) Isolation of receptor-ligand pairs by capture of long-lived multivalent interaction complexes. *Proc. Natl. Acad. Sci. U.S.A.* **99,** 8530–8535
5. Cain, S. A., McGovern, A., Small, E., Ward, L. J., Baldock, C., Shuttleworth, A., and Kielty, C. M. (2009) Defining elastic fiber interactions by molecular fishing: an affinity purification and mass spectrometry approach. *Mol. Cell. Proteomics* **8,** 2715–2732
6. Faye, C., Chautard, E., Olsen, B. R., and Ricard-Blum, S. (2009) The first draft of the endostatin interaction network. *J. Biol. Chem.* **284,** 22041–22047
7. Jiang, L., and Barclay, A. N. (2010) Identification of leucocyte surface protein interactions by high-throughput screening with multivalent reagents. *Immunology* **129,** 55–61
8. Gonzalez, L. C., Loyet, K. M., Calemine-Fenaux, J., Chauhan, V., Wranik, B., Ouyang, W., and Eaton, D. L. (2005) A coreceptor interaction between the CD28 and TNF receptor family members B and T lymphocyte attenuator and herpesvirus entry mediator. *Proc. Natl. Acad. Sci. U.S.A.* **102,** 1116–1121
9. Jiang, L., and Barclay, A. N. (2009) New assay to detect low-affinity interactions and characterization of leukocyte receptors for collagen including leukocyte-associated Ig-like receptor-1 (LAIR-1). *Eur. J. Immunol.* **39,** 1167–1175
10. Bushell, K. M., Söllner, C., Schuster-Boeckler, B., Bateman, A., and Wright, G. J. (2008) Large-scale screening for novel low-affinity extracellular protein interactions. *Genome Res.* **18,** 622–630
11. Voulgaraki, D., Mitnacht-Kraus, R., Letarte, M., Foster-Cuevas, M., Brown, M. H., and Barclay, A. N. (2005) Multivalent recombinant proteins for probing functions of leucocyte surface proteins such as the CD200 receptor. *Immunology* **115,** 337–346
12. Wojtowicz, W. M., Wu, W., Andre, I., Qian, B., Baker, D., and Zipursky, S. L. (2007) A vast repertoire of Dscam binding specificities arises from modular interactions of variable Ig domains. *Cell* **130,** 1134–1145
13. Tomschy, A., Fauser, C., Landwehr, R., and Engel, J. (1996) Homophilic adhesion of E-cadherin occurs by a co-operative two-step interaction of N-terminal domains. *EMBO J.* **15,** 3507–3514
14. Barclay, A. N. (2003) Membrane proteins with immunoglobulin-like domains–a master superfamily of interaction molecules. *Semin. Immunol.* **15,** 215–223
15. Dolan, J., Walshe, K., Alsbury, S., Hokamp, K., O'Keeffe, S., Okafuji, T., Miller, S. F., Tear, G., and Mitchell, K. J. (2007) The extracellular Leucine-Rich Repeat superfamily; a comparative survey and analysis of evolutionary relationships and expression patterns. *BMC Genomics* **8,** 320
16. Vogel, C., and Chothia, C. (2006) Protein family expansions and biological complexity. *PLoS Comput. Biol.* **2,** e48
17. Rougon, G., and Hobert, O. (2003) New insights into the diversity and function of neuronal immunoglobulin superfamily molecules. *Annu. Rev. Neurosci.* **26,** 207–238
18. Kidd, T., Bland, K. S., and Goodman, C. S. (1999) Slit is the midline repellent for the robo receptor in Drosophila. *Cell* **96,** 785–794
19. McEwan, P. A., Scott, P. G., Bishop, P. N., and Bella, J. (2006) Structural correlations in the family of small leucine-rich repeat proteins and proteoglycans. *J. Struct. Biol.* **155,** 294–305
20. Chen, Y., Aulia, S., Li, L., and Tang, B. L. (2006) AMIGO and friends: an emerging family of brain-enriched, neuronal growth modulating, type I transmembrane proteins with leucine-rich repeats (LRR) and cell adhesion molecule motifs. *Brain Res. Rev.* **51,** 265–274
21. Driever, W., Solnica-Krezel, L., Schier, A. F., Neuhauss, S. C., Malicki, J., Stemple, D. L., Stainier, D. Y., Zwartkruis, F., Abdelilah, S., Rangini, Z., Belak, J., and Boggs, C. (1996) A genetic screen for mutations affecting embryogenesis in zebrafish. *Development* **123,** 37–46
22. Driever, W., Stemple, D., Schier, A., and Solnica-Krezel, L. (1994) Zebrafish: genetic tools for studying vertebrate development. *Trends Genet.* **10,** 152–159
23. Haffter, P., and Nüsslein-Volhard, C. (1996) Large scale genetics in a small vertebrate, the zebrafish. *Int. J. Dev. Biol.* **40,** 221–227
24. Nasevicius, A., and Ekker, S. C. (2000) Effective targeted gene 'knockdown' in zebrafish. *Nat. Genet.* **26,** 216–220
25. Wienholds, E., Schulte-Merker, S., Walderich, B., and Plasterk, R. H. (2002) Target-selected inactivation of the zebrafish rag1 gene. *Science* **297,** 99–102

26. Thisse, C., and Thisse, B. (2008) High-resolution in situ hybridization to whole-mount zebrafish embryos. *Nat. Protoc.* **3,** 59–69

27. Söllner, C., and Wright, G. J. (2009) A cell surface interaction network of neural leucine-rich repeat receptors. *Genome Biol.* **10,** R99

28. Kerrien, S., Alam-Faruque, Y., Aranda, B., Bancarz, I., Bridge, A., Derow, C., Dimmer, E., Feuermann, M., Friedrichsen, A., Huntley, R., Kohler, C., Khadake, J., Leroy, C., Liban, A., Lieftink, C., Montecchi-Palazzi, L., Orchard, S., Risse, J., Robbe, K., Roechert, B., Thorneycroft, D., Zhang, Y., Apweiler, R., and Hermjakob, H. (2007) IntAct—open source resource for molecular interaction data. *Nucleic Acids Res.* **35,** D561–D565

29. Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13,** 2498–2504

30. Csárdi, G., and Nepusz, T. (2009) *igraph*, MTA RMKI, Budapest, Hungary

31. R Development Core Team (2006) *R, a Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria

32. Flicek, P., Aken, B. L., Beal, K., Ballester, B., Caccamo, M., Chen, Y., Clarke, L., Coates, G., Cunningham, F., Cutts, T., Down, T., Dyer, S. C., Eyre, T., Fitzgerald, S., Fernandez-Banet, J., Gräf, S., Haider, S., Hammond, M., Holland, R., Howe, K. L., Howe, K., Johnson, N., Jenkinson, A., Kähäri, A., Keefe, D., Kokocinski, F., Kulesha, E., Lawson, D., Longden, I., Megy, K., Meidl, P., Overduin, B., Parker, A., Pritchard, B., Prlic, A., Rice, S., Rios, D., Schuster, M., Sealy, I., Slater, G., Smedley, D., Spudich, G., Trevanion, S., Vilella, A. J., Vogel, J., White, S., Wood, M., Birney, E., Cox, T., Curwen, V., Durbin, R., Fernandez-Suarez, X. M., Herrero, J., Hubbard, T. J., Kasprzyk, A., Proctor, G., Smith, J., Ureta-Vidal, A., and Searle, S. (2008) Ensembl 2008. *Nucleic Acids Res.* **36,** D707–D714

33. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23,** 2947–2948

34. Li, L., Stoeckert, C. J., Jr., and Roos, D. S. (2003) OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* **13,** 2178–2189

35. Keino-Masu, K., Masu, M., Hinck, L., Leonardo, E. D., Chan, S. S., Culotti, J. G., and Tessier-Lavigne, M. (1996) Deleted in Colorectal Cancer (DCC) encodes a netrin receptor. *Cell* **87,** 175–185

36. Menon, S. D., Osman, Z., Chenchill, K., and Chia, W. (2005) A positive feedback loop between Dumbfounded and Rolling pebbles leads to myotube enlargement in Drosophila. *J. Cell Biol.* **169,** 909–920

37. Reed, J., McNamee, C., Rackstraw, S., Jenkins, J., and Moss, D. (2004) Diglons are heterodimeric proteins composed of IgLON subunits, and Diglon-CO inhibits neurite outgrowth from cerebellar granule cells. *J. Cell Sci.* **117,** 3961–3973

38. Sakisaka, T., Ikeda, W., Ogita, H., Fujita, N., and Takai, Y. (2007) The roles of nectins in cell adhesions: cooperation with other cell adhesion molecules and growth factor receptors. *Curr. Opin. Cell Biol.* **19,** 593–602

39. Kim, T., Kim, K., Lee, S. H., So, H. S., Lee, J., Kim, N., and Choi, Y. (2009) Identification of LRRc17 as a negative regulator of receptor activator of NF-kappaB ligand (RANKL)-induced osteoclast differentiation. *J. Biol. Chem.* **284,** 15308–15316

40. Mizuno, T., Kawasaki, M., Nakahira, M., Kagamiyama, H., Kikuchi, Y., Okamoto, H., Mori, K., and Yoshihara, Y. (2001) Molecular diversity in zebrafish NCAM family: three members with different VASE usage and distinct localization. *Mol. Cell. Neurosci.* **18,** 119–130

41. Smalla, K. H., Matthies, H., Langnäse, K., Shabir, S., Böckers, T. M., Wyneken, U., Staak, S., Krug, M., Beesley, P. W., and Gundelfinger, E. D. (2000) The synaptic glycoprotein neuroplastin is involved in long-term potentiation at hippocampal CA1 synapses. *Proc. Natl. Acad. Sci. U.S.A.* **97,** 4327–4332

42. Barabási, A. L., and Oltvai, Z. N. (2004) Network biology: understanding the cell's functional organization. *Nat. Rev. Genet.* **5,** 101–113

43. Pereira-Leal, J. B., and Teichmann, S. A. (2005) Novel specificities emerge by stepwise duplication of functional modules. *Genome Res.* **15,** 552–559

44. Sprague, J., Clements, D., Conlin, T., Edwards, P., Frazer, K., Schaper, K., Segerdell, E., Song, P., Sprunger, B., and Westerfield, M. (2003) The Zebrafish Information Network (ZFIN): the zebrafish model organism database. *Nucleic Acids Res.* **31,** 241–243

45. Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B., and Schilling, T. F. (1995) Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203,** 253–310

46. Karaulanov, E., Böttcher, R. T., Stannek, P., Wu, W., Rau, M., Ogata, S., Cho, K. W., and Niehrs, C. (2009) Unc5B interacts with FLRT3 and Rnd1 to modulate cell adhesion in Xenopus embryos. *PLoS One* **4,** e5742

47. Braun, P., Tasan, M., Dreze, M., Barrios-Rodiles, M., Lemmens, I., Yu, H., Sahalie, J. M., Murray, R. R., Roncari, L., de Smet, A. S., Venkatesan, K., Rual, J. F., Vandenhaute, J., Cusick, M. E., Pawson, T., Hill, D. E., Tavernier, J., Wrana, J. L., Roth, F. P., and Vidal, M. (2009) An experimentally derived confidence score for binary protein-protein interactions. *Nat. Methods* **6,** 91–97

48. de Lichtenberg, U., Jensen, L. J., Brunak, S., and Bork, P. (2005) Dynamic complex formation during the yeast cell cycle. *Science* **307,** 724–727

49. Komurov, K., and White, M. (2007) Revealing static and dynamic modular architecture of the eukaryotic protein interaction network. *Mol. Syst. Biol.* **3,** 110

50. Han, J. D., Bertin, N., Hao, T., Goldberg, D. S., Berriz, G. F., Zhang, L. V., Dupuy, D., Walhout, A. J., Cusick, M. E., Roth, F. P., and Vidal, M. (2004) Evidence for dynamically organized modularity in the yeast protein-protein interaction network. *Nature* **430,** 88–93

51. Ravasi, T., Suzuki, H., Cannistraci, C. V., Katayama, S., Bajic, V. B., Tan, K., Akalin, A., Schmeier, S., Kanamori-Katayama, M., Bertin, N., Carninci, P., Daub, C. O., Forrest, A. R., Gough, J., Grimmond, S., Han, J. H., Hashimoto, T., Hide, W., Hofmann, O., Kamburov, A., Kaur, M., Kawaji, H., Kubosaki, A., Lassmann, T., van Nimwegen, E., MacPherson, C. R., Ogawa, C., Radovanovic, A., Schwartz, A., Teasdale, R. D., Tegnér, J., Lenhard, B., Teichmann, S. A., Arakawa, T., Ninomiya, N., Murakami, K., Tagami, M., Fukuda, S., Imamura, K., Kai, C., Ishihara, R., Kitazume, Y., Kawai, J., Hume, D. A., Ideker, T., and Hayashizaki, Y. (2010) An atlas of combinatorial transcriptional regulation in mouse and man. *Cell* **140,** 744–752

52. Dupuy, D., Bertin, N., Hidalgo, C. A., Venkatesan, K., Tu, D., Lee, D., Rosenberg, J., Svrzikapa, N., Blanc, A., Carnec, A., Carvunis, A. R., Pulak, R., Shingles, J., Reece-Hoyes, J., Hunt-Newbury, R., Viveiros, R., Mohler, W. A., Tasan, M., Roth, F. P., Le Peuch, C., Hope, I. A., Johnsen, R., Moerman, D. G., Barabási, A. L., Baillie, D., and Vidal, M. (2007) Genome-scale analysis of in vivo spatiotemporal promoter activity in Caenorhabditis elegans. *Nat. Biotechnol.* **25,** 663–668

53. Bossi, A., and Lehner, B. (2009) Tissue specificity and the human protein interaction network. *Mol. Syst. Biol.* **5,** 260

54. Charoensawan, V., Adryan, B., Martin, S., Söllner, C., Thisse, B., Thisse, C., Wright, G. J., and Teichmann, S. A. (2010) The impact of gene expression regulation on evolution of extracellular signaling pathways. *Mol. Cell. Proteomics* **9,**

55. Sanderson, C. M. (2008) A new way to explore the world of extracellular protein interactions. *Genome Res.* **18,** 517–520