Spresso:
An ultrafast compound pre-screening method based on compound decomposition

Keisuke Yanagisawa\textsuperscript{1,2}  
Shunta Komine\textsuperscript{2,3}  
Shogo D. Suzuki\textsuperscript{2,3}  
Takashi Ishida\textsuperscript{1,2,3,4}  
Masahito Ohue\textsuperscript{1,3,4}  
Yutaka Akiyama\textsuperscript{1,2,3,4}

\textsuperscript{1}. Dept. CS, School of Computing, Tokyo Tech.  
\textsuperscript{2}. Education Academy of Computational Life Sciences (ACLS), Tokyo Tech.  
\textsuperscript{3}. Dept. CS, Graduate School of Information Science and Engineering, Tokyo Tech.  
\textsuperscript{4}. Advanced Computational Drug Discovery Unit (ACDD), Tokyo Tech.

* The software was renamed from ESPRESSO to Spresso (same pronunciation) in October 2016.
**S**peedy **PRE-**S**creening method**

with **S**egmented **cOMPounds**

( [http://www.bi.cs.titech.ac.jp/spresso/](http://www.bi.cs.titech.ac.jp/spresso/) )
Docking-based virtual screening

Virtual screening

Compound DB  →  Drug candidates

Docking calculation

A protein & a compound  →  A conformation
Conformation search

Translation
3 dimensions

Rotation
3 dimensions

Internal rotation
N dimensions

Problem: **Computationally expensive**
Compound pre-screening

Decreasing calculation with pre-screening

- Compound DB
- Pre-screening
- Docking
- Drug candidates

Ex: 10,000,000 compounds × 1/100 → 100,000 compounds → 1,000 compounds
Existing pre-screening methods

Pre-screening

Ligand-based
- ML\(^1\), Shape matching\(^2\) etc.
- Requiring known compounds resulting in difficulty finding novel drug candidates
- Light weight calculation

Structure-based
- Rough docking (Glide HTVS\(^3,4\))
- No bias toward known actives
- Insufficient speed (1 sec / compound = 1 CPU year for ZINC DB)

---

1) H. Geppert, et al. *J Chem Inf Model*, 2010.
2) A. Vuorinen, et al. *J Med Chem*, 2014.
3) S. B. Mirza, et al. *J Mol Graph Model*, 2016.
4) A. Grover, et al. *Biochim Biophys Acta*, 2012.
Our approach

- **Structure-based**
- **Ultrafast** compared to existing method

|                  | Ligand-based | Structure-based | Spresso |
|------------------|--------------|-----------------|---------|
| wo/ known compound | 😞           | ☺              | ☻       |
| pre-screening speed | ☻            | 😞              | ☻       |

© 2016 Keisuke YANAGISAWA
Ideas for acceleration

Idea I. Compound decomposition

Idea II. Rough compound evaluation

Fragment

score = x  
score = y  
score = z

Compound

score = w
Idea I. Compound decomposition

Creating **fragments** without any rotatable bond

1) S. Komine et al., *IPSJ SIG Technical Report*, 2015-BIO-42, 2015.
Another benefit of decomposition

Sharing of fragment docking results for duplication

Example of duplicated fragments

Decomposition result compounds: ZINC “drugs now” subset (10,639,555 entries)
Idea II. Rough compound evaluation

Compound evaluation without re-construction

Generalized Sum-3 (GS₃) of fragment scores is adopted

\[ GS₃ = \sqrt[3]{\sum_f (\text{score}_f)^3} \]
Method

Decomposition

Fragment Docking by Glide SP or Glide HTVS

Evaluation

 compounds

fragments

fragments

compounds

© 2016 Keisuke YANAGISAWA
Spresso is **open-sourced** under GPLv3 license (http://www.bi.cs.titech.ac.jp/spresso)

**Acknowledgements**

東京工業大学  Tokyo Institute of Technology

東京工業大学 情報生命博士教育院 Education Academy of Computational Life Sciences

国立研究開発法人 科学技術振興機構 Japan Science and Technology Agency

科研費 KAKENHI

© 2016 Keisuke YANAGISAWA