Assessment of EMA in CASP14-COVID

Sohee Kwon, Jonghun Won, Andriy Kryshtafovych, and Chaok Seok

Department of Chemistry, Seoul National University
Genome Center, University of California, Davis

Prediction Center
CASP-COVID EMA Statistics

Targets with released structures: C1905 and C1908 (=T1064)

**Number of models (N{GDT ≥40})**

<table>
<thead>
<tr>
<th>Targets</th>
<th>C1905</th>
<th>C1905-D1</th>
<th>C1905-D2</th>
<th>C1908</th>
</tr>
</thead>
<tbody>
<tr>
<td>CASP-COVID</td>
<td>153 (6) [5 AF1]</td>
<td>83 (38) [0 AF]</td>
<td>79 (0) [0 AF]</td>
<td>181 (0) [0 AF]</td>
</tr>
<tr>
<td>Post-analysis</td>
<td>+ 5 AF1 (5)</td>
<td>+ 5 AF1 (5)</td>
<td>+ 5 AF2 (5)</td>
<td></td>
</tr>
</tbody>
</table>

**Number of QA groups**

<table>
<thead>
<tr>
<th></th>
<th>Full</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>CASP-COVID</td>
<td>29/12</td>
<td>19/10</td>
</tr>
<tr>
<td>Post-analysis</td>
<td>10/10</td>
<td>10/10</td>
</tr>
</tbody>
</table>
1. C1905

Modeling results

QA results

Selected by QA: [BAKER]
QA results (C1905)

<table>
<thead>
<tr>
<th></th>
<th>GDT-TS</th>
<th>IDDT</th>
<th>MolP</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF1_2</td>
<td>59</td>
<td>59</td>
<td>2.0</td>
</tr>
<tr>
<td>FEIGLAB-R</td>
<td>42</td>
<td>43</td>
<td>1.2</td>
</tr>
</tbody>
</table>
Global QA: BAKER

Local QA: BAKER

<table>
<thead>
<tr>
<th></th>
<th>ASE</th>
<th>AUC</th>
<th>ULR F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF1_2</td>
<td>71</td>
<td>89</td>
<td>0.2</td>
</tr>
<tr>
<td>FEIGLAB-R</td>
<td>51</td>
<td>65</td>
<td>0.2</td>
</tr>
</tbody>
</table>
2. C1905-D1

Modeling results

QA results

Selected by QA:
Selected by QA: (including AF1)

3/4 groups
3. C1905-D2

Modeling results

QA results

Selected by QA: ■
Selected by QA: ■
(including AF1)
EMA of AF1 models for by 10 QA groups

- For both C1905-D1 and C1905-D2, none of 10 groups selected the best models as top1.
- QA score and GDT-TS/IDDT show mostly negative correlations.

Best correlation: QMEANDisCo C1905-D1 AF1 models
4. C1908 (T1064)

Modeling results

QA results

Selected by QA: ■ (including AF2)
Global QA score

VoroMQA-dark

IDDT

GDT-TS

MolProbity ~ 1

MolProbity ~ 2
C1908 structure in detail

AF2_1: GDT-TS = 87

Crystal contact

Five models of AF2
C1908: Local QA

Distance Error (Å)

Predicted distance error by PrQ3D

AUC 0.888

Actual error

Predicted error

True distance error

Predicted distance error by PrQ3D
Conclusion

- Some QA methods can discriminate AF models from non-AF models.
- Current QA methods have difficulty in selecting models of high global structural accuracy involving domain orientation.
- Current QA methods are only partially successful in identifying inaccurately/accurately modelled regions.