Supplementary Information

Integrative interactome analysis for mechanistic discovery and drug repurposing in rare diseases: GPX4-related rare skeletal dysplasia

Kalyani B. Karunakaran[1], N. Balakrishnan[2] and Madhavi K. Ganapathiraju[3]*

[1] Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, 560012; kalyanithepebble@gmail.com
[2] Supercomputer Education and Research Centre, Indian Institute of Science, Bangalore, 560012; balki@iisc.ac.in
[3] Department of Biomedical Informatics, School of Medicine, and Intelligent Systems Program, School of Computing and Information, University of Pittsburgh, Pittsburgh, PA, USA; madhavi@pitt.edu
* Correspondence: madhavi@pitt.edu

Pathophenotypes and functional modules enriched in the interactome of GPX4 and genes putatively associated with SMDS

This interactome was enriched in genes associated with myocarditis (P-value = 5.33E-04, odds ratio = 55.8, genes: VHL and GPX4), atrial septal defect (P-value = 0.016, odds ratio = 3.37, genes: USP9X, SMAD4, GPX4, RPS19 and EP300) and platyspondyly (P-value = 0.018, odds ratio = 5.38, genes: GPX4, SMAD4 and TONSL). Myocarditis-associated VHL was connected to GPX4 via ZNF197. Platypondyly-associated TONSL was predicted to be a novel interactor of RPS19.

A *cellular oxidant detoxification* module was identified in both global (M4: *Q*-value = 2.15E-03) and fetus-specific (M3: *Q*-value = 2.77E-03) contexts. GPX4 and 3 of its novel interactors (APBA3, GAMT and GTF2F1), RPS19 and 3 of its novel interactors (NEDD8, PSMD8 and XRCC1) and a novel interactor of AGRP called DXO were detected in the global functional module. The fetus-specific module did not contain DXO, but instead contained a novel interactor of ARTN called LDB1. The *adenylate cyclase-activating G-protein coupled receptor signaling pathway* was detected both as a global (M3: *Q*-value = 4.25E-04) and fetus-specific (M2: *Q*-value = 3.4E-04) module. 3 novel interactors of GPX4 (EGR4, MATK and ZNF197), 3 novel interactors of RPS19 (CHRNB4, PSG9 and TONSL), 2 novel interactors of ARNTL (CAMK4 and FCGBP), ARTN and its novel interactor DIO1, and AGRP and its novel interactor DHODH were detected in the global module. The fetus-specific module did not contain AGRP and DHODH, but instead contained VHL, a novel intermediate interactor between GPX4 and AGRP. *Cell growth* (M1: *Q*-value = 2.38E-04) and *response to redox state* (M2: *Q*-value = 2.38E-04) were detected only as global modules. The *cell growth module* contained 2 novel interactors of ARNTL (EIF4G2 and MSH2) and a novel interactor of ARTN (TUBGCP3). The *redox state response* module contained a novel interactor of AGRP called UPP1. *N-terminal peptidyl–lysine acetylation* (M1: *Q*-value = 1.21E-04)/*histone acetylation* (M1: *Q*-value = 2.84E-04) and *regulation of DNA metabolic process* (M4: *Q*-value = 7.84E-04) were detected only as tissue-specific modules. A novel interactor of ARNTL (TRIM22) and a novel interactor of AGRP (RNF19A) belonged to the *acetylation* module, whereas MSH2 and TUBGCP3 that were earlier detected in the *cell growth* module, seemed to also belong to the *DNA metabolic process* module.

Pathophenotypes and functional modules enriched in the interactome of GPX4 and genes related to other skeletal dysplasias

This interactome was enriched in genes associated with arrhythmia (*P*-value = 0.041, odds ratio = 1.88, genes: TAB2, COL4A1, HSPG2, FLNC, CALM1, GPX4, CBL, VHL, GSN and TGFB1), cardiorespiratory arrest (*P*-value = 9.25E-03, odds ratio = 13.38, genes: COL2A1 and GPX4), myocarditis (*P*-value = 4.64E-03, odds ratio = 19.11, genes: GPX4 and VHL), metaphyseal cupping (*P*-value = 5.9E-03, odds ratio = 3.68, genes: GPX4, PCYT1A, INPPL1, TRIP11, MMP13 and COL2A1), platyspondyly (*P*-value = 4.98E-13, odds ratio = 10.43, genes: GPX4, PCYT1A, INPPL1, TRIP11, MMP13, COL2A1, HSPG2, SPARC, BGN, COMP, COL1A1, COL9A2, COL9A3, CSF1R, TRPV4, COL1A2 and SMAD4) and abnormal ribs (*P*-value = 8.9E-03, odds ratio = 3.93, genes: GPX4, PCYT1A, HSPG2, TRPV4 and SMAD4).

Three pairs of closely related modules were detected in global and fetal contexts: *blood vessel development* (M1: *Q*-value = 9.6E-06)/*positive regulation of cell migration* (M1: *Q*-value = 2.77E-06), *extracellular organization* (M3: *Q*-value = 3.53E-04)/*cartilage development* (M3: *Q*-value = 2.72E-03) and *regulation of cysteine-type endopeptidase activity involved in apoptotic process* (M6: *Q*-value = 6.68E-03)/*execution phase of apoptosis* (M4: *Q*-value = 6.75E-04). Three unique global modules were identified: *leukocyte tethering or rolling* (M2: *Q*-value = 3.7E-05), *response to bacterium* (M4: *Q*-value = 6.4E-03) and *morphogenesis of an epithelium* (M5: *Q*-value = 6.57E-03). Four unique fetus-specific modules were identified: *cellular component maintenance* (M2: *Q*-value = 3.39E-04), *actin filament organization* (M5: *Q*-value = 1.36E-03), *negative regulation of intrinsic apoptotic signaling pathway* (M6: *Q*-value = 2.51E-03) and *regulation of protein catabolic process* (M7: *Q*-value = 1.05E-02). 3 novel interactors of GPX4 (EGR4, FUT5 and ZNF197) co-occurred with 10 novel interactors of genes associated with other SD genes (SPDYA, MST1, DRD2, KMT2D, GP5, LY6H, SERPINA4, CLSPN, IKZF2 and TM6SF2) and 2 SD genes themselves (COL2A1 and TRPV4) in the *extracellular organization/cartilage development* modules in both global and fetus-specific contexts. The global module additionally contained the novel interactors THRA and ADAM23, and the fetus-specific module contained MATK (a novel interactor of GPX4), CD3G, FICD, NPTX1 and RASGRP4. 2 novel interactors of GPX4 (APBA3 and GTF2F1) co-occurred with 3 novel interactors of SD-associated genes (RXRB, SLC48A1 and SF3A2) in the *epithelial morphogenesis* (global) module. APBA3, a novel interactor of GPX4, co-occurred with 2 novel interactors of SD-associated genes (RXRB and SF3A2) in the *protein catabolic process* (fetus-specific) module.

Pathophenotypes and functional modules enriched in the interactome of GPX4 and genes linked to disorders phenotypically similar to SMDS

This interactome was enriched in genes associated with arrhythmia (*P*-value = 0.042, odds ratio = 3.05, genes: TGFB1, IARS2, GPX4 and ELN), cardiorespiratory arrest (*P*-value = 5.9E-04, odds ratio = 54.2, genes: COL2A1 and GPX4), cerebellar hypoplasia (*P*-value = 8.63E-03, odds ratio = 4.95, genes: BMP4, EPG5, GPX4 and DAG1), agenesis of corpus callosum (*P*-value = 1.88E-03, odds ratio = 5.65, genes: BMP4, EPG5, GPX4, PTCH1 and DAG1), metaphyseal cupping (*P*-value = 3.73E-05, odds ratio = 45.18, genes: GPX4, COL2A1 and PAM16), platyspondyly (*P*-value = 5.9E-09, odds ratio = 19.9, genes: GPX4, COL2A1, PAM16, BGN, DYM, SMAD4, COL1A1 and COL1A2) and abnormal ribs (*P*-value = 3.8E-03, odds ratio = 9.57, genes: GPX4, SMAD4 and PTCH1). Arrhythmia-associated IARS2 is associated with X-linked spondyloepimetaphyseal dysplasia and separated from GPX4 only by 3 edges and 5 intermediate interactors (including 2 novel interactors of GPX4, namely, APBA3 and ZNF197). EPG5, the gene associated with cerebellar hypoplasia and corpus callosum agenesis, is a novel interactor of DYM, a Dyggve-Melchior-Clausen disease gene, and is an intermediate interactor connecting DYM to GPX4. PAM16, associated with metaphyseal cupping and platyspondyly, is a Megarbane type SMD gene, and is separated from GPX4 only by 3 edges and 2 intermediate interactors (including the novel interactor of GPX4, GTF2F1). Both BGN associated with X-linked spondyloepimetaphyseal dysplasia and DYM associated with Dyggve-Melchior-Clausen disease were separated from GPX4 by 3 edges, and appeared to be linked to platyspondyly.

Five modules were detected in both global and fetus-specific contexts from the interactome: *skeletal system development* (global M1: *Q*-value = 2.76E-06)/*embryo development/gastrulation* (fetus-specific M1: *Q*-value = 2.28E-04), *lipid storage* (global M2: *Q*-value = 1.69E-04)/*reactive oxygen species biosynthetic process* (fetus-specific M2: *Q*-value = 9.17E-04), *humoral response* (global M3: *Q*-value = 2E-03 and fetus-specific M3: *Q*-value = 1.28E-03), *cellular response to toxic substance* (global M4: *Q*-value = 2.59E-03 and fetus-specific M4: *Q*-value = 3.76E-03) and *membrane organization* (global M5: *Q*-value = 8.47E-03 and fetus-specific M6: *Q*-value = 6.29E-03). We noted another *embryo development* module that was fetus-specific and enriched for *neurogenesis* (M5: *Q*-value = 0.01). GPX4 and 2 of its novel interactors (APBA3 and GTF2F1) co-occurred with a novel interactor of a gene associated with a bone dysplasia (DNASE1L1) in the *toxic substance response* module in both global and fetus-specific contexts. 1 novel interactor (AMDHD2) and 2 novel interactors (CDIP1 and C1orf115) uniquely occurred in this module in global and fetus-specific contexts respectively. ZNF197, a novel interactor of GPX4, co-occurred with 4 novel interactors of bone dysplasia genes (TEX28, HHIPL2, SUSD4 and SLC30A10) and a bone dysplasia gene itself (IHH) in the global *humoral immune response* module. ZNF197 also co-occurred with 4 novel interactors of bone dysplasia genes (CCL20, AMDHD2, HHIPL2 and SLC30A10) in the fetus-specific *lipid storage/reactive oxygen species biosynthetic process* module.

Pathophenotypes and functional modules enriched in the interactome of GPX4 and genes that affect GPX4 expression when perturbed

This interactome was enriched in genes associated with cardiorespiratory arrest (*P*-value = 3.2E-03, odds ratio = 23.06, genes: GPX4 and KIT) and myocarditis (*P*-value = 1.52E-03, odds ratio = 32.9, genes: GPX4 and VHL).

Telomere maintenance/organization was detected both as global (M1: *Q*-value = 1.16E-04) and fetus-specific (M1: *Q*-value = 1.84E-06) modules. *Positive regulation of neurogenesis* was also detected in global (M7: *Q*-value = 5.88E-03) and fetal (M5: *Q*-value = 5.34E-03) contexts. 6 additional global modules were detected: *cellular response to retinoic acid* (M2: *Q*-value = 1.73E-03), *detoxification* (M3: *Q*-value = 1.73E-03), *peptidyl-threonine phosphorylation* (M4: *Q*-value = 1.73E-03), *regulation of cellular protein localization* (M5: *Q*-value = 3.5E-03), *DNA replication* (M6: *Q*-value = 5.35E-03) and *protein autophosphorylation* (M8: *Q*-value = 1.14E-02). 3 additional fetal modules were detected: *negative regulation of myeloid cell differentiation* (M2: *Q*-value = 1.33E-04), *transforming growth factor beta receptor signaling pathway* (M3: *Q*-value = 3.25E-04) and *positive regulation of endopeptidase activity* (M4: *Q*-value = 7.27E-04). 3 novel interactors of GPX4 (EGR4, FUT5 and ZNF197) co-occurred with 2 perturbed genes (NTRK1 and DNMT3L) in the *protein autophosphorylation* global module and in the *neurogenesis* module that was detected in both global and fetus-specific contexts.

Interconnections of GPX4 and genes associated with Joubert syndrome

From our Phenogrid analysis, we noted that Joubert syndrome with JATD showed 64% phenotypic similarity with SMDS. The developmental disorders shared pathophenotypes such as agenesis of corpus callosum, generalized hypotonia, cerebellar hypoplasia and atrial septal defects. Joubert syndrome with JATD shows an amalgamation of several key traits associated with Joubert syndrome such as ataxia-inducing brain stem malformations, hypotonia and cognitive impairment, and skeletal traits characteristic of JATD such as a narrow thorax that leads to respiratory failure, and rib, limb, and metaphyseal dysplasia (Lehman et al., 2010).

We sought to determine whether GPX4 was closely connected to the genes associated with Joubert syndrome, namely, IFT140, KIAA0586, and CSSP1, extracted from DisGeNET. IFT140, KIAA0586, and CSSP1 shared 48 intermediate interactors, including 14 novel interactors with GPX4 (**Fig. 6a**). These included four novel interactors of IFT140 (TRAP1, TELO2, IL32 and DNAJA3), one novel interactor of KIAA0586 (DACT1), five novel interactors of GPX4 (APBA3, GAMT, GTF2F1, MATK and ZNF197), and four intermediate interactors (KPNA5, SP100, SNRPC and MAPK13). Additionally, we identified three genes within this network whose perturbation led to underexpression of Joubert syndrome-associated genes (**Fig. 6a**). Interestingly, these genes (TRIM28, ABL1 and BCR1) were known interactors of the novel interactors (ZNF197, GTF2F1 and APBA3) that we had predicted for GPX4. Knockout of TRIM28 resulted in the underexpression of IFT140, whereas mutations in ABL1 and BCR1 resulted in the underexpression of KIAA0586.

Since GPX4 showed close interconnections with IFT140, a ciliary protein, we further examined whether GPX4 was connected to other ciliary proteins. For this purpose, we examined the shortest paths connecting 165 ciliary proteins (Karunakaran et al., 2020b) to GPX4. The ciliary protein that showed the closest interconnection with GPX4 was TERF1 (**Fig. 6b**). GAMT, a novel interactor of GPX4, and PRDX6, a known interactor of GPX4, served as intermediate interactors between GPX4 and TERF1. Overall, GPX4 was connected to TERF1 through 13 intermediate interactors, including four novel interactors of GPX4 (APBA3, GAMT, GTF2F1 and MATK), one novel interactor of TERF1 (LYN) and two intermediate interactors (TRIP10 and SP100).

Detoxification was detected as both a global and fetus-specific module (**Supplementary Note 5**). Other enriched global modules included positive regulation of cell development, positive regulation of actin filament bundle assembly, regulation of G1/S transition of mitotic cell cycle, regulation of JAK-STAT cascade, response to growth factor, regulation of proteasomal protein catabolic process and protein localization of organelle. Additional fetus-specific modules included regulation of binding, cellular response to organonitrogen compound and telomere capping.

Functional modules enriched in the interactome of GPX4 and ciliary genes

Detoxification was detected both as a global (M4: *Q*-value = 1.26E-03) and fetus-specific (M4: *Q*-value = 1.68E-03) module. Other enriched global modules included *positive regulation of cell development* (M1: *Q*-value = 1.18E-03),

positive regulation of actin filament bundle assembly (M2: *Q*-value = 1.18E-03), *regulation of G1/S transition of mitotic cell cycle* (M3: Q-value = 1.18E-03), *regulation of JAK-STAT cascade* (M5: *Q*-value = 1.26E-03), *response to growth factor* (M6: *Q*-value = 1.32E-03), *regulation of proteasomal protein catabolic process* (M7: *Q*-value = 1.84E-03) and *establishment of protein localization of organelle* (M8: *Q*-value = 1.94E-03). Other enriched fetus-specific modules included *regulation of binding* (M1: *Q*-value = 4.62E-05), *cellular response to organonitrogen compound* (M2: *Q*-value = 8.91E-04) and *telomere capping* (M3: *Q*-value = 1.05E-03).