

**Table S1: Priorities of compound names during compound unification.** During chemical compound unification based on names the following priorities are applied.

1. InChI	9. Generic Name
2. InChI Key	10. Systematic Name
3. SMILES	11. Traditional Name
4. Formula	12. ChEBI Name
5. CAS Registry Number	13. Open Eye Name
6. IUPAC Name	14. INN
7. CAS Name	15. Mesh
8. Brand Name	16. Synonym

**Table S2: Number of gene and compound entries extracted from all source databases regardless their participation in relations.** After unification these entries form the namespace used for the creation of new relations (upon proprietary data uploads) or for the update process of existing relations.

<b>Database</b>	<b>Genes/Proteins</b>	<b>Drugs/Compounds</b>
HGNC	781,818	
NCBI Gene	11,232,904	
CGDCP	6,071	2,095
ChEMBL	5115	1,216,322
CTD	27,314	6,690
DrugBank	3,726	6,683
IUPHAR	114	1,796
MICAD	249	116
PDSP_Ki	605	40,154
PharmGKB	22,677	1,777
TTD	1,518	3,190
UniProt	86,605	23,515
ChEBI		20,345
Ligand Expo		10,857
PubChem Compound		35,292,189
PubChem Substance		88,222,926
<b>Total entries before integration</b>	12,168,716	124,848,656
<b>Total entries after integration</b>	11,232,904	28,730,299