

1. Stallinga, S. & Rieger, B. *Opt. Express* **18**, 24461–24476 (2010).
2. Mortensen, K.I., Churchman, L.S., Spudich, J.A. & Flyvbjerg, H. *Nat. Methods* **7**, 377–381 (2010).
3. van de Linde, S. *et al. Nat. Protoc.* **6**, 991–1009 (2011).
4. Ram, S., Prabhat, P., Chao, J., Ward, E.S. & Ober, R.J. *Biophys. J.* **95**, 6025–6043 (2008).
5. Löscherger, A. *et al. J. Cell Sci.* **125**, 570–575 (2012).
6. Smith, C.S., Joseph, N., Rieger, B. & Lidke, K.A. *Nat. Methods* **7**, 373–375 (2010).

ORCAE: online resource for community annotation of eukaryotes

To the Editor: Genome sequencing and annotation projects are no longer restricted to large sequencing institutes. With high-throughput next-generation sequencing technologies, draft genome sequences are now being determined at an unprecedented pace. However, more and more genome sequences are poorly annotated. To bridge this gap, there is an increasing need for easy-to-use, online genome-annotation tools¹. Moreover, genome projects are being conducted by smaller consortia, and thus genome annotation should benefit maximally from community annotation efforts² that reflect the knowledge available in the scientific community³. Therefore, we developed ORCAE (<http://bioinformatics.psb.ugent.be/orcae/>, formerly known as BOGAS), a wiki-style annotation portal. ORCAE offers public access to a wide variety of plant, fungal and animal genomes (**Supplementary Table 1**).

The basic setup of ORCAE is highly comparable to wiki systems such as MediaWiki, and the information page for each gene can be seen as a ‘topic’ page of a traditional text wiki. We developed ORCAE with a gene-centric vision in mind, which is reflected by the fact that the gene information pages (**Fig. 1** and **Supplementary Figs. 1** and **2**) are the central pages in the system (**Supplementary Methods**), in contrast to chromosome-centric systems such as GNPannot⁴. The system is ideally suited to allow multiple people to simultaneously curate the initial automatic predictions, a typical setup in many genome projects where curators with diverse expertise join forces to improve the genome annotation. ORCAE offers functionality to browse, annotate and manage eukaryotic genome annotation projects. The system can be used both to present published genomes to the scientific community and to coordinate manual annotation and curation efforts in ongoing annotation projects. This is one of ORCAE’s main advantages because users have a central access point where annotations can be queried or improved.

For publicly available genomes, everyone is allowed to view all gene pages. Restricted genome projects are only accessible for registered users. Editing privileges are limited to participants in the particular genome project, whereas nonmembers of the consortium can receive editing privileges when approved by the consortium (**Supplementary Methods**).

The portal will keep track of all the modifications to both functional and structural annotations. This is done in a wiki-compliant manner: all edits are stored in the database, and thus a history

The screenshot displays the ORCAE interface for the gene MD10G000070.1 in *Malus domestica*. The top section shows the gene ID, locus, functional description, gene type, contig, and history. Below this is a 'Gene Structure' section with a graphical representation of the gene's structure, showing exons as boxes and introns as lines. The interface also includes navigation options like 'VIEW IN GENOMVIEW' and 'VIEW IN ARTEMIS', and a 'Modify This Record' button.

Figure 1 | Gene page in the ORCAE resource. Through the extensive use of graphical representations, a clear overview of the data is provided to assist users in easily assessing the quality and accuracy of the offered annotations for a given gene locus.

of modifications is available for each locus (**Supplementary Methods**). This allows ORCAE to eliminate a bottleneck found in other online annotation tools, such as yrGATE⁵, where a curator needs to approve modifications before they become available. Indeed, in the spirit of community annotation, the quality of the presented annotation is the responsibility of the whole community. However, to ensure a certain level of quality, several automatic checks have been put in place that encompass all existing knowledge on the gene structure and functions (for example, whether a KOG or EC number exist) (**Supplementary Methods**). Finally, ORCAE is highly dynamic: upon modification of a gene model, all the available information is immediately updated and presented on the gene page.

Note: Supplementary information is available at <http://www.nature.com/doi/funder/10.1038/nmeth.2242>.

ACKNOWLEDGMENTS

We thank S. Rombauts, Y.C. Lin, B. Verhelst, S. De Bodt and T. Van Parys for their help in developing the portal and in the preparation of this manuscript. This work was supported by the Ghent University Multidisciplinary Research Partnership ‘Bioinformatics: from nucleotides to networks’ and Research Foundation Flanders (FWO) (awarded to T.A.).

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Lieven Sterck^{1,2}, Kenny Billiau^{1,2}, Thomas Abeel^{1,2}, Pierre Rouzé^{1,2} & Yves Van de Peer^{1,2}

¹Department of Plant Systems Biology, VIB, Ghent, Belgium. ²Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium. e-mail: lieven.sterck@psb.vib-ugent.be

1. Parkhill, J., Birney, E. & Kersey, P. *Genome Biol.* **11**, 402 (2010).
2. Salzberg, S.L. *Genome Biol.* **8**, 102 (2007).
3. Schlueter, S.D., Wilkerson, M.D., Huala, E., Rhee, S.Y. & Brendel, V. *Trends Plant Sci.* **10**, 9–14 (2005).
4. Guignon, V. *et al. Bioinformatics* **28**, 1054–1056 (2012).
5. Wilkerson, M.D., Schlueter, S.D. & Brendel, V. *Genome Biol.* **7**, R58 (2006).