
SUPPLEMENTARY MATERIAL

Splice2Deep: An Ensemble of Deep Convolutional Neural Networks for Improved Splice Site Prediction in Genomic DNA

Somayah Albaradei^{1,2*}, Arturo Magana-Mora^{1,3*}, Maha Thafar^{1,4}, Mahmut Uludag¹, Vladimir B. Bajic¹, Takashi Gojobori^{1,5}, Magbubah Essack^{1#}, Boris R. Jankovic^{1#}

¹Computer, Electrical and Mathematical Sciences and Engineering Division (CEMSE), Computational Bioscience Research Center, Computer (CBRC), King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia.

²Faculty of Computing and Information Technology, King Abdulaziz University, Kingdom of Saudi Arabia.

³Saudi Aramco, EXPEC-ARC, Drilling Technology Team, Dhahran, 31311, Kingdom of Saudi Arabia.

⁴Faculty of Computers and Information Systems, Taif University, Kingdom of Saudi Arabia.

⁵Biological and Environmental Sciences and Engineering Division (BESE), King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia.

*First author.

#To whom correspondence should be addressed. Tel: +966 (54) 470 0746; Email: magbubah.essack@kaust.edu.sa; Email: boris.jankovic@kaust.edu.sa

Supplementary Table S6. Accuracy (Acc), Sensitivity (Sn), and Specificity (Sp) from the cross-organism validation obtained by using the *Caenorhabditis elegans* model and testing on the other organisms. N/A indicates that the tool is not available for that specific organism.

| Splice site | Model | <i>Homo sapiens</i> | | | <i>Arabidopsis thaliana</i> | | | <i>Oryza sativa japonica</i> | | | <i>Drosophila melanogaster</i> | | | <i>Caenorhabditis elegans</i> | | |
|-------------|-----------------|---------------------|--------|--------|-----------------------------|--------|--------|------------------------------|--------|--------|--------------------------------|--------|--------|-------------------------------|--------|--------|
| | | Sp. | Sn. | Acc. | Sp. | Sn. | Acc. | Sp. | Sn. | Acc. | Sp. | Sn. | Acc. | Sp. | Sn. | Acc. |
| AcSS | Genesplicer | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| | SplicePredictor | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| | DeepSS | 0.8904 | 0.8914 | 0.8909 | 0.8904 | 0.8954 | 0.8929 | 0.8704 | 0.8804 | 0.8919 | 0.9404 | 0.9309 | 0.9207 | 0.9049 | 0.9086 | 0.9332 |
| | Splicerover | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| | Splice2Deep | 0.9404 | 0.9414 | 0.9409 | 0.9304 | 0.9354 | 0.9329 | 0.9104 | 0.9134 | 0.9119 | 0.9604 | 0.9509 | 0.9407 | 0.9778 | 0.9838 | 0.9808 |
| DoSS | Genesplicer | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| | SplicePredictor | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| | DeepSS | 0.9104 | 0.8934 | 0.9019 | 0.9204 | 0.9054 | 0.9129 | 0.8804 | 0.8904 | 0.9019 | 0.9504 | 0.9309 | 0.9405 | 0.9194 | 0.9168 | 0.940 |
| | Splicerover | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A | N/A |
| | Splice2Deep | 0.9604 | 0.9434 | 0.9519 | 0.9604 | 0.9454 | 0.9529 | 0.9204 | 0.9234 | 0.9219 | 0.9704 | 0.9509 | 0.9497 | 0.9774 | 0.9763 | 0.9768 |