Managing the life cycle of a portfolio of Open Data Resources

Chiara Gabella PhD
Swiss Institute of Bioinformatics
Lausanne, Switzerland
https://orcid.org/0000-0002-7104-5025

Séverine Duvaud
Swiss Institute of Bioinformatics
Lausanne, Switzerland
https://orcid.org/0000-0001-7892-9678

Christine Durinx PhD
Swiss Institute of Bioinformatics
Lausanne, Switzerland
https://orcid.org/0000-0003-4237-8899

Abstract—*Data resources are essential for the long-term preservation of scientific data and the reproducibility of science. The SIB Swiss Institute of Bioinformatics provides the life science community with a stable and reliable data infrastructure. The Institute provides a portfolio of high-quality databases and software platforms, which vary from expert curated knowledgebases such as UniProtKB/Swiss-Prot (part of the UniProt consortium), neXtProt, STRING and Bgee, to online tools such as SWISS-MODEL and SwissDrugDesign. SIB aims to ensure that these SIB Resources are available on the long term, i.e. as long as their scientific return-on-investment and impact are high. SIB's vision is that data and research results should be freely accessible to all. For this reason, the Institute promotes the adoption of open licenses. This paper describes the processes that support the identification, evaluation, and development of SIB Resources. For over ten years, SIB employs a set of indicators that reflect the multiple facets and complexity of data resources. These indicators set quality standards, and monitor usage trends and scientific impact. They guide and inform both the SIB Scientific Advisory Board in its evaluation, and the managers of the databases and tools in the development of their respective resources towards scientific excellence and Open Science. Through a professional management framework with central services such as user-centric design and a license advisory committee, SIB supports the promotion of excellence in resource development and operation.*

Keywords—Bioinformatics, Open Data, Database, Software tools, Infrastructure, User Experience (UX), High-quality integrated resources, Sustainability.

I. INTRODUCTION

The SIB Swiss Institute of Bioinformatics (www.sib.swiss) is an internationally recognized non-profit organization, dedicated to biological and biomedical data science. It is present in the main academic institutions in Switzerland (Fig. 1) and leads numerous national and international projects with a major impact on life science research and health.

SIB's scientists are passionate about creating knowledge and converting complex questions into solutions in many fields, from biodiversity and evolution to medicine. They provide essential resources, such as databases and software platforms, as well as data management, software engineering, biocuration services, computational biology know-how and training in bioinformatics. The Institute delivers this expertise to academic groups and clinicians as well as to private companies.

SIB fosters collaboration and knowledge sharing among some 800 scientists across Switzerland and represents the Swiss bioinformatics community, collaborating with international institutions on global research infrastructures. The Institute contributes to keeping Switzerland at the forefront of innovation by promoting progress in biological research and enhancing health.

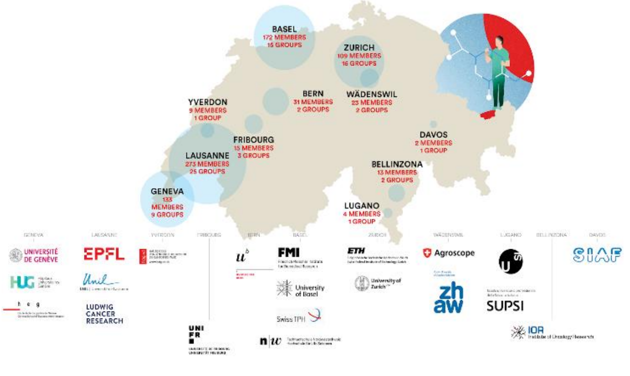

Fig. 1: Map of the SIB partner institutions as of December 2020. In January 2021, EMPA (the Swiss Federal Laboratories for Materials Science and Technology) in St. Gallen and the University Hospital of Zurich became new institutional members of SIB.

A study by Attwood et al. looked at the 18-year survival of 326 publicly available biological databases (Attwood et al., 2015). Over 60% "died" within that time period, and a further 14% were archived, i.e. no longer updated. The instability of the existence of major data resources is associated with the risk of losing an immense wealth of biological information, and the associated investments. The study shows that a viable, sustainable framework for long-term data stewardship is sorely needed. Indeed, the databases that survived were, for the most part, important to their institution's main focus, and had core institutional support. Database longevity depends on the existence of infrastructures that are underpinned by long-term strategies. Moreover, although data resources play an essential role in scientific research, a sustainable funding model that ensures their maintenance and development remains a critical challenge (Gabella et al., 2018). Within the limit of available funding, SIB's commitment is to ensure the long-term existence of the SIB Resources in order to provide a stable environment for the development, enhancement and maintenance of high-quality databases and software tools to support the life science community.

II. THE LIFE CYCLE OF A BIOINFORMATICS RESOURCE

A new database or software tool typically starts with a research project, leading to a proof-of-concept. Through further development, the resource evolves towards maturity and may become part of the research infrastructure of the scientific community (Fig. 2).

The SIB Groups develop and maintain numerous cutting-edge resources that are made available to the community. Among them, several key resources - the SIB Resources - benefit from the Institute's specific support after a careful selection process and are included in the SIB portfolio.

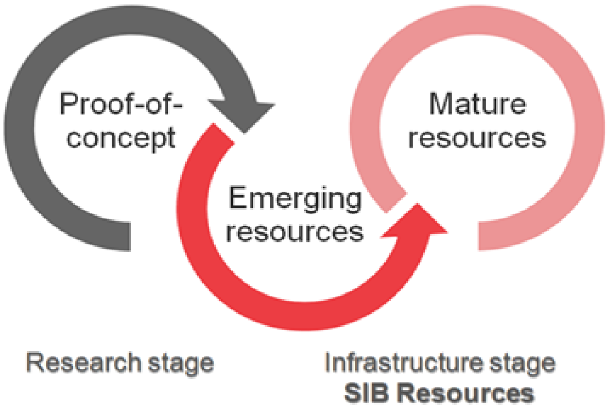

Fig. 2: The life cycle of a bioinformatics resource.

SIB's resource portfolio contains promising emerging resources, such as SwissLipids (Aimo et al., 2015) or V-pipe (Posada-Céspedes et al., 2020), as well as fully mature ones. To promote new developments at the highest level of scientific excellence, SIB is committed to identifying and supporting additions to its resource portfolio among the emerging or already well-established resources developed by SIB groups, and which have a (potentially) high impact on the life science community. These resources must have passed the proof-of-concept stage and reached a sufficient level of development and usage to be considered as infrastructure. They should demonstrate their uniqueness, a strong demand from the target community and alignment with Open Data practices, in addition to high scientific standards.

SIB's high-quality infrastructure, i.e. databases and software tools, as well as the associated services, are driven by excellence in research. SIB Resources are typically linked to research work and often embedded in a research group, in order to remain state-of-the-art (SIB Swiss Institute of Bioinformatics Members, 2016).

Among the SIB Resources we count the UniProtKB/Swiss-Prot database, which is part of the UniProt Consortium (Poux et al., 2017; The UniProt Consortium, 2019). The knowledgebase contains a reviewed collection of high-quality annotated and non-redundant protein sequences, bringing together experimental results, computed features and scientific conclusions to provide information related to a protein's function, structure and subcellular location, specific features and interactions. The UniProtKB/Swiss-Prot database contains over half a million expert-curated protein sequences. A team of highly qualified scientists —called biocurators — who are based in Geneva, select, review and annotate the information. The biocurators are supported by advanced machine learning models that automatically identify and classify relevant publications for review (Lee et al., 2018). With a million unique users per month, UniProt is the most widely used protein information resource in the world.

Another example of a SIB Resource is the SWISS-MODEL Workspace (Waterhouse et al., 2018). This fully automated web-based service assists and guides the user in building a three-dimensional structure of a protein, based on its homology with proteins for which experimentally determined structures are available. SWISS-MODEL receives over a million model requests every year.

STRING (Szklarczyk et al., 2019), another SIB Resource, is a knowledgebase and software tool for known and predicted protein-protein interactions. It includes direct (physical) and indirect (functional) associations derived from various sources, such as genomic context, high-throughput experiments, (conserved) co-expression and the literature. STRING networks cover over 5,000 different organisms with over 25 million high-confidence links between proteins.

The full list of SIB Resources is available in appendix A.

III. A COMPREHENSIVE SET OF INDICATORS AS GUIDE ALONG A RESOURCE'S LIFE CYCLE

SIB is bound to the State Secretariat for Education, Research and Innovation SERI through a Service Level Agreement (SLA), which covers the funding to SIB for the provision of bioinformatics resources to the life science community. Every four years, the SIB Board of Directors (BoD) is responsible for selecting best-in-class resources (i.e. SIB Resources), as well as for the allocation of the federal funds to said resources[9]. Decisions are based on the recommendations of the Institute's external Scientific Advisory Board (SAB) that, in some cases, also takes input from evaluations by external reviewers. Members of the BoD do not take part in decisions for which they have a conflict of interest.

Three criteria define a best-in-class resource at SIB: scientific impact, scientific return-on-investment, and its fit within the resource portfolio. Scientific impact is defined as a combination of the scientific state-of-the-art, utility, and use. The (expected) scientific ROI of funding is estimated, i.e. the impact in terms of serving more users, of filling an important unmet need of the scientific community, and of accelerating science. And last but not least, careful attention goes also to the alignment of SIB's portfolio of databases and software tools with the Institute's core competencies and strategic focus.

The SAB assesses whether best-in-class criteria are met by using a set of 28 indicators that are grouped in six categories (see below). The responsible resource is required to submit a workplan based on these indicators (a template of the workplan for candidate SIB Resources, with the full list of indicators, is available in Table I. For more details see Gabella et al. (2020)).

Databases and software tools being very diverse, it is therefore crucial to take into account their many different facets into the indicators for evaluation. Providing precise information and figures allows then the panel of external reviewers and the SAB to make an objective recommendation.

These indicators can also be helpful for the scientists developing a resource to guide the development process. Indicators are assessed differently depending on the type of resource: it is the whole body of indicators together that reflects the quality and impact of a specific bioinformatics resource.

Indicators for SIB Resources are grouped into the following six categories:

Category I, Scientific focus and quality of science: Demonstrate high quality of data and metadata, respond to a clear scientific need, and be unique. This implies benchmarking against other resources, and being an authority in its field compared to the major competitors.

TABLE I. INDICATORS FOR A CANDIDATE SIB RESOURCE. MORE INFORMATION ARE AVAILABLE IN Gabella et al. (2020)

I: Scientific focus and quality of science		
a	Definition	The Resource is a ☐ Database ☐ Software tool ☐ Database & Software tool
b	Scope statement	Describe the scope and scientific coverage of the Resource (for example, all species or a subset of species, families, outputs from a particular experimental method), as well as the scientific need to which it responds.
c	Uniqueness of the resource	Indicate who are the major competitors of the Resource and how the Resource compares to them including benchmarking efforts.
d	Potential synergies and collaborations with other SIB Resources	Describe the potential synergies and collaborations with other SIB Resources that could increase the impact of the (respective) Resource(s).
e	Objectives and implementation plan for the next funding period	Describe the objectives and implementation plan for the next funding period.
II: Community		
a	User community	Whom is the Resource addressed to? Describe the current user community and the size of the potential user community. Are there other user communities that are currently not yet reached? Include quantitative measurements, if possible.
b	Overall usage - web access	Web access as measured by Google Analytics. Please join the following extract from Google Analytics (period: from 2018 to present, or from more recently if not yet available in 2018; view by month): - Audience > overview - Acquisition > overview

---

[9] Article 7.2 of the SIB Statutes

<table>
<tr><td></td><td></td><td>- Behaviour > overview</td></tr>
<tr><td></td><td>Overall usage - additional access methods</td><td>Please give figures to quantify any additional access methods: visits, unique visitors, hits, and downloads (includes FTP downloads and programmatic access)</td></tr>
<tr><td>c</td><td>Overall usage (software tools / service platforms)</td><td>If available, provide statistics about the number of jobs submitted and/or the number of downloads of the software tool and/or any other measurements that assess the usage of the Resource.

Example:
<table>
<tr><td>Type of measurement</td><td>Value at 31/12/2017</td><td>Value at 31/12/2018</td><td>Value at 31/12/2019</td><td>Value at 31/07/2020</td></tr>
<tr><td>Jobs submitted</td><td>100</td><td>200</td><td>400</td><td>500</td></tr>
</table></td></tr>
<tr><td>d</td><td>Usage in research as measured through citation in the literature</td><td>Go to https://europepmc.org/ and run a query using the relevant keywords, which best identify the Resource. Copy and paste the resulting URL, together with the number of citations (example: https://europepmc.org/search?query=resourceAAB%20%2B%20organism: 35 citations). If necessary, indicate more than one URL, but do not exceed 3.
If any extra explanation is needed, you can also add it to the text field. Moreover, if other citation counting tools are more representative of the usage of the Resource, please feel free to report the additional statistics.</td></tr>
<tr><td>e</td><td>Dependency of other resources</td><td>Indicate which resources depend on the Resource.

Examples:
● The UniLectin platform depends on SWISS-MODEL to display molecular interactions between small molecules and protein structures.
● PeptideAtlas uses information about single amino acid variants and post-translational modifications in neXtProt to enable processing and display of data</td></tr>
<tr><td colspan="3"><i>III: Quality of Service</i></td></tr>
<tr><td>a</td><td>Unique ID</td><td>Indicate if the Resource provides persistent and unique identifiers and if yes, describe it.</td></tr>
<tr><td>b</td><td>Data entries/records (data resources)</td><td>Cumulative total number of entries or records to indicate the growth of the Resource. Please provide a description and explanation on what an “entry” is.

Example:
<table>
<tr><td>Entry</td><td>Description</td><td>Value at 31/121/2017</td><td>Value at 31/12/2018</td><td>Value at 31/12/2019</td><td>Value at 31/07/2020</td></tr>
<tr><td>Species</td><td>Number of Species in the Resource</td><td>100</td><td>200</td><td>300</td><td>400</td></tr>
<tr><td>Genes</td><td>Number of genes annotated in the Resource</td><td>1000</td><td>2000</td><td>3000</td><td>4000</td></tr>
<tr><td>Variants</td><td>Number of sequence variations present in the Resource</td><td>200</td><td>250</td><td>400</td><td>410</td></tr>
</table></td></tr>
<tr><td>c</td><td>Use of community-recognized standards for (meta)data</td><td>Which community-recognized standards are used for metadata and data (e.g. MIAME, JATS, INSDC features, ontologies)? Provide a link to documentation.</td></tr>
<tr><td>d</td><td>Data availability & access (data resources)</td><td>Data sharing services: list services through which data is shared (e.g. website, APIs, FTP, TripleStore)

Data sharing formats: list formats for available data (e.g. plain text, FASTA, XML, RDF, Dublin Core, tsv, JSON)</td></tr>
<tr><td>e</td><td>Customer service</td><td>Helpdesk: does the Resource run a helpdesk?

User feedback: how does the Resource seek and incorporate user input into service design decisions?

User training: which training is offered by the Resource? (face2face training, eLearning, etc.)</td></tr>
</table>

VI: Legal and funding infrastructure		
a	Scientific advisory board	Does the Resource have an external independent Advisory Committee with scientists and/or users (other than the SIB SAB)?
b	Legal framework supporting Open Science?	Does the Resource have terms of use or a licence that enables the reuse and remixing of data? (see Open Definition for a list of open licenses) If yes, please include a link to terms of use or state license designation (e.g. CC0, CC-BY, CC-BY-SA). Is the access to, and/or usage of, the Resource restricted in any way to the user or certain categories of users? If yes, please explain why. If the current legal framework does not support Open Science, do you plan to adopt an open licence?
c	Licensing code: is it published on open access?	Is the source code of the Resource published in open access (e.g. in GitHub)? If yes, please indicate under which licence and the relevant URLs.
d	Sustainable support and funding	Describe what has been, and what will be, undertaken to seek additional funds from other sources.
e	Estimation of funds	Please indicate an estimation of funds that will be needed from SIB (in CHF) for the next funding period (4 years). Detail (in person-months) for what kind of work SIB funding will be used (for example: XX person-months for biocuration, YY person-months for code development, ZZ person-months for dataset selection):
V: Impact and visibility		
a	Counterfactual	Description of how science would be affected if the Resource had not existed or was to disappear and not be replaced.
b	Accelerating science	Description of how much the Resource accelerates science: present one or more selected examples of how the Resource has been used by its user community, showcasing the importance of the Resource for advancing science. These “translational stories” help SIB to show its impact to audiences such as policy makers and funders and help the SAB to assess the Resource.
c	Visibility	What actions has the Resource taken to increase its visibility within its potential user community?
VI: SIB		
a	Benefits	How does your Resource contribute to SIB in terms of scientific credibility, visibility or any other aspect?
b	Participation	Describe your contributions to SIB in the last 2 years, and how you see your involvement in the coming years
c	SIB Portfolio of Resources	How does the Resource align with the current portfolio of SIB Resources?

Category II, Community: Know the community to whom it is addressed, its size and the usage: web statistics, user reach, and community size. Candidate resources with a valid track record of usage, responding to a clear need within the scientific community are more likely to be included as SIB Resources. However, emerging resources are encouraged to submit as well. The scientific context in which the resource operates should be taken into account. A resource that serves a small scientific community may not have as many users as a resource serving a broader interest, and yet it may reach 90% of the community it supports (coverage) and be crucial for the scientific work of that community.

Category III, Quality of the service: Demonstrate a high level of service and reliability with the integration of features such as persistent and unique identifiers, community-recognized standards, user support and training, as well as the integration of user feedback.

Category IV, Legal and funding infrastructure and governance: Have a sound legal framework supporting Open Science and seek complementary funds from other sources in order to ensure sustainable long-term funding.

Category V, Impact and translational stories: Have a significant impact on the life science community and be impact driven.

Category VI, SIB: Group Leaders who manage a SIB Resource are expected to show strong involvement in SIB. Therefore, this aspect is also taken into consideration in the evaluation.

For each category and for each of the new candidate resources, the external reviewers assign a score according to a 9-point scale, going from Poor (1) to Exceptional (9), supplemented by a short commentary. At the SAB meeting, both the existing and the short-listed candidate SIB Resources are evaluated, based on these categories. The SAB scores their workplans and makes recommendations regarding their funding level.

SIB's support to the identified best-in-class Resources takes two forms. On the one hand, SIB offers a portfolio of services that support professional infrastructure provision, including User Experience (UX) studies and design, hosting, best practices and knowledge sharing, a security audit, user training, as well as licensing and legal advice. On the other hand, SIB provides funds that allow hiring skilled personnel to develop and maintain the resource. The mission of SIB is to maintain a portfolio of SIB Resources over the long-term and fund the SIB Resources within this portfolio as long as they have a high impact in the life science community. It is the role of the SAB to evaluate this impact and give recommendations to the Board of Directors regarding the level of funding.

In principle, SIB commits to supporting resources along the entire funding period of 4 years. However, the SIB Resources are also evaluated by the SAB at mid-term. This mid-term review includes an evaluation of the progress made since the beginning of the period. Based on the outcome, the BoD can decide to adjust the funding level or stop funding the resource.

CONCLUSION

Thanks to its coherent portfolio, including both emerging and already well-established resources, SIB is a key driver of innovation in bioinformatics. The indicators developed for the evaluation and selection process, through continuous monitoring of usage trends and scientific impact of the Resources, inform their life-cycle management by providing strategic recommendations for mature resources and allowing promising resources to develop to their full potential.

The provision of a professional solid infrastructure ranging from user-centred design, user research, licensing consulting and funding, enables the SIB's resource portfolio to be at the forefront of scientific excellence and ensures its long term sustainability in a context of Open Science.

ACKNOWLEDGMENT

We thank the SIB Scientific Advisory Board for their precious and expert advices and the SIB Board of Directors for their commitment within SIB in taking all the decisions necessary to achieve the aims of the Institute. Overall, we want to thank all the SIB Group Leaders and the SIB Resources managers for their commitment in making a better science.

REFERENCES

Aimo, L., Liechti, R., Hyka-Nouspikel, N., Niknejad, A., Gleizes, A., Götz, L., Kuznetsov, D., David, F. P. A., van der Goot, F. G., Riezman, H., Bougueleret, L., Xenarios, I., & Bridge, A. (2015). The SwissLipids knowledgebase for lipid biology. *Bioinformatics*, 31(17), 2860–2866. https://doi.org/10.1093/bioinformatics/btv285

Attwood, T. K., Agit, B., & Ellis, L. B. M. (2015). Longevity of Biological Databases. *EMBnet.Journal*. https://doi.org/10.14806/ej.21.0.803

Gabella, C., Durinx, C., & Appel, R. (2018). Funding knowledgebases: Towards a sustainable funding model for the UniProt use case. *F1000Research*. https://doi.org/10.12688/f1000research.12989.2

Gabella, C., Durinx, C., & Appel, R. (2020). Selection of SIB Resources for the period 2021-2025. *Zenodo*. https://doi.org/10.5281/zenodo.3937334

Lee, K., Famiglietti, M. L., McMahon, A., Wei, C.-H., MacArthur, J. A. L., Poux, S., Breuza, L., Bridge, A., Cunningham, F., Xenarios, I., & Lu, Z. (2018). Scaling up data curation using deep learning: An application to literature triage in genomic variation resources. *PLoS Computational Biology*. https://doi.org/10.1371/journal.pcbi.1006390

Posada-Céspedes, S., Seifert, D., Topolsky, I., Metzner, K., & Beerenwinkel, N. (2020). V-pipe: a computational pipeline for assessing viral genetic diversity from high-throughput sequencing data. *BioRxiv*, 2020.06.09.142919. https://doi.org/10.1101/2020.06.09.142919

Poux, S., Arighi, C. N., Magrane, M., Bateman, A., Wei, C.-H., Lu, Z., Boutet, E., Bye-A-Jee, H., Famiglietti, M. L., Roechert, B., & UniProt Consortium, T. (2017). On expert curation and scalability: UniProtKB/Swiss-Prot as a case study. *Bioinformatics*, 33(21), 3454–3460. https://doi.org/10.1093/bioinformatics/btx439

SIB Swiss Institute of Bioinformatics Members. (2016). The SIB Swiss Institute of bioinformatics' resources: Focus on curated databases. *Nucleic Acids Research*, 44(D1), D27–D37. https://doi.org/10.1093/nar/gkv1310

Szklarczyk, D., Gable, A. L., Lyon, D., Junge, A., Wyder, S., Huerta-Cepas, J., Simonovic, M., Doncheva, N. T., Morris, J. H., Bork, P., Jensen, L. J., & Von Mering, C. (2019). STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Research*, *47*(D1), D607–D613. https://doi.org/10.1093/nar/gky1131

The UniProt Consortium. (2019). UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Research*, *47*(D1), D506–D515. https://doi.org/10.1093/nar/gky1049

Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., De Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., & Schwede, T. (2018). SWISS-MODEL: Homology modelling of protein structures and complexes. *Nucleic Acids Research*, *46*(W1), W296–W303. https://doi.org/10.1093/nar/gky427

V. Appendix: The current SIB Resource Portfolio

Name of the SIB Resource		***Type***	***Description***	***Highlights***
Bgee	Gene expression expertise	Knowledgebase with expert curation and software tool	Gene expression data (including all types of transcriptomes), allowing retrieval and comparison of expression patterns between animals, humans, model organisms and diverse species of evolutionary or agronomical relevance.	Only resource to provide homologous gene expression between species.
EPD	Eukaryotic Promoter Database	Knowledgebase with expert curation and software tools	Quality-controlled information on experimentally defined promoters of higher organisms, as well as web-based tools for promoter analysis.	Over 180,000 promoters download- able, analysable over a web interface and viewable in the UCSC genome browser.
neXtProt	Human protein knowledgebase	Knowledgebase with expert curation and associated tools	Information on human proteins such as function, involvement in diseases, mRNA/protein expression, protein/protein interactions, post-translational modifications, protein variations and their phenotypic effects.	High data coverage through integration of multiple sources. Advanced semantic search functionalities. Tools specifically designed for the proteomics community.
STRING	Protein-protein Interaction Networks and Functional Enrichment Analysis	Knowledgebase and software tool	Resource for known and predicted protein-protein interactions, including direct (physical) and indirect (functional) associations derived from various sources, such as genomic context, high-throughput experiments, (conserved) co-expression and the literature.	An ELIXIR Core Data Resource. STRING networks cover over 5,000 different organisms with over 25 million high-confidence links between proteins

SwissDrugDesign	Widening access to computer-aided drug design	Software tools	Web-based computer-aided drug design tools, from molecular docking (SwissDock) to pharmacokinetics and druglikeness (SwissADME), through virtual screening (SwissSimilarity), lead optimization (SwissBioisostere) and target prediction of small molecules (SwissTargetPrediction).	Comprehensive and integrated web-based drug design environment
SWISS-MODEL	Protein structure homology-modelling	Software tool and repository	Automated protein structure homology-modelling platform for generating 3D models of a protein using a comparative approach, and database of annotated models for key reference proteomes based on UniProtKB.	Easy-to-use web-based platform processing over two million model requests per year, providing model information for experts and non-specialists.
SwissOrthology (OMA + OrthoDB)	One-stop shop for orthologs	Phylogenomic databases and software tools	Web portal of resources to infer orthologs, i.e. corresponding genes across different species, a key aspect to predicting gene function or reconstructing species trees. It includes OrthoDB, BUSCO as well as OMA and the Quest for Orthologs benchmark service.	World-leading orthology and comparative genomic resources.
SwissRegulon	Tools and data for regulatory genomics	Software tools and knowledgebases	Web portal for regulatory genomics, including genome-wide annotations of regulatory sites and motifs, the webserver ISMARA for automated inference of regulatory networks and CRUNCH for automated analysis of ChIP-seq data, and REALPHY for reconstructing phylogenies from raw sequence data.	ISMARA and Crunch web servers allow users to upload raw microarray, RNA-seq or ChIP-seq data to automatically infer the core regulatory networks acting in their system of interest.
UniProtKB/SwissProt	Protein knowledgebase	Knowledgebase with expert curation	Hundreds of thousands of protein descriptions, including function, domain structure, subcellular location, post-translational modifications and functionally characterized variants.	Expert-curated part of UniProt, the most widely used protein information resource in the world, with over six million pageviews per month. An ELIXIR Core Data Resource.
SwissLipids	A knowledge resource for lipids	Knowledgebase	Information about known lipids, including knowledge of lipid structures, metabolism and interactions, providing a framework for the integration of lipid and lipidomics data with biological knowledge and models.	Contains information on more than 590,000 lipid structures from over 640 lipid classes.
V-pipe	Viral genomics pipeline	Software tool	Pipeline integrating various open-source software packages for assessing viral genetic diversity from next-generation sequencing data.	Enabling reliable and comparable viral genomics and epidemiological studies and facilitating clinical diagnostics of viruses.