

Supplementary Information:

Supplementary Note 1	Differences in topological distance between proteins with interaction-disrupting and non-disrupting dnMis mutations to proteins from known ASD-associated classes
Supplementary Note 2	Ranking computational prediction candidates for experimental validation
Supplementary Note 3	Enrichment of dnMis mutations that are identified as population variants on protein interaction interfaces
Supplementary Note 4	Characteristic network and haploinsufficiency properties of genes with interaction-disrupting dnMis mutations across a wide range of developmental disorders (DDs)
Supplementary Note 5	Curation of high-throughput (HT)-derived human interactomes
Supplementary Note 6	Expanding interaction disruption predictions for all dnMis mutations
Supplementary Figure 1	Analyses of interaction-disrupting (Dis) and non-disrupting (Non-Dis) dnMis mutations measured experimentally (blue) or computationally (purple)
Supplementary Figure 2	Analyses of probably damaging (PBD) and benign (BN) dnMis mutations predicted by Polyphen-2 (PPH2)
Supplementary Figure 3	Analyses of interface and non-interface (Non-Int) dnMis mutations predicted by Interactome INSIDER
Supplementary Figure 4	Analyses of interaction-disrupting (Dis) and non-disrupting (Non-Dis) dnMis mutations predicted by our computational approach using 0.5 and 0.7 as Polyphen-2 cutoffs
Supplementary Figure 5	Analyses of dnMis mutations in developmental disorders (DDs) with respect to ExAC
Supplementary Figure 6	Network properties, haploinsufficiency, and pLI of genes with dnMis mutations
Supplementary Figure 7	Network properties of proteins with dnMis mutations in a chronologically-ordered series of high-throughput (HT)-derived human interactomes
Supplementary Figure 8	Analyses of interaction-disrupting (Dis) and all the other non-disrupting (Non-Dis) dnMis mutations predicted by our computational approach (no experimental data)
Supplementary Figure 9	Schematic of Clone-seq pipeline
Supplementary Figure 10	Uncropped Western blots for co-IP of RXRB with wild-type and mutant RARA in Fig. 4c
Supplementary Table 1	All dnMis mutations in SSC database (.xlsx)
Supplementary Table 2	Interaction disruption results tested in Y2H experiments (.xlsx)
Supplementary Table 3	Genes with interaction-disrupting (Dis) or non-disrupting (Non-Dis) dnMis mutations (.xlsx)
Supplementary Table 4	Lists of genes in seven ASD-associated functional classes (.xlsx)

Supplementary Note 1: Differences in topological distance between proteins with interaction-disrupting and non-disrupting dnMis mutations to proteins from known ASD-associated classes

Topological distance analyses revealed that in ASD probands proteins harboring interaction-disrupting dnMis mutations are significantly closer to proteins from seven ASD-associated classes in comparison to proteins with non-disrupting dnMis mutations (Table 1). We consider these observed differences in topological distance, though small in magnitude, as meaningful indicators of their functions in relation to known ASD genes and pathways. The human protein interactome overall is densely connected, where over 90% of the proteins are within the largest connected component of the network. As such, the distance between two protein sets averaged across all possible protein pairs between these two sets could appear small.

In our analyses, we note crucially that unaffected siblings serve as a strong negative control. In ASD probands, we observed ~0.25 average difference in topological distance between the “Dis” category and “Non-Dis” category across all seven gene classes, with *P*-values at the level of 10^{-6} (Table 1). In contrast, the average difference observed in unaffected siblings is only ~0.01, over an order of magnitude smaller than the difference observed in ASD probands and does not approach statistical significance. Thus the significant differences in topological distance between “Dis” and “Non-Dis” in ASD probands strongly imply that proteins with interaction-disrupting dnMis mutations are more likely to share common functions with known ASD-associated proteins.

Supplementary Note 2: Ranking computational prediction candidates for experimental validation

Our computational prediction approach identified a total of 81 interaction-disrupting dnMis mutations in ASD probands. To prioritize candidates for experimental validation, we ranked candidate mutations by their Interactome INSIDER prediction scores and their PPH2 scores. 17 mutations with Interactome INSIDER prediction score = 1.0 and PPH2 score = 1.0 were identified. We excluded genes that are loss-of-function tolerant ($pLI^1 \leq 0.10$) and further partitioned the remaining 11 genes by their probabilities of being haploinsufficient (pHI^2). For experimental validation, we need to have the clone for the gene available in our hORFeome v8.1 clone library. Upon applying these rankings, proband RARA p.Pro375Leu mutation emerged as the top experimentally testable candidate.

Gene	Uniprot	Mutation	Interactome INSIDER	PPH2	Proband	pHI	hORFeome v8.1
<i>RARA</i>	P10276	p.Pro375Leu	1.0	1.0	14108.p1	0.973	yes
<i>SFPQ</i>	P23246	p.Tyr470Cys	1.0	1.0	11193.p1	0.901	yes
<i>PRKCA</i>	P17252	p.Pro514Leu	1.0	1.0	12465.p1	0.827	yes
<i>PSMC5</i>	P62195	p.Arg258Trp	1.0	1.0	11768.p1	0.597	yes
<i>DPYSL2</i>	Q16555	p.Arg496Cys	1.0	1.0	11566.p1	0.456	yes

<i>KCND3</i>	Q9UK17	p.Arg86Pro	1.0	1.0	14020.p1	0.303	yes
<i>KCND3</i>	Q9UK17	p.Asp85Val	1.0	1.0	14020.p1	0.303	yes
<i>GPS1</i>	Q13098	p.Arg456Gln	1.0	1.0	13629.p1	0.226	yes
<i>RAN</i>	P62826	p.Thr93Ile	1.0	1.0	12437.p1	na	yes
<i>ABL1</i>	P00519	p.Thr117Met	1.0	1.0	14373.p1	0.995	no
<i>TRIO</i>	O75962	p.Lys1431Met	1.0	1.0	13621.p1	0.712	no
<i>MBD2</i>	Q9UBB5	p.Arg380Cys	1.0	1.0	12910.p1	0.474	no

We note that no literature evidence was used in our rankings. In fact, *RARA* has never been directly implicated in ASD in the literature. Moreover, an unaffected sibling was found to have a dnMis mutation, p.Arg83His, on *RARA*. *RARA* p.Arg83His was computationally predicted to occur away from the interaction interface. Since this mutation is found in a healthy sibling unaffected with ASD on the same gene as the proband mutation, *RARA* p.Arg83His served as a convincing negative control in our validation experiments for our computational predictions.

Supplementary Note 3: Enrichment of dnMis mutations that are identified as population variants on protein interaction interfaces

Although *de novo* mutations were specifically identified in parent-offspring trios, we recognize that a fraction of such spontaneously arising mutations coincide with variants segregating within human populations. As such, we examined whether dnMis mutations in developmental disorders (DDs) that were found independently in the Exome Aggregation Consortium (ExAC)¹ are also enriched on protein interaction interfaces in the human interactome. We found that while the enrichment pattern persisted for this subset of DD dnMis mutations as we previously observed in Figure 5, the extent of this enrichment weakened (Supplementary Fig. 5a). In fact, when we directly compared the fraction of interface residues between ExAC-present and ExAC-absent DD dnMis mutations, we found that mutations absent from ExAC occur more frequently on interaction interfaces than ExAC-present ones ($P = 5.9 \times 10^{-3}$ by one-tail Z-test, Supplementary Fig. 5b). This result is expected since ExAC serves as a proxy for standing variation in the human population and thus mutations absent from this reference panel are more likely to be deleterious³. We further note that the ExAC-present dnMis mutations distribute principally across very low allele frequencies (Supplementary Fig. 5c), which likely restricts any potential relationship between topological properties and population allele frequency for dnMis mutations (Supplementary Fig. 5d-5e).

Supplementary Note 4: Characteristic network and haploinsufficiency properties of genes with interaction-disrupting dnMis mutations across a wide range of developmental disorders (DDs)

To reinforce that our interactome perturbation framework is scalable and generalizable to prioritize missense mutations in human diseases, we applied our interaction-disruption predictor on ~7,500 dnMis

mutations previously identified across a wide range of DDs⁴ within the human interactome. We repeated our topology and haploinsufficiency analyses, and we found that the newly predicted interaction-disrupting dnMis mutations in DDs exhibit similar network and haploinsufficiency properties to those in ASD probands; interaction-disrupting dnMis mutations tend to impact hub proteins in the interactome network, and they occur more frequently on haploinsufficient genes in comparison to non-disrupting dnMis mutations (Supplementary Fig. 6). These results show that these characteristic network and haploinsufficiency properties are not unique to ASD but are shared features across different DDs. To further confirm that such features are truly disease-associated, we examined these same network and haploinsufficiency properties for dnMis mutations that are reported as pathogenic in ClinVar⁵ (downloaded on November 27, 2017). Indeed, we found that genes with potentially disease-contributing dnMis mutations tend to encode hub proteins in the human interactome and are more likely to be haploinsufficient in comparison to genes without known pathogenic dnMis mutations (Supplementary Fig. 6). In contrast, no differential patterns were observed when we compared genes corresponding to dnMis mutations in DDs to genes corresponding to dnMis mutations in unaffected siblings (Supplementary Fig. 6). These results underscore that the characteristic network and haploinsufficiency properties reported for ASD and DDs are specific to disease-associated dnMis mutations. Thus, our findings that interaction-disrupting dnMis mutations exhibit similar features to disease-associated dnMis mutations indicate their functional significance in disease etiology. Therefore, our interactome perturbation framework offers an effective and generalizable way to identify potential disease-associated dnMis mutations.

Supplementary Note 5: Curation of high-throughput (HT)-derived human interactomes

The current HT-derived human interactome consisting of 47,044 interactions was constructed through manual curation of 48 publications as listed below.

PubMed	Year	Number of interactions	PubMed	Year	Number of interactions
12421765	2002	144	20936779	2010	807
12614612	2003	116	21078624	2011	151
12805554	2003	222	21163940	2011	334
15231747	2004	498	21182203	2011	141
15232106	2004	135	21900206	2011	3,213
15231748	2004	1,030	21516116	2011	1,255
14667819	2004	126	21044950	2011	710
15383276	2004	273	21988832	2011	4,084
15761153	2005	580	22493164	2012	239
16169070	2005	4,241	22558309	2012	122
15604093	2005	462	22365833	2012	737

16189514	2005	3,393	22626734	2012	147
16273093	2006	160	22939624	2012	400
16713569	2006	904	23455924	2013	547
17043677	2007	294	23275563	2013	152
17474147	2007	962	23414517	2013	701
18624398	2008	268	25036637	2014	1,579
18654987	2008	122	24722188	2014	505
19060904	2009	264	25416956	2014	15,181
19953087	2009	155	24728074	2014	201
19549727	2009	815	24705354	2014	233
19690564	2009	365	25640309	2015	673
19167335	2009	185	25814554	2015	694
20211142	2010	737	27107012	2016	716

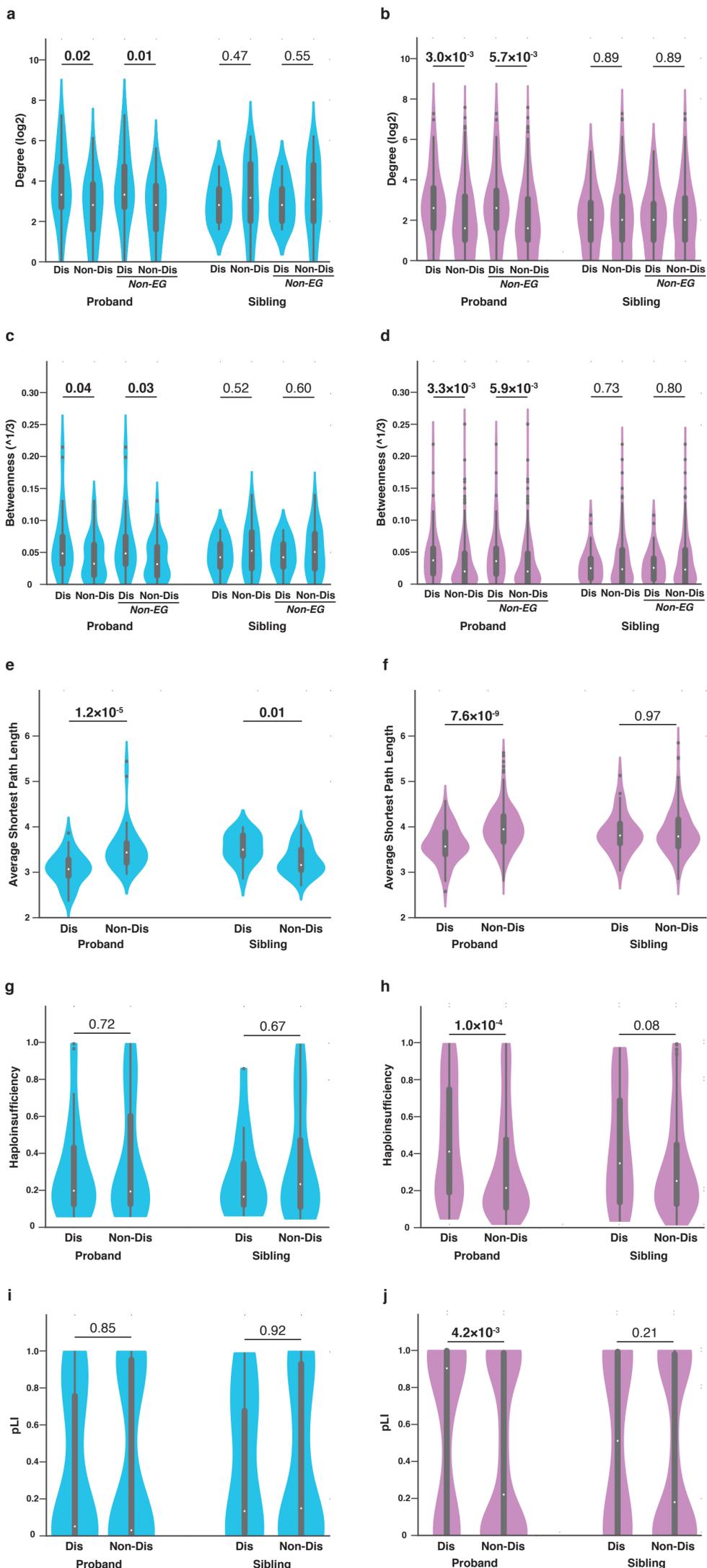
Supplementary Note 6: Expanding interaction disruption predictions for all dnMis mutations

To show that our computational approach can be readily applied to all dnMis mutations in the human interactome, we repeated all our analyses to include dnMis mutations that meet only one of the Interactome INSIDER and PPH2 criteria as non-disrupting. In other words, we consider mutations that are on the interface with a “probably damaging” PPH2 score (Interface+PPH2+) as interaction-disrupting, and all other mutations (Interface+PPH2-, Interface-PPH2+, Interface-PPH2-) as non-disrupting. We found that all our results remained the same (Supplementary Fig. 8). These results indicate that the predicted interaction-disrupting dnMis mutations exhibit highly distinguishable properties from all other dnMis mutations that are uniquely associated with ASD probands. Therefore, our computational approach can serve as an effective way towards identifying potential disease-contributing missense mutations.

References

1. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285-91 (2016).
2. Huang, N., Lee, I., Marcotte, E.M. & Hurles, M.E. Characterising and predicting haploinsufficiency in the human genome. *PLoS Genet* **6**, e1001154 (2010).
3. Robinson, E.B. *et al.* Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. *Nature Genetics* **48**, 552 (2016).
4. Turner, T.N. *et al.* denovo-db: a compendium of human de novo variants. *Nucleic Acids Res* **45**, D804-D811 (2017).
5. Landrum, M.J. *et al.* ClinVar: public archive of interpretations of clinically relevant variants. *Nucleic Acids Res* **44**, D862-8 (2016).

Experiment
Prediction



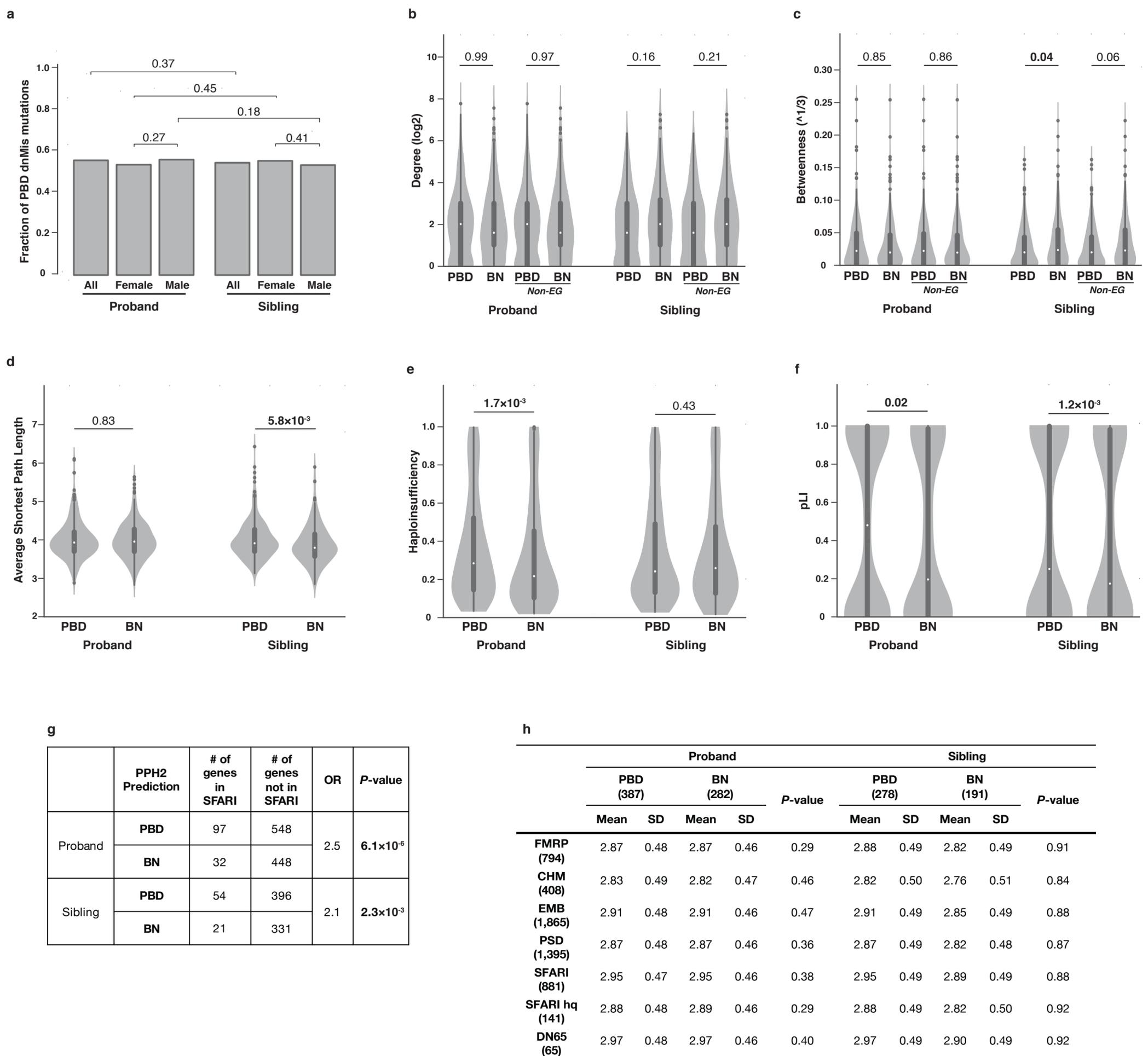
	Experiment	# of genes in SFARI	# of genes not in SFARI	OR	P-value
Proband	Dis	4	27	1.1	0.57
	Non-Dis	6	45		
Sibling	Dis	0	15	0.0	1.0
	Non-Dis	1	44		

	Prediction	# of genes in SFARI	# of genes not in SFARI	OR	P-value
Proband	Dis	17	56	3.2	1.1×10⁻³
	Non-Dis	26	270		
Sibling	Dis	6	42	1.7	0.23
	Non-Dis	16	185		

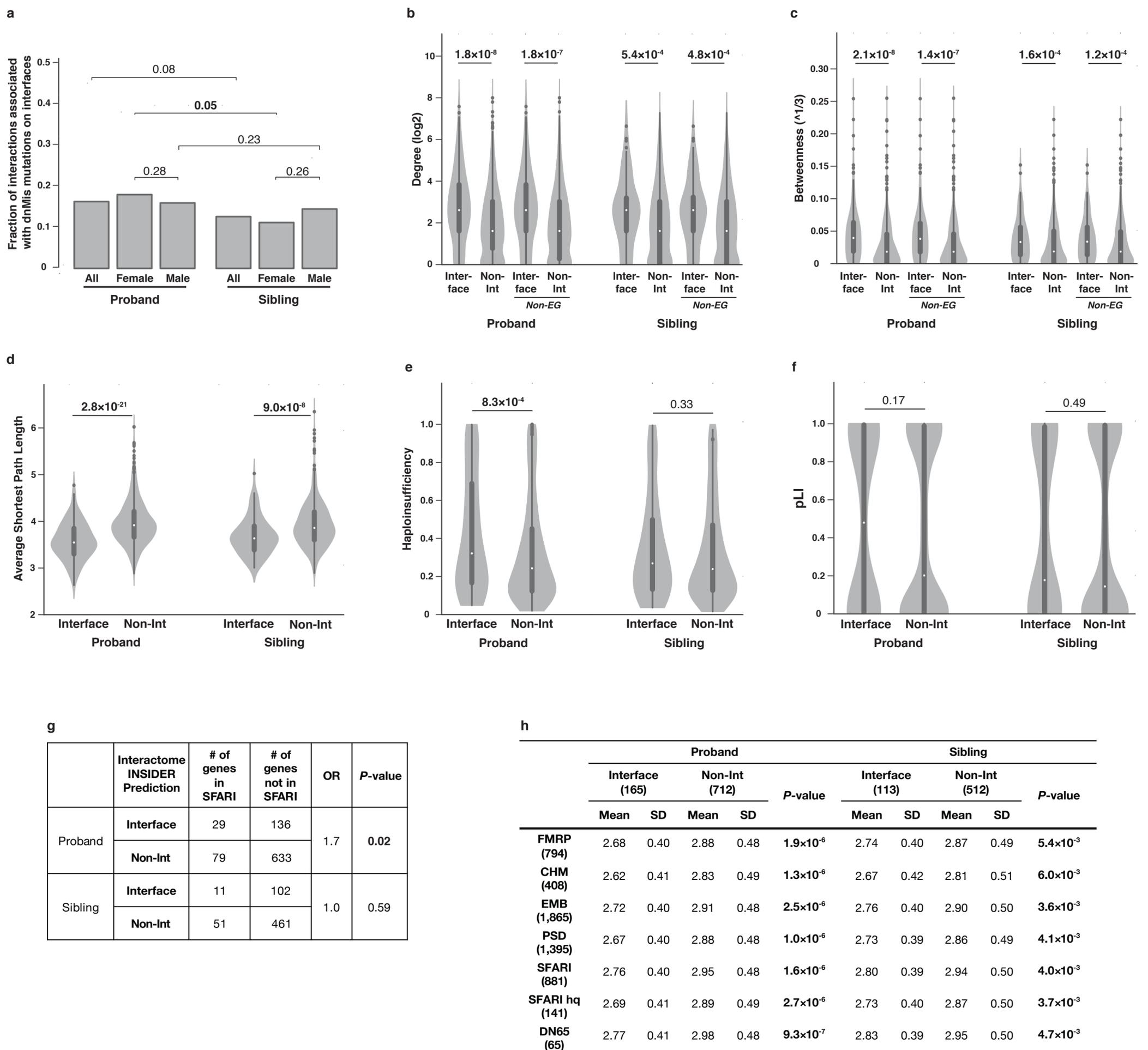
	Experiment									
	Proband					Sibling				
	Dis (31)		Non-Dis (51)		P-value	Dis (15)		Non-Dis (45)		P-value
	Mean	SD	Mean	SD		Mean	SD	Mean	SD	
FMRP (794)	2.47	0.32	2.69	0.47	0.025	2.71	0.29	2.57	0.39	0.91
CHM (408)	2.40	0.31	2.62	0.49	0.034	2.65	0.30	2.48	0.38	0.94
EMB (1,865)	2.47	0.31	2.69	0.47	0.025	2.70	0.30	2.57	0.37	0.90
PSD (1,395)	2.47	0.31	2.68	0.47	0.041	2.68	0.29	2.56	0.38	0.86
SFARI (881)	2.53	0.31	2.75	0.47	0.024	2.78	0.29	2.62	0.38	0.93
SFARI hq (141)	2.46	0.33	2.67	0.49	0.032	2.72	0.28	2.54	0.39	0.95
DN65 (65)	2.53	0.32	2.76	0.47	0.023	2.79	0.29	2.62	0.38	0.95

	Prediction									
	Proband					Sibling				
	Dis (73)		Non-Dis (296)		P-value	Dis (48)		Non-Dis (201)		P-value
	Mean	SD	Mean	SD		Mean	SD	Mean	SD	
FMRP (794)	2.69	0.38	2.87	0.45	1.4×10⁻³	2.80	0.40	2.82	0.48	0.49
CHM (408)	2.63	0.38	2.82	0.46	1.6×10⁻³	2.73	0.43	2.77	0.50	0.32
EMB (1,865)	2.74	0.37	2.90	0.45	4.0×10⁻³	2.83	0.41	2.86	0.49	0.40
PSD (1,395)	2.69	0.38	2.86	0.46	2.2×10⁻³	2.80	0.39	2.82	0.48	0.45
SFARI (881)	2.77	0.38	2.94	0.45	1.9×10⁻³	2.87	0.40	2.90	0.48	0.42
SFARI hq (141)	2.70	0.39	2.88	0.45	1.2×10⁻³	2.80	0.40	2.82	0.49	0.43
DN65 (65)	2.79	0.39	2.96	0.45	1.9×10⁻³	2.89	0.40	2.91	0.48	0.39

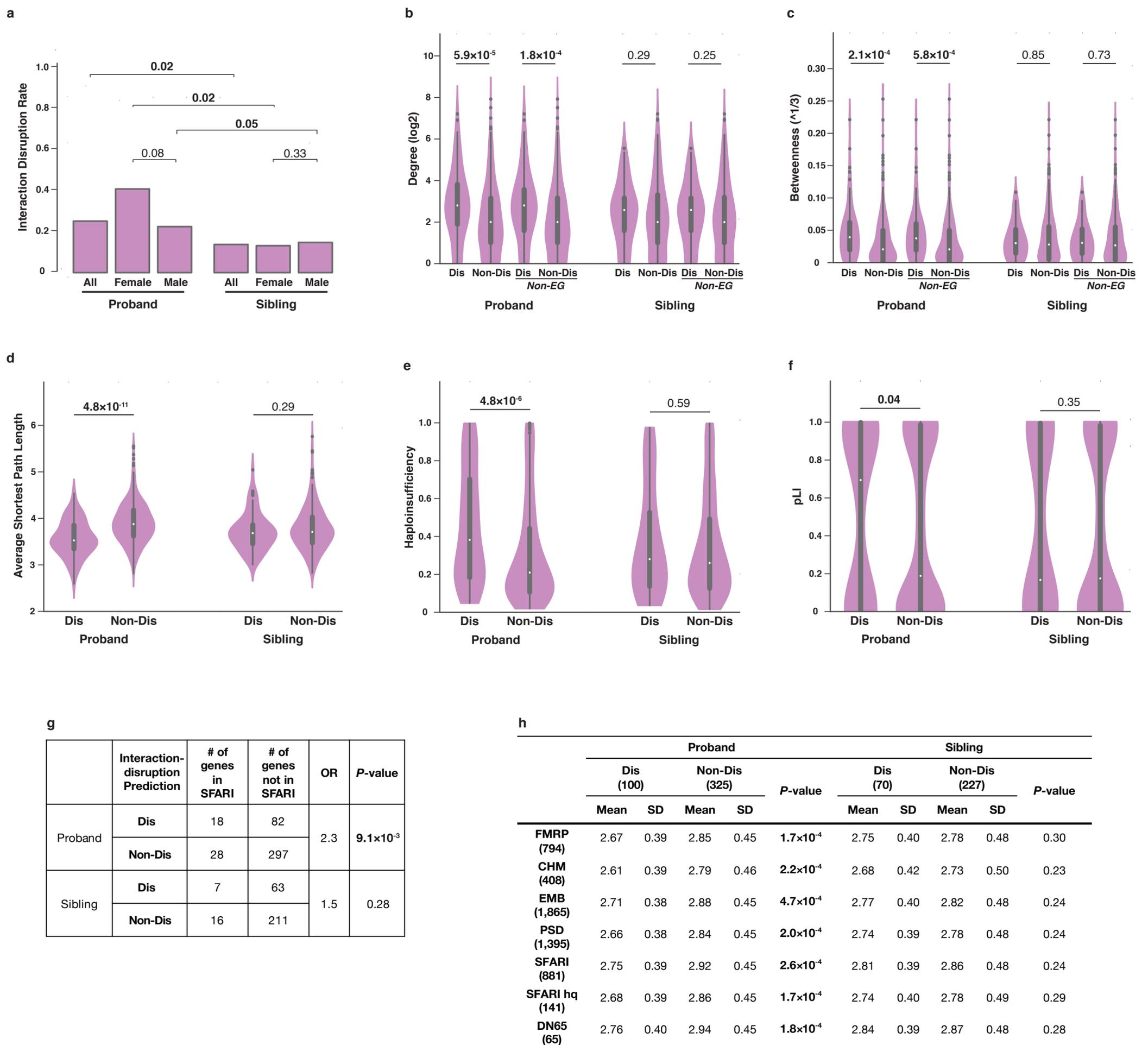
Supplementary Figure 1. Analyses of interaction-disrupting (Dis) and non-disrupting (Non-Dis) dnMis mutations measured experimentally (blue) or computationally (purple). (a-b) Degree and (c-d) betweenness distributions of proteins with Dis or Non-Dis dnMis mutations in ASD probands and unaffected siblings. Non-EG: non-essential genes encoded proteins. P-values were calculated using two-tail U-test ($P < 0.05$ in **bold**). (e-f) Average shortest path length distributions of proteins with Dis or Non-Dis dnMis mutations in ASD probands and unaffected siblings. P-values were calculated using two-tail U-test ($P < 0.05$ in **bold**). (g-h) Haploinsufficiency and (i-j) pLI of genes that harbor Dis or Non-Dis dnMis mutations in ASD probands and unaffected siblings. P-values were calculated using two-tail U-test ($P < 0.05$ in **bold**). (k-l) Contingency tables for the counts of genes harboring Dis or Non-Dis dnMis mutations in SFARI database. P-values were calculated using one-tail Fisher's exact test ($P < 0.05$ in **bold**, OR: Odds Ratio). (m-n) Distance of genes with Dis and Non-Dis dnMis mutations to seven classes of known ASD-associated genes in a protein interactome network background. Number of genes in each class is indicated in parentheses. P-values were calculated using one-tail U-test ($P < 0.05$ in **bold**).



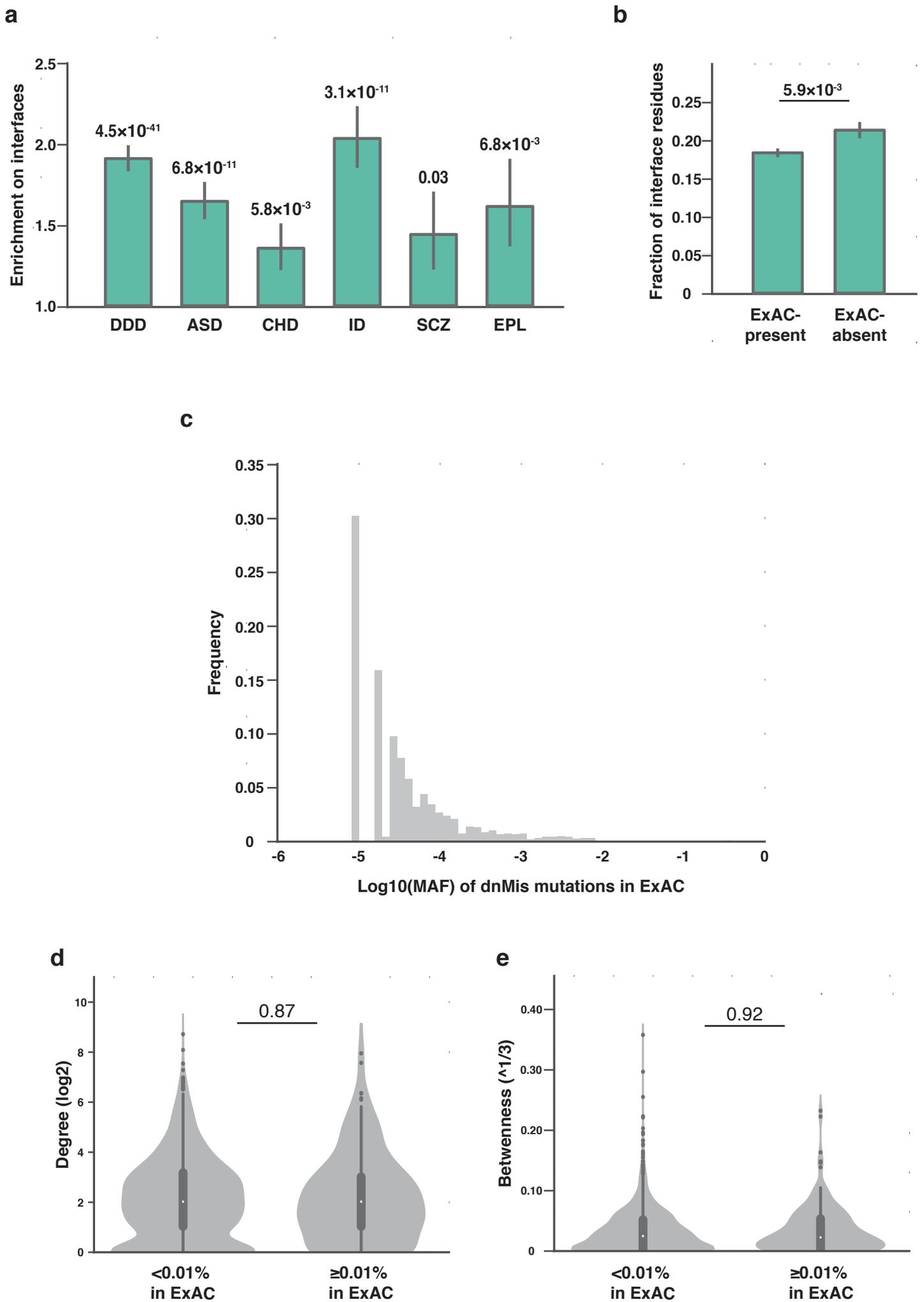
Supplementary Figure 2. Analyses of probably damaging (PBD) and benign (BN) dnMis mutations predicted by Polyphen-2 (PPH2). (a) Fraction of predicted PBD dnMis mutations in ASD probands and unaffected siblings. Probands and siblings are divided by sex. The count of PBD dnMis mutations per subject was modeled with a negative binomial model; two-tail *P*-values were calculated for comparing all probands versus all siblings, and one-tail *P*-values were calculated for comparing between and within sex (*P* < 0.05 in **bold**). (b) Degree and (c) betweenness distributions of proteins with PBD or BN dnMis mutations in ASD probands and unaffected siblings. Non-EG: non-essential gene-encoded proteins. *P*-values were calculated using two-tail *U*-test (*P* < 0.05 in **bold**). (d) Average shortest path length distributions of proteins with PBD or BN dnMis mutations in ASD probands and unaffected siblings. *P*-values were calculated using two-tail *U*-test (*P* < 0.05 in **bold**). (e) Haploinsufficiency and (f) pLI of genes that harbor PBD or BN dnMis mutations in ASD probands and unaffected siblings. *P*-values were calculated using two-tail *U*-test (*P* < 0.05 in **bold**). (g) Contingency table for the counts of genes harboring PBD or BN dnMis mutations in SFARI database. *P*-values were calculated using one-tail Fisher's exact test (*P* < 0.05 in **bold**, OR: Odds Ratio). (h) Distance of genes with PBD or BN dnMis mutations to seven classes of known ASD-associated genes in a protein interactome network background. Number of genes in each class is indicated in parentheses. *P*-values were calculated using one-tail *U*-test (*P* < 0.05 in **bold**).



Supplementary Figure 3. Analyses of interface and non-interface (Non-Int) dnMis mutations predicted by Interactome INSIDER. (a) Fraction of interactions associated with dnMis mutations on interfaces in ASD probands and unaffected siblings. Proband and siblings are divided by sex. The count of interactions associated with interface dnMis mutations per subject was modeled with a negative binomial model; two-tail *P*-values were calculated for comparing all probands versus all siblings, and one-tail *P*-values were calculated for comparing between and within sex (*P* < 0.05 in **bold**). (b) Degree and (c) betweenness distributions of proteins with interface or non-interface dnMis mutations in ASD probands and unaffected siblings. Non-EG: non-essential gene-encoded proteins. *P*-values were calculated using two-tail *U*-test (*P* < 0.05 in **bold**). (d) Average shortest path length distributions of proteins with interface or non-interface dnMis mutations in ASD probands and unaffected siblings. *P*-values were calculated using one-tail *U*-test (*P* < 0.05 in **bold**). (e) Haploinsufficiency and (f) pLI of genes that harbor interface or non-interface dnMis mutations in ASD probands and unaffected siblings. *P*-values were calculated using two-tail *U*-test (*P* < 0.05 in **bold**). (g) Contingency table for the counts of genes harboring interface or non-interface dnMis mutations in SFARI database. *P*-values were calculated using one-tail Fisher's exact test (*P* < 0.05 in **bold**, OR: Odds Ratio). (h) Distance of genes with interface or non-interface dnMis mutations to seven classes of known ASD-associated genes in a protein interactome network background. Number of genes in each class is indicated in parentheses. *P*-values were calculated using one-tail *U*-test (*P* < 0.05 in **bold**).

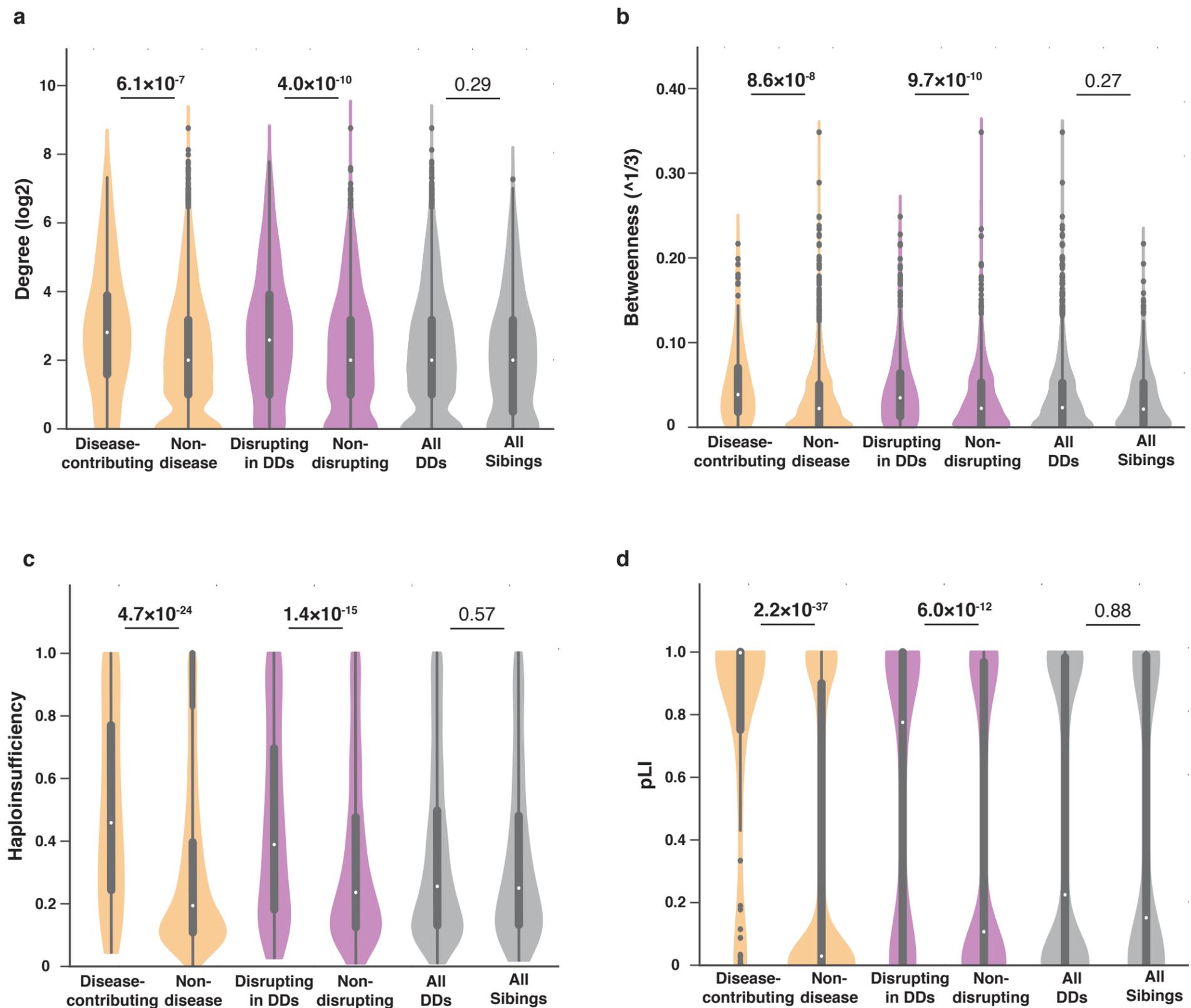


Supplementary Figure 4. Analyses of interaction-disrupting (Dis) and non-disrupting (Non-Dis) dnMis mutations predicted by our computational approach using 0.5 and 0.7 as Polyphen-2 cutoffs. (a) Interaction disruption rate of dnMis mutations in ASD probands and unaffected siblings. Proband and siblings are divided by sex. The count of disruptions per subject was modeled with a negative binomial model; two-tail *P*-values were calculated for comparing all probands versus all siblings, and one-tail *P*-values were calculated for comparing between and within sex (*P* < 0.05 in **bold**). (b) Degree and (c) betweenness distributions of proteins with Dis or Non-Dis dnMis mutations in ASD probands and unaffected siblings. Non-EG: non-essential gene-encoded proteins. *P*-values were calculated using two-tail *U*-test (*P* < 0.05 in **bold**). (d) Average shortest path length distributions of proteins with Dis or Non-Dis dnMis mutations in ASD probands and unaffected siblings. *P*-values were calculated using two-tail *U*-test (*P* < 0.05 in **bold**). (e) Haploinsufficiency and (f) pLI of genes that harbor Dis or Non-Dis dnMis mutations in ASD probands and unaffected siblings. *P*-values were calculated using two-tail *U*-test (*P* < 0.05 in **bold**). (g) Contingency table for the counts of genes harboring Dis or Non-Dis dnMis mutations in SFARI database. *P*-values were calculated using one-tail Fisher's exact test (*P* < 0.05 in **bold**, OR: Odds Ratio). (h) Distance of genes with Dis or Non-Dis dnMis mutations to seven classes of known ASD-associated genes in a protein interactome network background. Number of genes in each class is indicated in parentheses. *P*-values were calculated using one-tail *U*-test (*P* < 0.05 in **bold**).

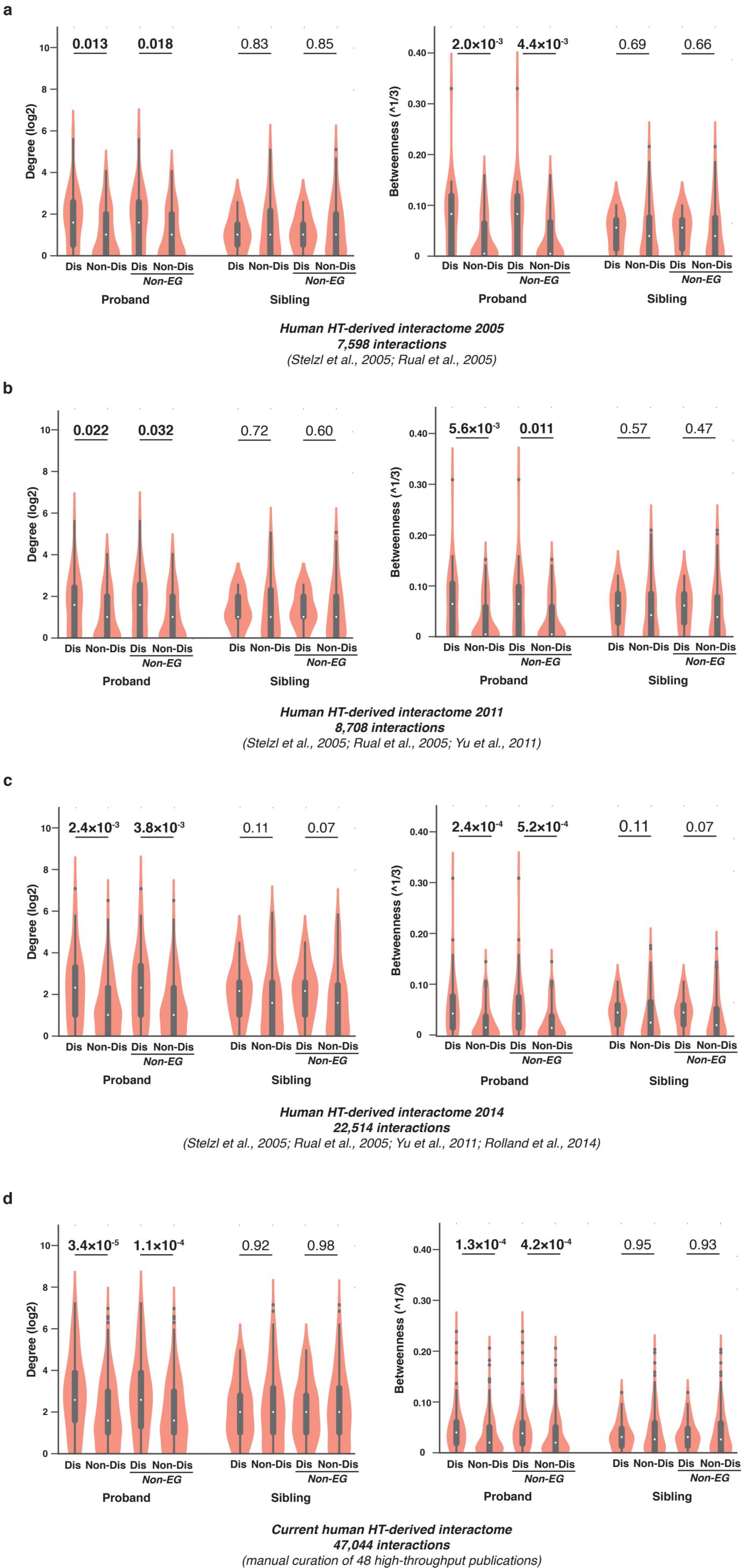


Supplementary Figure 5. Analyses of dnMis mutations in developmental disorders (DDs) with respect to ExAC. (a) Occurrence of dnMis mutations that are present in ExAC on protein interaction interfaces across different DDs. Enrichment was calculated by the ratio of the observed fraction of dnMis mutations that occur on interaction interfaces over the fraction of interface residues on corresponding proteins (expected fraction). Error bars indicate \pm standard error. *P*-values were calculated using two-tail exact binomial test. DDD: Deciphering Developmental Disorders project, ASD: autism spectrum disorder, CHD: congenital heart disease; ID: intellectual disability; SCZ: schizophrenia, EPL: epilepsy. (b) Comparison of the fraction of interface residues between dnMis mutations that are present in ExAC (ExAC-present) and those that are not present in ExAC (ExAC-absent). Error bars indicate \pm standard error of the fractions. *P*-value was calculated by one-tail Z-test. (c) Minor allele frequency (MAF) distribution of dnMis mutations that are present in ExAC and comparisons of (d) degree and (e) betweenness distributions between proteins with ultra-rare (MAF<0.01%) and other rare (MAF \geq 0.01%) dnMis mutations. *P*-values were calculated using two-tail *U*-test (*P* < 0.05 in **bold**).

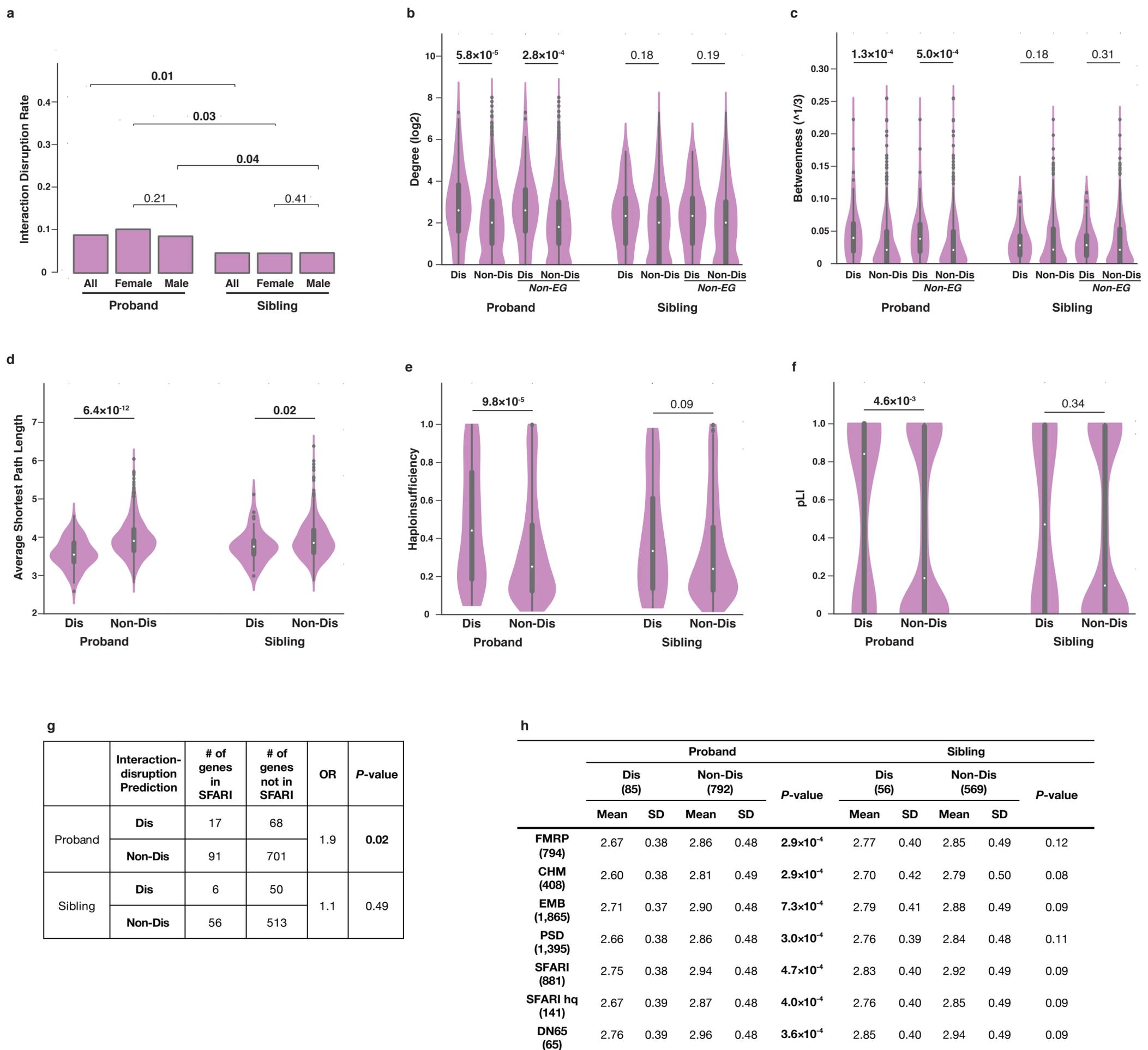
■ Disease-contributing
 ■ Prediction
 ■ Overall



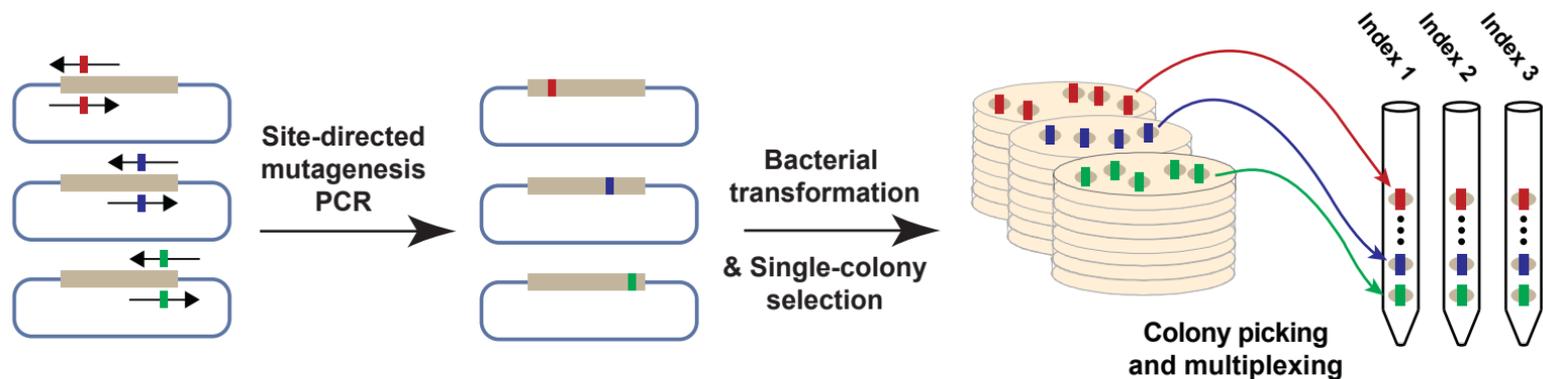
Supplementary Figure 6. Network properties, haploinsufficiency, and pLI of genes with dnMis mutations. Disease-contributing and Non-disease (yellow): genes with dnMis mutations that are reported in ClinVar as pathogenic and genes with dnMis mutations that are not ClinVar pathogenic, respectively. Disrupting in DDs and Non-disrupting (purple): genes with predicted interaction-disrupting and non-disrupting dnMis mutations in DDs, respectively. All DDs and All Sibings (grey): genes with any dnMis mutations in DDs and in unaffected siblings, respectively. (a) Degree and (b) betweenness distributions. (c) Haploinsufficiency and (d) pLI distributions. *P*-values were calculated using two-tail *U*-test (*P* < 0.05 in **bold**).



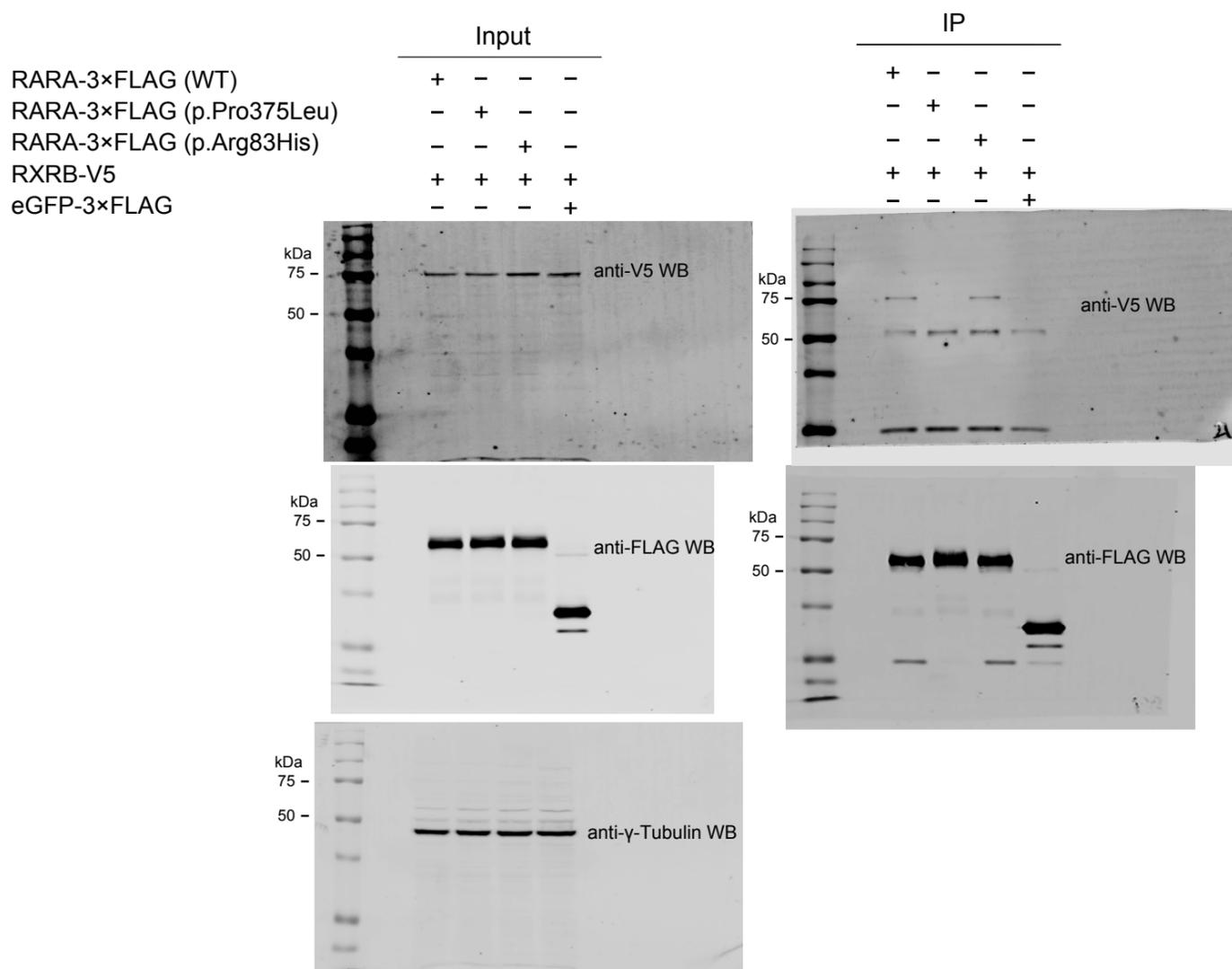
Supplementary Figure 7. Network properties of proteins with dnMis mutations in a chronologically-ordered series of high-throughput (HT)-derived human interactomes. (a-d) Degree and betweenness distributions of proteins with interaction-disrupting (Dis) or non-disrupting (Non-Dis) dnMis mutations in a series of HT-derived interactomes benchmarked at 2005, 2011, 2014, and the current HT-derived interactome. Non-EG: non-essential gene-encoded proteins. *P*-values were calculated using two-tail *U*-test (*P* < 0.05 in **bold**).



Supplementary Figure 8. Analyses of interaction-disrupting (Dis) and all the other non-disrupting (Non-Dis) dnMis mutations predicted by our computational approach (no experimental data). (a) Interaction disruption rate of dnMis mutations in ASD probands and unaffected siblings. Proband and siblings are divided by sex. The count of disruptions per subject was modeled with a negative binomial model; two-tail *P*-values were calculated for comparing all probands versus all siblings, and one-tail *P*-values were calculated for comparing between and within sex (*P* < 0.05 in **bold**). (b) Degree and (c) betweenness distributions of proteins with Dis or Non-Dis dnMis mutations in ASD probands and unaffected siblings. Non-EG: non-essential gene-encoded proteins. *P*-values were calculated using two-tail *U*-test (*P* < 0.05 in **bold**). (d) Average shortest path length distributions of proteins with Dis or Non-Dis dnMis mutations in ASD probands and unaffected siblings. *P*-values were calculated using two-tail *U*-test (*P* < 0.05 in **bold**). (e) Haploinsufficiency and (f) pLI of genes that harbor Dis or Non-Dis dnMis mutations in ASD probands and unaffected siblings. *P*-values were calculated using two-tail *U*-test (*P* < 0.05 in **bold**). (g) Contingency table for the counts of genes harboring Dis or Non-Dis dnMis mutations in SFARI database. *P*-values were calculated using one-tail Fisher's exact test (*P* < 0.05 in **bold**, OR: Odds Ratio). (h) Distance of genes with Dis and Non-Dis dnMis mutations to seven classes of known ASD-associated genes in a protein interactome network background. Number of genes in each class is indicated in parentheses. *P*-values were calculated using one-tail *U*-test (*P* < 0.05 in **bold**).



Supplementary Figure 9. Schematic of Clone-seq pipeline. In Clone-seq, mutagenesis PCR is performed in 96-well format using site-specific mutagenesis primers and full-length human ORF templates. DpnI-digested mutagenesis PCR product is then transformed into competent cells and single colonies specific to each mutagenesis PCR reaction are recovered on antibiotic-selective agar plates. After recovery, a unique single colony per mutagenesis attempt is hand-picked into a single bacterial pool and repeated for a total of three bacterial pools. DNA for each pool is then prepared by maxiprep, next-generation sequencing libraries are prepared, and then library DNA is multiplexed using unique indexing primers per pooled sample.



Supplementary Figure 10. Uncropped Western blots for co-IP of RXRB with wild-type and mutant RARA in Fig. 4c. Membranes on left represent lysate input. Membranes on right represent co-IP results. Proteins were detected using the antibodies labeled on each membrane. eGFP-3×FLAG is used as a negative control for co-IP. For γ -Tubulin loading control, input membrane on upper left was stained using anti- γ -Tubulin, as shown on the bottom left. 50 and 75 kDa markers are labeled.

Supplementary Table 1. All dnMis mutations in SSC database (.xlsx). All 2,821 dnMis mutations identified in 1,180 ASD probands and 841 unaffected siblings (from ref. 12) are listed. For each mutation, we show: the SFARI database ID of the child (for example, 12115.p1 and 12115.s1 are the ASD proband and the sibling of family 12115); the affection status and the sex of the child; the gene affected by the mutation; and the amino acid change caused by the mutation. Mutations that occur in more than one child are listed in separate lines.

Supplementary Table 2. Interaction disruption results tested in Y2H experiments (.xlsx). All interactions tested in Y2H experiments are listed. For each protein-protein interaction, we show: the protein that harbors the dnMis mutation and the amino acid change caused by the mutation; the SFARI database ID, affection status, and sex of the child; the interaction partner of the protein with dnMis mutation; and whether the interaction was disrupted (“1”) or not (“0”) by the dnMis mutation.

Supplementary Table 3. Genes with interaction-disrupting (Dis) or non-disrupting (Non-Dis) dnMis mutations (.xlsx). 109 genes with Dis, 342 genes with Non-Dis in ASD probands and 68 genes with Dis, 241 genes with Non-Dis in unaffected siblings are listed; “1” indicates the gene is found in SFARI database.

Supplementary Table 4. Lists of genes in seven ASD-associated functional classes (.xlsx). Genes used in the calculations of network distance are listed: (1) FMRP target genes (FMRP); (2) genes encoding chromatin modifiers (CHM); (3) genes expressed preferentially in embryos (EMB); (4) genes encoding postsynaptic density proteins (PSD); (5) 881 genes in the SFARI database (downloaded on June 6, 2017); (6) a high-quality SFARI subset (SFARI-hq, 141 genes scored as syndromic, high confidence, or strong candidate); and (7) the latest set of 65 ASD genes discovered *by de novo* mutations (DN65).