DELvision: A Protein-DEL Interaction Database Established by Using a High-throughput Workflow and Its Application in Tractability Assessment

Wanfa Yang, Huan Liu, Jinming Guan, Weiren Cui, Zhongyao Ma, Wenji Su, Letian Kuai

WuXi AppTec Headquarters, 288 Fute Zhong Road Waigaoqiao Free Trade Zone, Pudong District, Shanghai, 200131, China.

Abstract

Describing the characteristics of protein-ligand interactions (PLIs) and finding novel interactions are essential in the fields of studying life science topics and developing new therapeutics [1-3]. However, there lack efficient experimental strategies for collection of proteome-level PLI data of high quality and consistent formats that refer to a wide range of overlapping small molecule chemical space. In this work, we demonstrate that the DEL technology can be applied to profiling vast chemical space for their binding capabilities with thousands of proteins for PLI data collection. We established the database named DELvision, which consists of 1511 protein-DEL interaction datasets by using an in-house built high-throughput DEL screening workflow coupled with high-throughput protein production platform. Post-selection data analysis proved that the high-throughput strategy was consistent across different selection batches, and that the protein background noise was manageable. Further On-DNA and OFF-DNA ASMS validations on p38 family targets proved the reliability of DELvision data. Our study demonstrates the possibility to profile protein-ligand interactions by high-throughput DNA-encoded library screening.

Overview of DELvision Workflow

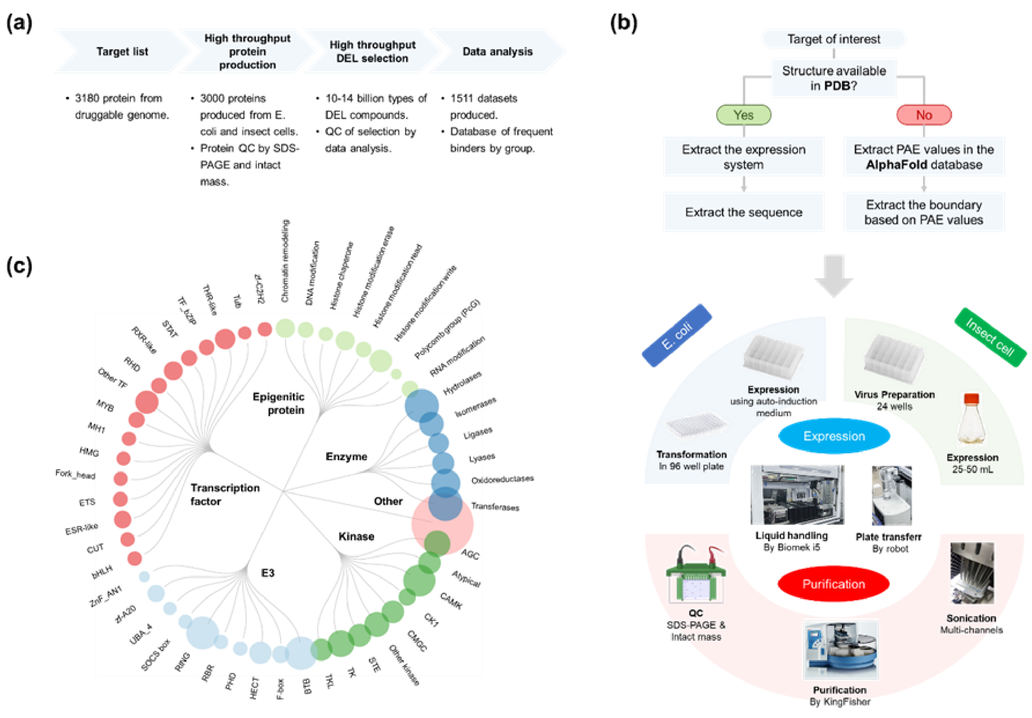


Figure 1. Schematic diagram of DELvision database generation workflow and covered protein classes. (a) High-throughput workflow for data production of DELvision. (b) Graphical representation of the design, expression and purification of protein constructs. (c) Overview of DELvision datasets by protein classes.

Quality Control of High Throughput DEL Selection Processes

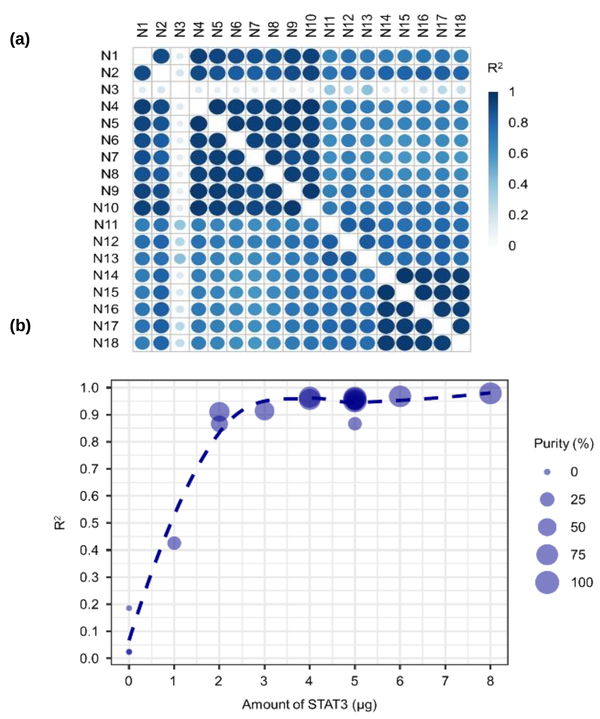


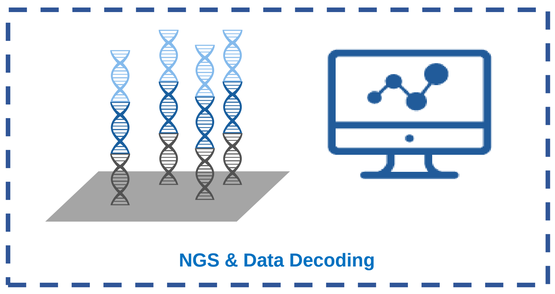
Table 1. Detailed Protein Sample input For Figure.2b

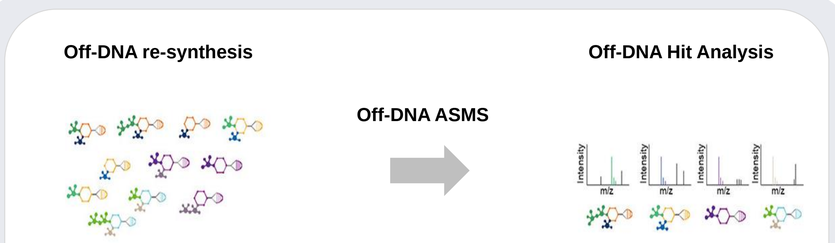
Condition	Amount of STAT3 (μg)	Amount of RIPL (μg)	Total amount (μg)	Purity of STAT3 (%)
C1	10	0	10	100.0
C2	8	1	9	88.9
C3	6	1	7	85.7
C4	4	1	5	80.0
C5	2	1	3	66.7
C6	0	1	1	0.0
C7	5	0	5	100.0
C8	4	1	5	80.0
C9	3	2	5	60.0
C10	2	3	5	40.0
C11	1	4	5	20.0
C12	0	5	5	0.0
C13	5	0	5	100.0
C14	5	1.25	6.25	80.0
C15	5	3.33	8.33	60.0
C16	5	7.5	12.5	40.0
C17	5	20	25	20.0
C18*	0	0	0	0.0

Figure 2. Pairwise comparison demonstrate consistency in high-throughput selection and selection results is correlated to input protein amount (a) The R^2 of chemotype enrichment scores across 18 NTC datasets obtained by using the same DEL selection set-up.. (b) The correlation between the amount of STAT3 and pairwise R^2 values for C1 versus C2-C18.

Discovery and Validation of Potential Selective Compound hits against p38 Family(MAPK12/13/14) by DELvision Database

Figure 3. Discovery of potential selective compound hits against p38 family (MAPK12/13/14) by DELvision database. (a) The general scaffold of the cluster Vision020-0-17-285-162 and structure of Compound 1. The selection output of the cluster for (b) MAPK12, showing 160 compounds (41474.5≤enrichment≤3649752.0), (c) MAPK13, showing 29 compounds (33655.7≤enrichment≤100967.0), and (d) MAPK14, showing 3 compounds (enrichment=14515.3), respectively (BB4=162).





Schematic diagram of Off-DNA validation

Off-DNA/ASMS	Recovery Rate (no heating)	Recovery Rate (heating)
MAPK12	28.27%	32.32%
MAPK13	23.89%	27.31%
MAPK14	/	/
SUMO	/	/
Beads matrix	/	/

Table 2. OFF-DNA ASMS validation of potential DEL-vision discovered hits against P38 family. The OFF-DNA ASMS validation shows that DEL vision database generated hits has good selectivity against target protein, with significant binding signals can be detected against Target Protein (MAPK12,MAPK13) and not binding signals can be detected against MAPK14.

Summary

To our knowledge, this study represents the first use of DEL in establishing protein-DEL interaction database, where we experimentally generated DELvision, comprising of 1511 datasets of various types of proteins through a high-throughput strategy. The count of signals in each dataset ranges from tens of thousands to millions. In the future, the DELvision database can contribute in bridging the gap between a large chemical space and the druggable genome.

References

[1] Carter A J, Kraemer O, Zwick M, et al. Target 2035: probing the human proteome[J]. Drug Discovery Today, 2019, 24(11): 2111-2115.
[2] Machutta C A, Kollmann C S, Lind K E, et al. Prioritizing multiple therapeutic targets in parallel using automated DNA-encoded library screening[J]. Nature communications, 2017, 8(1): 16081.
[3] Zhao, Z.; Bourne, P. E. Harnessing Systematic Protein-Ligand Interaction Fingerprints for Drug Discovery. Drug Discov. Today 2022, 27 (10), 103319.

Contact

www.wuxibiology.com

Business Contact: Mahnaz_Arjomand@wuxiapptec.com (US)
Business Contact: dave_madge@wuxiapptec.com (EU and Israel)
Business Contact: xu_longji@wuxiapptec.com (China)
Business Contact: sycho@wuxiapptec.com (Korea)
Business Contact: fumio_itoh@wuxiapptec.com (Japan)
Technical Contact: DB_Early_Discovery_Business_Transformation@wuxiapptec.com