

TGF BETA PROJECT

The TGF beta Protein Interaction Map Transforming growth factor beta (TGF-beta) superfamily members, which signal through Smad proteins, are involved in autoimmune diseases, fibrotic disorders and cancer. Hybrigenics has launched a research program to better understand this important signaling pathway. A total of 44 bait proteins were screened against a highly complex placental cDNA library using Hybrigenics' high-throughput proprietary yeast two-hybrid-based technology. The protein-protein interaction map described here connects 591 human proteins potentially involved in the Smad signaling pathway. Over 750 interactions have been identified in this network. A new version of the PIMRider®, an annotation and navigation software created by Hybrigenics, made this complex network easy to explore and to analyze visually. Biological significance as well as novelty of this network are illustrated by the presence of 18 known Smad-connected proteins and by the identification of 179 proteins hitherto poorly known or unannotated. siRNA knockdown and over-expression assays performed in mammalian cells identified eight novel proteins, out of 14 tested, involved in Smad signaling (Colland et al., July 2004 in *Genome Research*).

List of baits used in the TGF-beta/BMP

project SARA Smad5 SnoN LAPTm5 PTPN12 Smad1 Smad7 FLJ20037 LMO4 RNF11 Smad2 Smad8 HIPK3 P621 ZNF8 Smad murf2 HYPA PKD2 Smad4 SNIP1 KIAA1196 PPP1CA Baits and cDNA libraries construction Baits were either directly sub-cloned in the pB27 plasmid derived from the original pBTM116 or were first PCR-amplified and then transferred (Vojtek and Hollenberg, *Methods Enzymol.* 1995, 255:331-342). Inserts were submitted to quality control by full insert sequencing and those who matched the predicted sequence were used further. In some circumstances, specific domains such as transmembrane domain, membrane-anchoring motifs, signal peptide or transactivation domain were deleted from the baits. Randomly primed cDNA libraries from human placenta poly (A+) RNA were constructed into the pP6 plasmid derived from the original pGADGH [Bartel et al., 1993, Using the two-hybrid system to detect protein-protein interactions. In *Cellular Interactions in Development: A Practical Approach*, D. A. Hartley, Ed., Oxford University Press, Oxford; pp 153-179]. The libraries were then transformed into the Y187 yeast strain and ten millions independent yeast colonies were collected, pooled and stored at -80°C as equivalent aliquot fractions of the same library. Yeast Two-Hybrid Screening Procedure and Prey Identification The screens were performed using a proprietary mating method (Rain et al. *Nature* 2001, 409:211-215). Each screen was performed to ensure a minimum of 50 millions interactions tested. Up to 384 positive clones per independent screen were picked and the corresponding prey fragments were amplified by PCR and sequenced at their 5' and 3' junctions on a PE3700 Sequencer. 5' and 3' sequences were then filtered by using PHRED (Ewing et al. *Genome Res* 1998, 8:175-185) and masking ALU repeats. Sequence contigs are built using CAP3 (Huang & Madan *Genome Res* 1999, 9:868-877) and compared to the latest release of GenBank using BLASTN (Altschul et al. *Nucleic Acids Res* 1997, 25:3389-3402). If GenBank entries corresponding to the complete mRNA are found, the best annotated entry, preferentially from the RefSeq division of GenBank, is assigned to every overlapping prey fragment family. Interactions Scoring and Results Display For each interaction, the minimal interacting domain or SID® (Selected Interacting Domain) was deduced from the individual fragments obtained in two-hybrid screens. The SID corresponds to the smallest identified protein domain necessary for a particular protein-protein interaction. A PBS® (Predicted Biological Score) was calculated for each interaction thanks to Hybrigenics proprietary algorithms allowing assessment of their reliability. These informations as well as protein annotations were integrated in a proprietary application, PIMRider®, designed to explore the Protein Interaction Map and which includes four viewers: the ProteinViewer™, which displays textual annotations for a given protein, the list of interacting partners, and the sequence; the PIMViewer™, which displays a graphical and dynamic view of protein interaction networks; the InteractionViewer™, which gives access to raw experimental data on prey, bait and SID sequences; and the DomainViewer™, which displays and makes it possible to compare, for one protein and all its partners in the map, domains and motifs extracted from both experimental (bait and SID) and calculated (transmembrane segments, signal peptides, and functional InterPro domains) analyses.