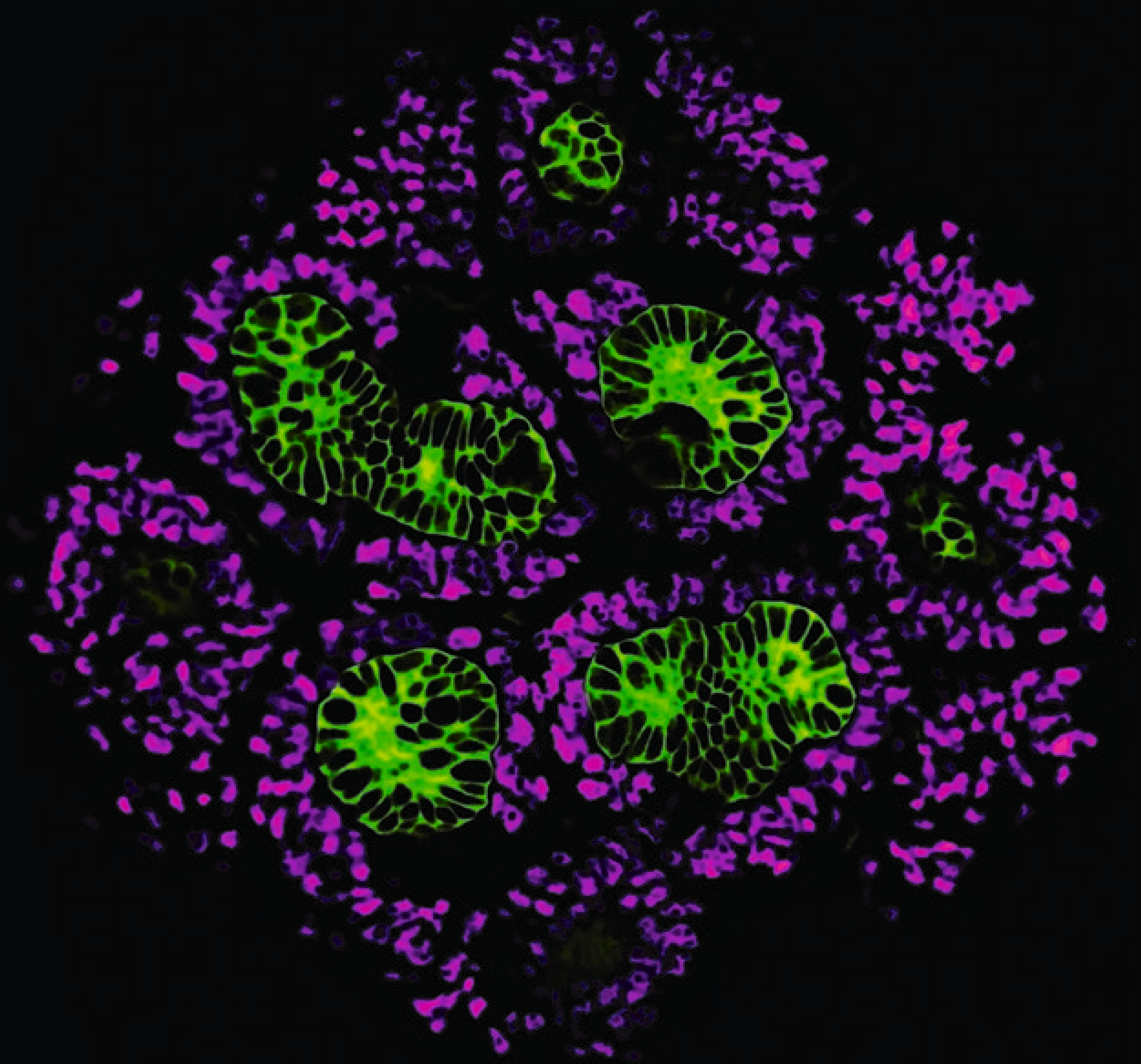


Swiss Institute of
Bioinformatics

SIB Profile 2019

Empowering advances in life sciences and health

WHY BIOINFORMATICS? | FOCUS ON THE INSTITUTION | ACTIVITIES 2018

COVER IMAGE
How do organs develop?
To better understand the mechanisms that shape organs and that enable cells to organize into functional tissues, the ETH's Computational Biology Group (CoBi - SEE P. 65) uses advanced microscopy techniques to build computational models of organogenesis. Such *in silico* models are paving the way for promising medical applications, such as tissue engineering and bioartificial organs. The image shown here is a section of an embryonic mouse kidney, obtained by light-sheet fluorescence microscopy. Cells that form the kidney's collecting system and cells that will differentiate into nephrons – the filtering units of the kidney – are labelled in green and purple respectively, thus enabling measurements of cell morphology during organ development.

SOURCE: HAROLD GÓMEZ, PHD STUDENT (COBI GROUP - ETH ZURICH). ALL RIGHTS RESERVED

Foreword

Felix Gutzwiller
President of the Foundation Council

Jérôme Wojcik
Chairman of the Board of Directors

Christine Durinx
Joint Executive Director

Ron Appel
Joint Executive Director

2018 was a hallmark year for Swiss Bioinformatics. In 20 years, the Institute has grown into a fully fledged organization, recognized nationally and internationally, with 70 research and service groups scattered across the country in its most distinguished schools of higher education. Once a case of creating a Swiss Bioinformatics Institute to secure funding and build a community, two decades on, SIB has definitively shed its adolescence to dive into adulthood. And the occasion was not overlooked: several celebratory events and projects honoured those who have shaped SIB over all these years.

> **"In 20 years, the Institute has grown into a fully fledged organization, recognized nationally and internationally"**

The most memorable event was the 20th anniversary evening in June, placed under the patronage of State Secretary Mauro Dell'Ambrogio. The celebration took place on the lakeside in Biel during SIB's two-day conference and was attended by hundreds of SIB Members as well as eminent guests. Several outreach projects were also launched, in particular a short offbeat institutional video, an animated history of SIB, a photography book and an educational mobile game. All these initiatives, fully financed by our sponsors, are still nurturing visibility in the press and have helped to consolidate our ties with our partners.

On another note, a few major changes occurred on the Institute's structural front – notably the co-directorship of its executive management in June. The Board of Directors also changed; Manuel Peitsch left to be succeeded by Jérôme Wojcik (undersigned), and Torsten Schwede stepped down – with Christophe Dessimoz succeeding him – to devote himself to his new role as Vice President for Research at the University of Basel. As for SIB's pioneer core facility and competence centre Vital-IT, Mark Ibberson, one of its Senior Scientists, has been appointed as new director.

In 2019, SIB will pursue its mission by providing the life science community with state-of-the-art bioinformatics resources, expertise, services and training. Dealing with major challenges raised by the development of personalized health, the Institute will continue to lay the foundations for a data exchange infrastructure within the context of the Swiss Personalized Health Network, while lending biocuration expertise to health-related projects.

We are grateful to our partner institutions who have helped SIB thrive, as we are to the State Secretariat for Education, Research and Innovation, the Federal Assembly, the Swiss National Science Foundation and those in funding roles. SIB, however, would not have reached its 20th anniversary without the expertise, work and dedication of its employees and members, to whom we express our sincere gratitude.

And now, let us dive headfirst into the next two decades!

Table of contents

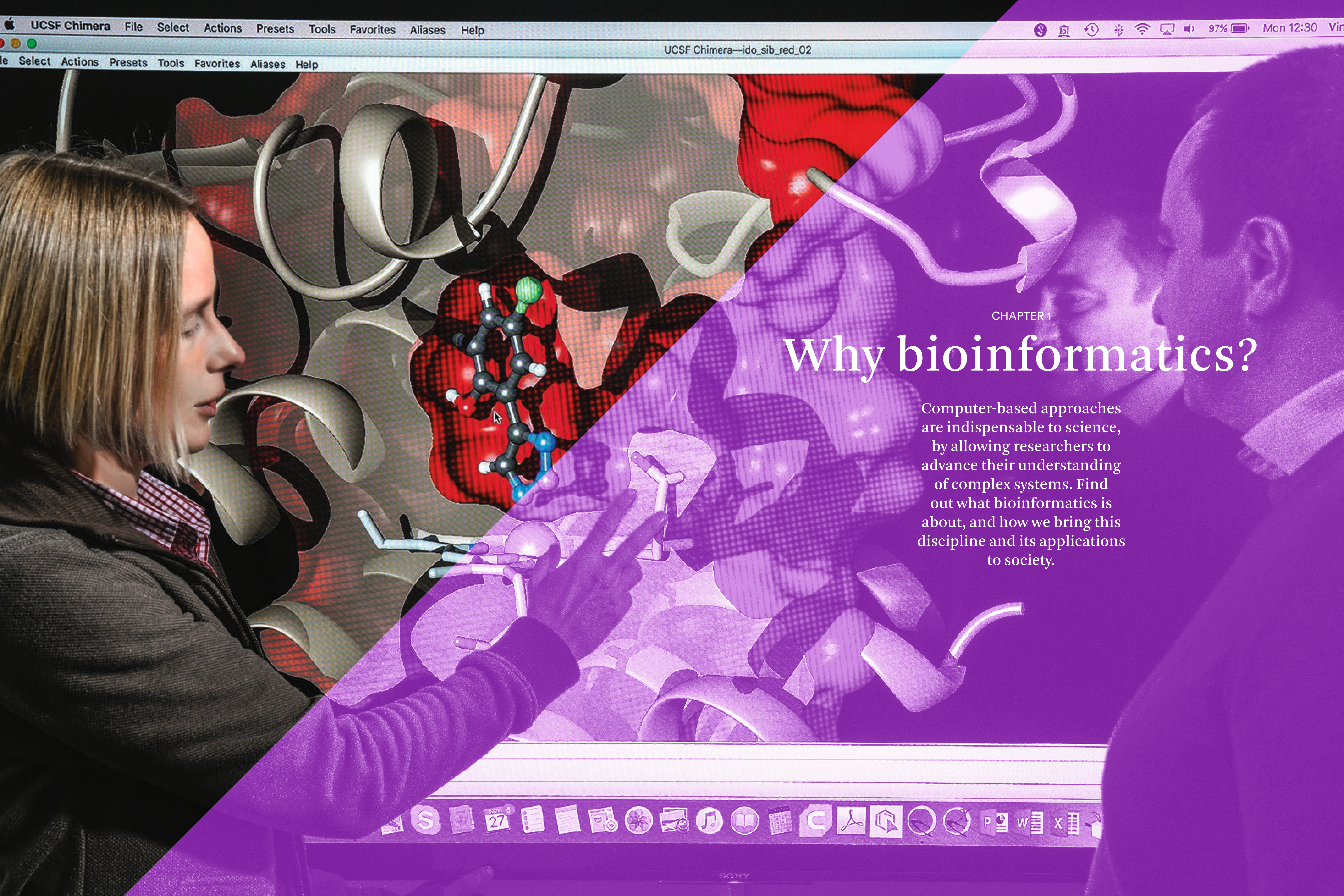
CHAPTER 1

CHAPTER 2

CHAPTER 3

A complete bioinformatics infrastructure

A strong bioinformatics community



CHAPTER 1

Why bioinformatics?

Computer-based approaches are indispensable to science, by allowing researchers to advance their understanding of complex systems. Find out what bioinformatics is about, and how we bring this discipline and its applications to society.

Bioinformatics: a definition

FIGHTING DISEASES. DISCOVERING SPECIES. UNDERSTANDING LIFE.

Bioinformatics is the application of computer technology to the understanding and effective use of biological and clinical data

Life scientists and clinicians have always tried to assemble data and evidence to find the right answers to fundamental questions. In 2019, data are no longer lacking. But a different kind of problem has emerged. Nowadays, new technologies are producing data at an unprecedented rate. Indeed, so much data – and of so many kinds – that they can no longer be interpreted by the human mind alone.

Enter bioinformatics.

Bioinformatics is the application of computer technology to the understanding and effective use of biological and clinical data. It is the discipline that stores, analyses and interprets the 'big data' generated by life science experiments, or collected in a clinical context, using computer science and dedicated data experts. This multidisciplinary field brings together scientists from a variety of backgrounds: biologists, computer scientists, mathematicians, statisticians and physicists.

Bioinformatics encompasses:

DATABASES for storing, retrieving and organizing information to maximize the value of biological data;

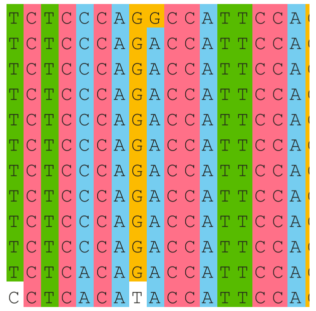
SOFTWARE TOOLS for modelling, visualizing, analysing, interpreting and comparing biological data;

COMPUTING AND STORAGE INFRASTRUCTURE to process and safeguard large amounts of data;

ANALYSIS of complex biological datasets or systems in the context of particular research projects;

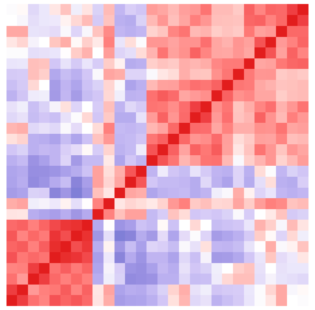
RESEARCH in a wide variety of biological fields using computer- and data science and leading to applications in diverse areas, from agriculture to precision medicine. (SEE P. 48)

What sort of data are we talking about?
Bioinformatics deals with a broad spectrum of complex data types.

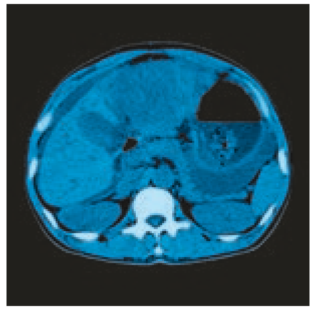


TCTCCCAGGCCATTCCA
TCTCCCAGACCATTCCA
TCTCCCAGACCATTCCA
TCTCCCAGACCATTCCA
TCTCCCAGACCATTCCA
TCTCCCAGACCATTCCA
TCTCCCAGACCATTCCA
TCTCCCAGACCATTCCA
TCTCCCAGACCATTCCA
TCTCCCAGACCATTCCA
TCTCACAGACCATTCCA
CCTCACATACCATTCCA

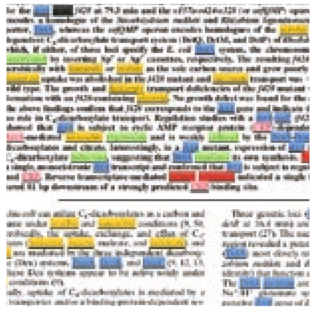
Sequence data from DNA, RNA or proteins



Expression data, such as the level of expression of a gene in a sample



Imaging data



Text data
And more...

Converting biological questions …

Massive amount of data and data types: genetics, text, biochemical, imaging, etc.

Hospitals and clinics

Research institutes

Private sector

International institutions

Life sciences and health actors

SIB
Swiss Institute
of Bioinformatics

Data management planning

Secure computing and storage infrastructure

Databases and software tools

Cutting-edge analyses and research

Dedicated multidisciplinary experts

… into answers

with various applications

Basic research

Medicine

Environmental sciences

Agriculture

Tailoring treatment to cancer patients

Understanding how chameleons change colour

Predicting the spread of bird flu

Bringing bioinformatics to society

Bioinformatics is increasingly tied to health, ethics, and societal issues. Through its outreach activities, SIB informs the public about the potential applications of the discipline.

> **"I now see how absurdly large the human genome truly is!"**
>
> **"Using the scientific tool, I was able to unleash my creativity and design a molecule with a potential use against cancer"**

Feedback from high school students who attended the Drug Design workshop

Explaining bioinformatics to the public matters greatly to SIB. Science fairs, activities within the classroom, a mobile game or hackathons… our team's imagination has no limits!

In 2018, the SIB Outreach Team met you at:

PLANÈTE SANTÉ LIVE
…the largest health fair in Switzerland: hundreds of visitors learned about the latest advances in clinical bioinformatics in the field of personalized health.
planetesante.ch/salon

NUIT DE LA SCIENCE
…two days and nights to discover the role of bioinformatics in drug design, personalized health and metagenomics.
ville-ge.ch/lanuitdelascience

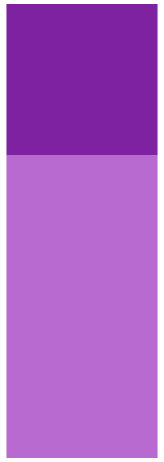
SIB BIOHACKATHON
…at the second edition of the Open Geneva Innovation Festival, children aged 9-14 used the SCRATCH programming language to tackle biological problems.
opengeneva.org

In 2019, meet the team at:

MYSTÈRES DE L'UNIL
University of Lausanne, 23-26 May
The University's open days will once again feature bioinformatics activities!
wp.unil.ch/mysteres

SCIENTIFICA
Zurich Science Days, ETH Zurich, 30 August – 1 September
What is the usefulness of bioinformatics in the identification of variants in oncology?
scientifica.ch

EXPANDING YOUR HORIZONS
University of Geneva, 16 November
Biennial event to encourage girls aged 11-14 to study science, technology, engineering and mathematics.
elargisteshorizons.ch/en



About
2,400
participants in our activities each year
including
1,600 students
(12-18 years old)

More events and news on Facebook, our dedicated outreach channel in French and English
goo.gl/4c6xCZ

BIOINFORMATICS IN THE CLASSROOM

SIB has designed several websites and workshops to introduce bioinformatics to schoolchildren, teachers and students:

CHROMOSOME WALK
An award-winning website to explore human chromosomes and the world of DNA, genes and proteins, with a new thematic journey on precision medicine coming up in 2019.
chromosomewalk.ch

DRUG DESIGN
A web-based workshop to acquire a simple yet realistic picture of how drugs are designed with the help of computers.
drug-design-workshop.ch

THE METAGENOMIC PIZZA
A workshop to grasp the concept of metagenomics, and identify all the species present in a pizza, based on their DNA.

TECDAYS
Since 2013, SIB has taken part in these Switzerland-wide events organized by the Swiss Academy of the Technical Sciences (SATW) for students aged 15-20.
satw.ch/en/tecdays

CHAPTER 2

Focus on the institution

Take a look back on the 20th anniversary celebrations of the SIB Swiss Institute of Bioinformatics and discover its key themes for the future. But first, take a backstage tour of the Institute’s organization, partners, people and finances.

SIB in brief

OUR MISSION AND OBJECTIVES

The mission of our Institute is to lead and coordinate the field of bioinformatics in Switzerland, in order to advance biological and medical research and to enhance health.

Infrastructure

SIB provides the national and international life science community with a state-of-the-art bioinformatics infrastructure, including resources, expertise and services.

DATABASES AND SOFTWARE TOOLS

SIB creates, maintains and disseminates worldwide a large portfolio of databases and software tools.

CORE FACILITIES AND COMPETENCE CENTRES

It offers easily accessible, in-depth expertise and support in bioinformatics through core facilities and competence centres.

PERSONALIZED HEALTH

It supports researchers and clinicians with know-how, resources and a secure data infrastructure dedicated to personalized health.

70
research and service groups

810
members, including

190
employees

20
institutional partners across Switzerland

13
core facilities and competence centres

Over
150
databases and software tools developed by our members and accessible via the ExPASy web portal

---

Over
2,450
peer-reviewed articles published since SIB's creation in 1998

Community

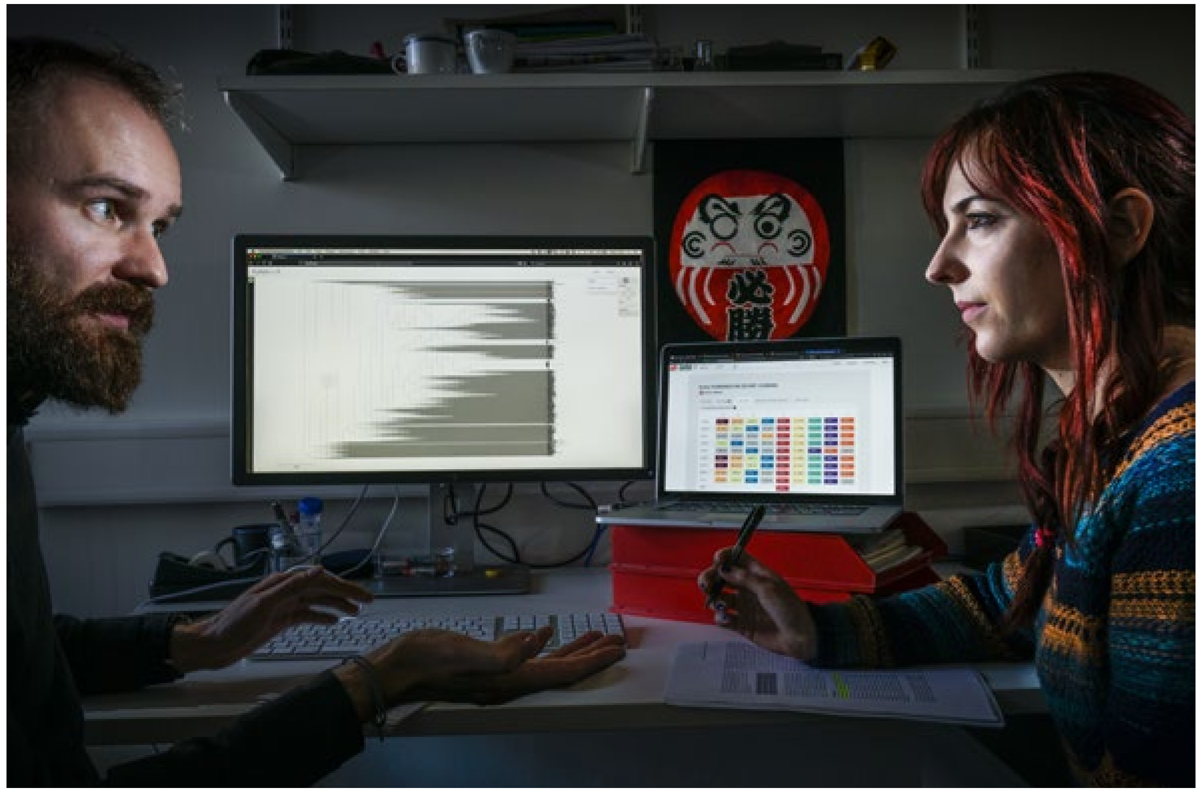
SIB brings together world-class researchers based in Switzerland and delivers training in bioinformatics.

SCIENTIFIC COLLABORATION

SIB represents and promotes Swiss bioinformatics nationally and internationally, by bringing together bioinformatics research and service groups from Swiss universities and research institutes. SIB fosters scientific collaboration and innovation at the highest level of excellence.

TRAINING IN BIOINFORMATICS

It provides life scientists and clinicians with a large portfolio of courses and workshops. It fosters exchanges among bioinformatics and computational biology PhD students, and trains them using the most up-to-date methods necessary for their research.



Top
Life scientists rely on various types of infrastructure to perform their research: knowledge-based resources, computing power, software, etc.

Bottom
Collaboration is at the heart of science. And it all starts with a strong community.

Organization and governance

Federating and leading such a pervasive domain as bioinformatics, even across a modestly sized country such as Switzerland, requires a unique structure with multiple bridges across institutions.

A dual management at the head of SIB

As of 22 June 2018, Christine Durinx, previously Associate Director, and Ron Appel have been leading SIB together as Joint Executive Directors. Christine is also a member of the Executive Board of the BioMedIT project, and co-chair of the ELIXIR Data Platform.

Becoming an SIB Group Leader

The main criterion for groups to join SIB is that they are producing new bioinformatics methods or software, or act as a core facility providing computing infrastructure or bio-informatics support. In addition, their Group Leader must have a professional appointment in one of SIB's 20 partner institutions, or be a senior scientist leading her/his own group.

As a non-profit foundation unifying bioinformatics across Switzerland and with 20 partner institutions (SEE OPPOSITE), SIB has a robust governance structure ensuring both its scientific independence and optimal internal functioning.

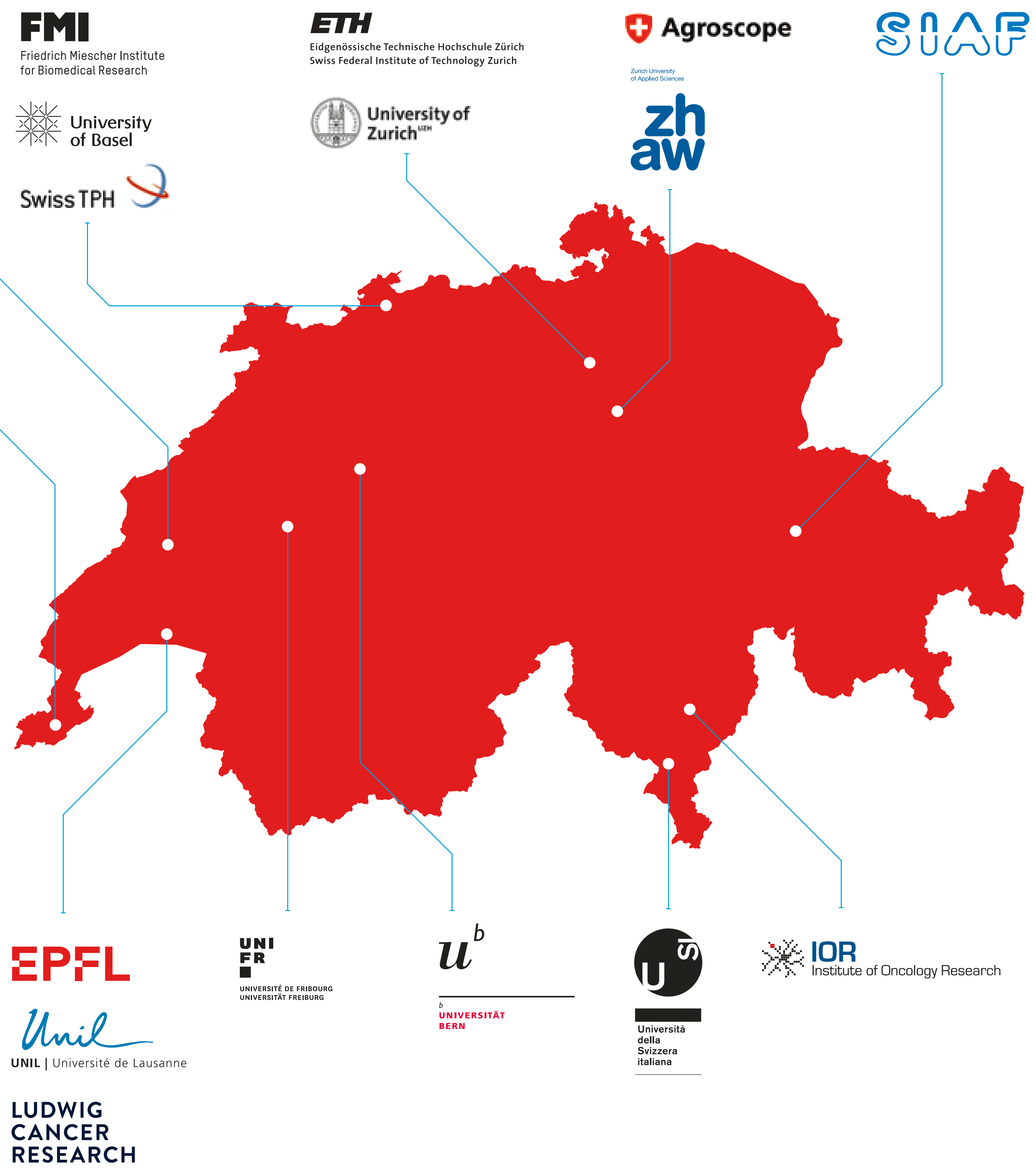
A Foundation Council, a Board of Directors, an external Scientific Advisory Board and a Council of Group Leaders complement the Management and support teams of the Institute.

SIB Internal Groups, which are entirely composed of SIB Employees, include: Clinical Bioinformatics (SEE P. 38), Core-IT (SEE P. 34), Digital Humanities+, Personalized Health Informatics (SEE P. 40), Swiss-Prot (SEE P. 56), Training (SEE P. 42) and Vital-IT (SEE P. 71).

Finally, SIB Affiliated Groups include academic groups from partner institutions, including those maintaining and developing an SIB-supported infrastructure, such as an SIB Resource or an SIB Core facility – they can thus include SIB Employees as well.

> **"The SIB Swiss Institute of Bioinformatics has considerable influence in Europe and successfully operates a national bioinformatics community with over 20 partner institutes."**
>
> Evaluation of SIB (Swiss Node of ELIXIR) by ELIXIR's Scientific Advisory Board, March 2019

PARTNER INSTITUTIONS



LIST OF GOVERNING BODIES

The Foundation Council

Each of SIB's partner institutions is represented on the Council.

President
Prof. Felix Gutzwiller
Former Senator

Founding Members
Prof. Ron Appel
Joint SIB Executive Director, University of Geneva

Prof. Amos Bairoch
Group Leader, SIB and University of Geneva

Dr Philipp Bucher
Group Leader, SIB and EPFL

Prof. Denis Hochstrasser
Former Vice-Rector, University of Geneva

Prof. C. Victor Jongeneel
Carl R. Woese Institute for Genomic Biology, University of Illinois, USA

Prof. Manuel Peitsch
Chief Scientific Officer Research at Philip Morris International

***Ex officio* Members**
Prof. Cezmi A. Akdis
Director, Swiss Institute of Allergy and Asthma Research (SIAF)

Mr Thomas Baenninger
Chief Financial Officer, Ludwig Institute for Cancer Research

Dr Claire Baribaud
Director, School of Business Administration (HEG-Geneva), HES-SO

Prof. Henri Bounameaux
Dean, Faculty of Medicine, University of Geneva

Prof. Edouard Bugnion
Vice-President for Information Systems, EPFL

Prof. François Bussy
Vice-Rector for Research, International Relations and Continuing Education, University of Lausanne

Prof. Daniel Candinas
Vice-Rector Research, University of Bern

Prof. Carlo Catapano
Director, IOR Institute of Oncology Research

Prof. Boas Erez
Rector, Università della Svizzera Italiana

Prof. Nicolas Fasel
Vice-Dean for Research and Innovation, Faculty of Biology and Medicine, University of Lausanne

Prof. Brigitte Galliot
Vice-Rector, University of Geneva

Prof. Susan Gasser
Director, Friedrich Miescher Institute for Biomedical Research (FMI)

Prof. Antoine Geissbühler
Head of eHealth and Telemedicine Division, Geneva University Hospitals (HUG)

Prof. Detlef Günther
Vice-President Research and Corporate Relations, ETH Zurich

Prof. Rolf Ingold
Vice-Rector for Research and Information Technology, University of Fribourg

Dr Corinne Jud
Head of the Competence Division Method Development and Analytics, Agroscope

Dr Caroline Kant
Executive Director, EspeRare Foundation Switzerland

Prof. Jérôme Lacour
Dean, Faculty of Science, University of Geneva

Prof. Jean-Marc Piveteau
President, Zurich University of Applied Sciences (ZHAW)

Prof. Alexandre Reymond
Director, Centre for Integrative Genomics, Faculty of Biology and Medicine, University of Lausanne

Prof. Vincent Peiris
Dean, School of Management and Engineering Vaud (HEIG-VD), HES-SO

Prof. Michael Schaepman
Vice-President for Veterinary Medicine and Natural Sciences, University of Zurich

Prof. Torsten Schwede
Vice President of Research and Talent Promotion, University of Basel

Prof. Juerg Utzinger
Director, Swiss Tropical and Public Health Institute

Co-opted Member
Dr Laurent Duret
CNRS Research Director, Laboratory of Biometry and Evolutionary Biology, Claude Bernard-Lyon 1 University, France

The Board of Directors (BoD)

The BoD consists of two Group Leaders elected jointly by the Council of Group Leaders and the BoD, two external members elected by the Foundation Council on the recommendation of the BoD, and the SIB Executive Directors. Members of the BoD are appointed for a renewable five-year period.

Prof. Ron Appel and
Dr Christine Durinx
Joint SIB Executive Directors

Ms Martine Brunschwig Graf
Former National Councillor

Prof. Christophe Dessimoz
Group Leader, SIB and University of Lausanne

Prof. Christian von Mering
Group Leader, SIB and University of Zurich

Dr Jérôme Wojcik (Chairman)
Senior Vice President, Translational Informatics and Biometrics, Europe, Precision for Medicine

The Scientific Advisory Board (SAB)

The SAB is made up of at least five members, who must be internationally renowned scientists from the Institute's fields of activity.

Prof. Søren Brunak
Founder of the Centre for Biological Sequence Analysis, Technical University of Denmark, Denmark

Dr Laurent Duret
CNRS Research Director, Laboratory of Biometry and Evolutionary Biology, Claude Bernard-Lyon 1 University, France

Prof. Melissa Haendel
Director of the Ontology Development Group, Oregon Health & Science University, Portland, USA

Prof. Alexey I. Nesvizhskii
Department of Pathology and Department of Computational Medicine & Bioinformatics, University of Michigan, Ann Arbor, USA

Prof. Christine Orengo
Department of Structural and Molecular Biology, University College London, United Kingdom

Prof. Ron Shamir
Computational Genomics Group at the Blavatnik School of Computer Science, Tel Aviv University, Israel

Prof. Alfonso Valencia (Chairman)
Life Sciences Department Director, Barcelona Supercomputing Centre, Spain

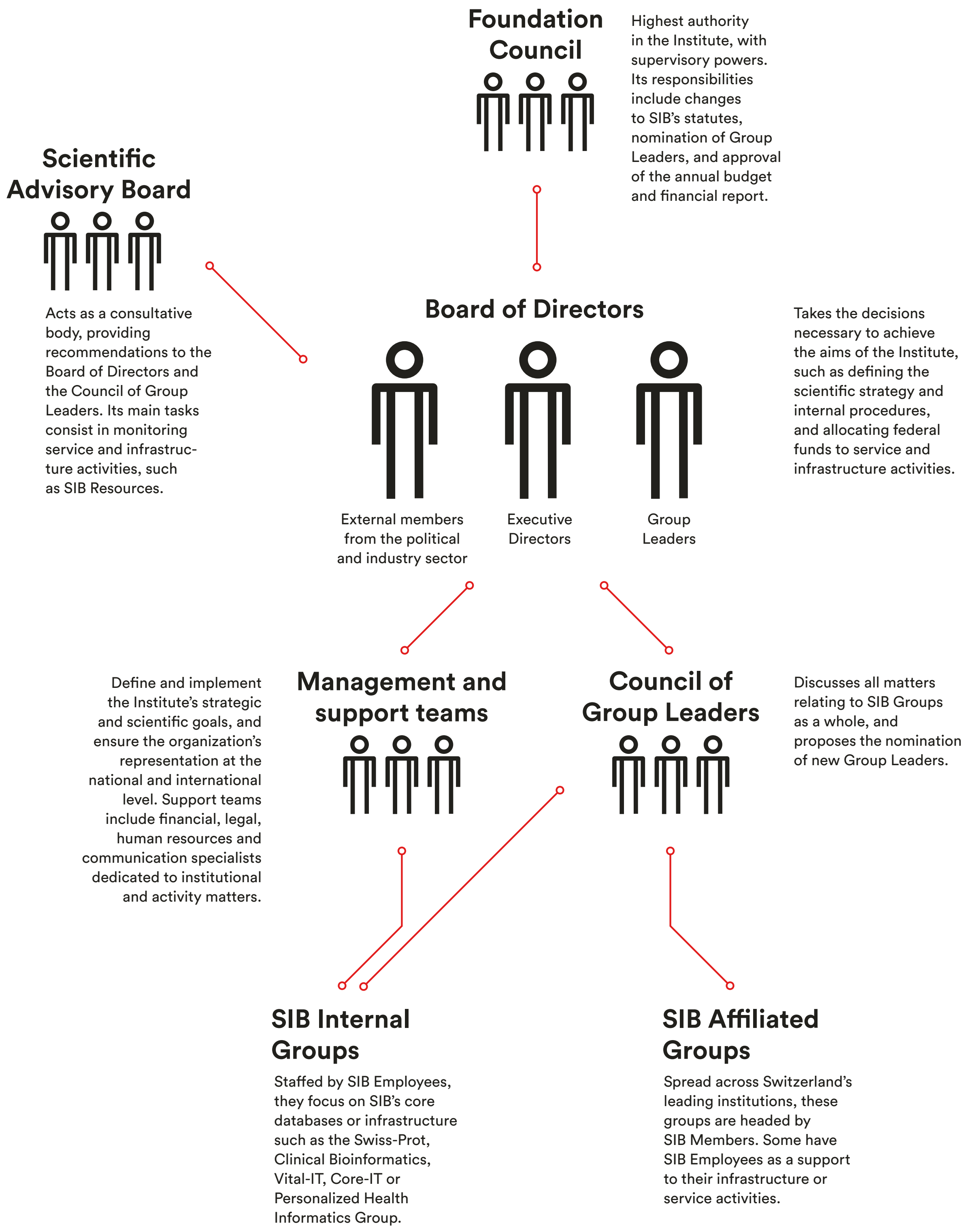
Council of Group Leaders

The Council consists of the Group Leaders and the SIB Executive Directors.

Honorary SIB Members

Prof. Ernest Feytmans
SIB Honorary Director

Dr Johannes R. Randegger
Former National Councillor, Honorary President of the SIB Foundation Council



Foundation Council

Highest authority in the Institute, with supervisory powers. Its responsibilities include changes to SIB's statutes, nomination of Group Leaders, and approval of the annual budget and financial report.

Scientific Advisory Board

Acts as a consultative body, providing recommendations to the Board of Directors and the Council of Group Leaders. Its main tasks consist in monitoring service and infrastructure activities, such as SIB Resources.

Board of Directors

Takes the decisions necessary to achieve the aims of the Institute, such as defining the scientific strategy and internal procedures, and allocating federal funds to service and infrastructure activities.

Management and support teams

Define and implement the Institute's strategic and scientific goals, and ensure the organization's representation at the national and international level. Support teams include financial, legal, human resources and communication specialists dedicated to institutional and activity matters.

Council of Group Leaders

Discusses all matters relating to SIB Groups as a whole, and proposes the nomination of new Group Leaders.

SIB Internal Groups

Staffed by SIB Employees, they focus on SIB's core databases or infrastructure such as the Swiss-Prot, Clinical Bioinformatics, Vital-IT, Core-IT or Personalized Health Informatics Group.

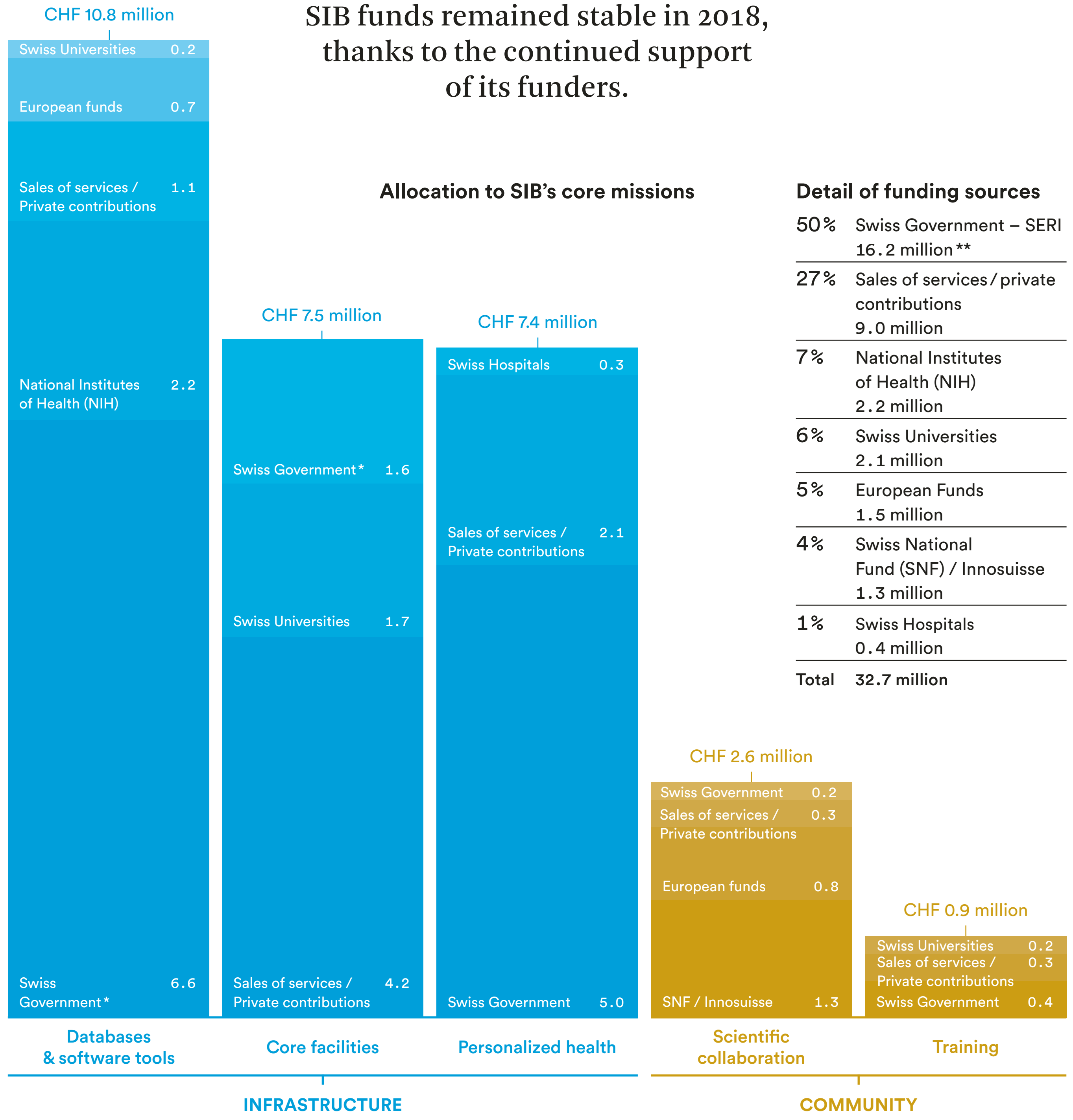
SIB Affiliated Groups

Spread across Switzerland's leading institutions, these groups are headed by SIB Members. Some have SIB Employees as a support to their infrastructure or service activities.

Key figures

SIB funds remained stable in 2018, thanks to the continued support of its funders.

Allocation to SIB's core missions

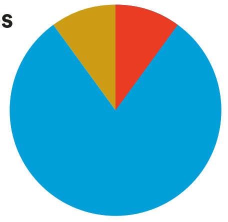


Detail of funding sources

50 %	Swiss Government – SERI 16.2 million **
27 %	Sales of services / private contributions 9.0 million
7 %	National Institutes of Health (NIH) 2.2 million
6 %	Swiss Universities 2.1 million
5 %	European Funds 1.5 million
4 %	Swiss National Fund (SNF) / Innosuisse 1.3 million
1 %	Swiss Hospitals 0.4 million
Total	**32.7 million**

Allocation by activities

- 🟦 78 % Infrastructure
- 🟨 11 % Community
- 🟥 11 % Management & support teams ***



71%

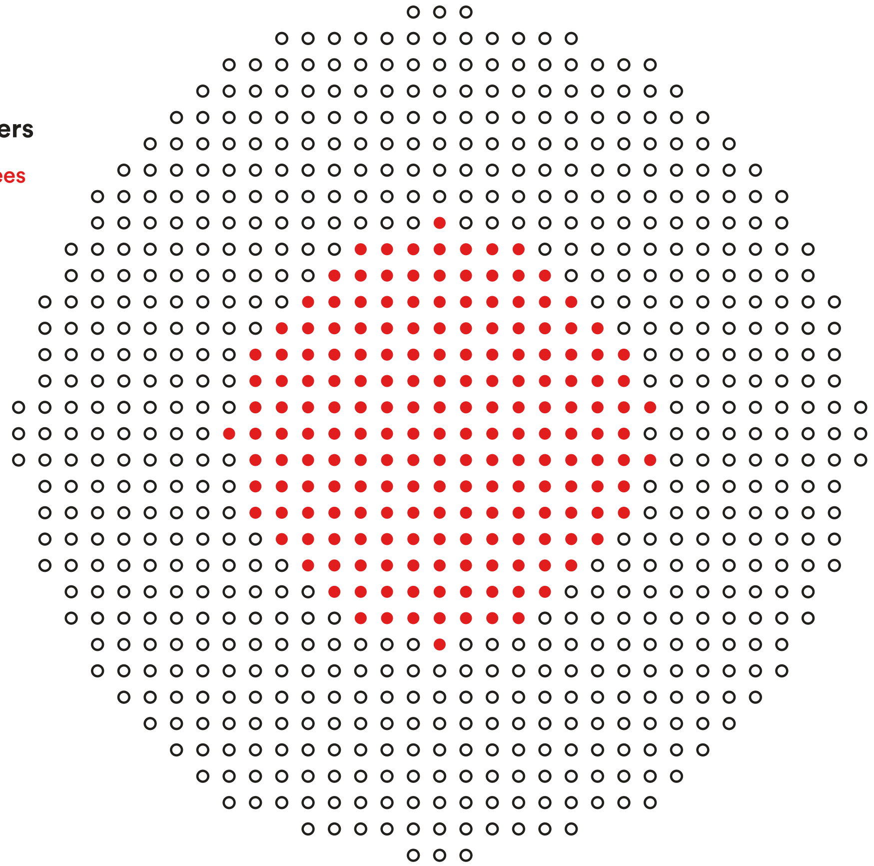
of financial resources are allocated to the payment of salaries, reflecting the expertise-driven activities of the Institute.

* Swiss Government funds are allocated to SIB Resources and Core facilities as per the recommendations of SIB's external Scientific Advisory Board (SEE P. 19) ** This amount includes a special financing of CHF 4.5 million dedicated to the BioMedIT project (SEE P. 40) reflected in the Personalized Health pillar. *** SIB's Management and support teams (SEE P. 19) are financed by the Swiss Government as well as overheads on external funds. The groups having entrusted SIB with the management of their funds benefit from a specific support in legal, human resources and financial monitoring domains.
As per 2018 audited figures

SIB MEMBERS*

SIB has 810 Members

including 190 Employees from 23 different nationalities



They are based in:

Basel	182
Bellinzona	10
Bern	31
Davos	2
Fribourg	11
Geneva	145
Lausanne	286
Lugano	4
Wädenswil	19
Yverdon	10
Zurich	110

* As of 31 December 2018

Perspectives

Where will SIB be adding value in the years to come? Here are some of the Institute's major development themes for the near future, resulting from the 2018 strategic roadmap.

Strengthening capacities in data security and developments for the health sector

Sensitive data in bioinformatics resources or across hospitals and clinics require the highest levels of security. SIB is further developing its abilities in this context.

SECURING HEALTH DATA INFRASTRUCTURE

In the context of the Swiss Personalized Health Network (SPHN), SIB plays a central role as Data Coordination Centre and was mandated by the Swiss Confederation to implement a secure network, enabling data to be shared for research in personalized health. The first steps in this secure data architecture are being implemented through the BioMedIT project (SEE P. 40), thanks to the joint work of SIB Employees and affiliated core facilities known as BioMedIT nodes. The nodes provide access to a remote data analysis environment, thereby enabling scientists to analyse data securely.

Together with strong analytical capacities and comprehensive reference databases, the demand for high security and reliable storage of sensitive data is continuously increasing. These are all part of SIB's competence portfolio.

SAFEGUARDING DATA RESOURCES

For SIB Resources with sensitive data, penetration testing is conducted on a regular basis and guidelines on how to resolve security vulnerabilities have been established. Further work is underway to ensure security policies are adapted to the sensitivity of databases and extended to SIB Resources outside the health field. This is all part of SIB's mission to add value to bioinformatics resources dedicated to the scientific community.

A TRUSTED PARTNER FOR HEALTH ACTORS

SIB has become a partner of choice for health professionals, both for its strategic position between research and clinical needs, and for its expertise related to data security.

Some obvious examples are the many projects of the Clinical Bioinformatics Group (SEE P. 39). A few years ago, the Institute launched this activity to bring bioinformatics advances directly to clinicians and their patients.

Today, the Group is recognized for its ability to partner with Swiss hospitals in projects supporting clinical diagnosis, unite them around common guidelines on interpretation practices for next-generation sequencing and collaborate on new data resources, such as the one for interpreting genetic variants (SVIP, SEE P. 23). All this is the result of a strong collaborative spirit with scientists and clinicians, the unique expertise of SIB's Developers and SIB's capacity to ensure the highest standards of security for the data involved.

Leveraging biocuration expertise for health projects

Be it at the genomic level or increasingly in other omics areas, SIB is leveraging its biocuration expertise in the context of personalized health.

BIOCURATION, A KEY BUILDING BLOCK OF PERSONALIZED HEALTH

In a tumour sample, up to several hundred variants can be detected. But how can one filter out those that are particularly pathogenic, or code for an accrued sensitivity – or resistance – to a specific treatment?

With the advent of molecular medicine, clinicians from around the world are turning to carefully annotated knowledgebases to assess the pathogenicity of a specific variant, guide their diagnosis, or pinpoint associated drugs.

In this context, the role of biocurators – those expert scientists who collect, annotate and validate the most relevant and valuable information from the growing biomedical literature and enter it into knowledgebases – will take on increasing importance.

Biocurators will form an essential building block of pluridisciplinary personalized medicine programs in Switzerland and internationally.

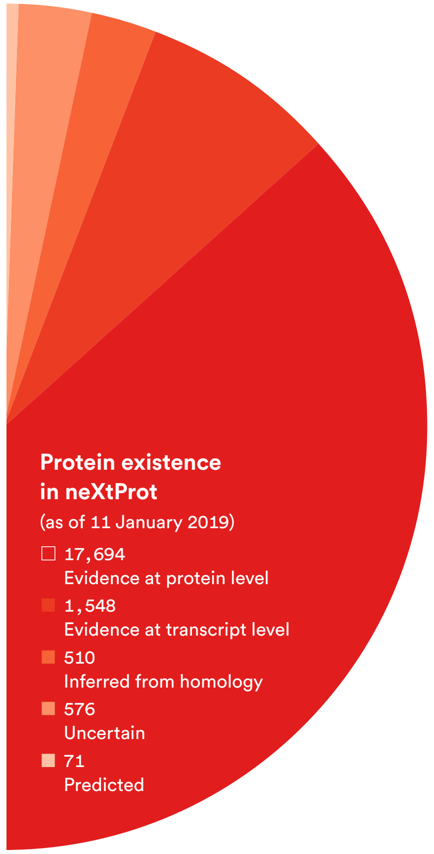
S V I P

A CURATED PLATFORM TO HARMONIZE VARIANT INTERPRETATION

A concrete example is the SVIP project that has recently started under the co-leadership of the SIB Clinical Bioinformatics Group. Its purpose is to harmonize the clinical interpretation of variants across Swiss hospitals.

Initially targeting variants associated with cancer, the platform will combine clinical data from partner hospitals with publicly available information from international projects (e.g. ClinVar, CIViC) and published literature, sourced by expert biocurators from SIB Groups such as Swiss-Prot (SEE P. 56).

svip.ch



FROM GENOMIC VARIANTS TO PROTEINS – AND BEYOND

The human protein knowledgebase neXtProt has also been working on leveraging its information towards the clinical world: initially designed to help researchers find out what human proteins do in our bodies (SEE ABOVE), the Group developing the resource, CALIPHO (SEE P. 56), has been generating a corpus of annotations on the impact of protein sequence variations in cancers and genetic diseases. These so-called 'phenotypic' effects encompass changes in the activity of the protein, its stability and its interactions with other proteins or drugs.

While most information currently used by clinicians is at the genomic level, the rise and integration of other 'omics' such as proteomics, lipidomics or transcriptomics means that expert-curated resources developed by SIB Groups will be crucial to advance our understanding of diseases, improve biomarker discovery or support the development of personalized medicine programmes.

Increasing the impact of our bioinformatics resources

What does it imply to increase the impact of a bioinformatics resource when its scientific relevance is already established? Since February 2018, a dedicated SIB Team has been working hand-in-hand with resource developers to create a best-practice toolkit that will help them to grow their user base and share their know-how with the community.

To conduct their daily work, life scientists and clinicians have to wade through an ever-thickening swamp of tools, software and databases. At the same time, the next generation of researchers – both tech-savvy and ardent app-consumers – is raising the bar of expectations in terms of resource usability. “Scientific excellence is no longer the sole criterion for a resource to be competitive in the long term,” says Chiara Gabella, SIB Scientific Coordinator and User eXperience (UX) Architect.

To keep pace with this highly competitive context, SIB – whose mission includes providing the life science community with state-of-the-art resources, as well as ensuring their long-term sustainability – launched an initiative to further increase the impact of its resources.

Indeed, beyond an established reputation for scientific excellence, much can still be done to ensure that the user community is aware of these tools, and that they are used to their full potential while generating a positive experience. Led by Chiara Gabella and Séverine Duvaud, SIB Software Developer and UX Designer, the initiative aims to develop a best-practice toolkit available to SIB Resources’ teams, as well as to the wider community of SIB Developers.

In addition to improving resource design by taking into account the experience of the end user, this toolkit will also include instruments for the in-depth interpretation of analytics used to monitor and improve a resource. Among these: analysing the user base, finding out how much time is necessary for a user to find relevant information, or what is the most useful type of information on the platform. Cross-referencing between resources involved in the same analytical workflow will also be part of this strategic review. “Ultimately, such a toolkit will help target development efforts and increase the return rate of users,” says Duvaud.

To achieve its goal, the team is relying on a wide set of interactive approaches that have been successful in similar contexts, such as workshops and live user testing at conferences. It also released several community-based initiatives, including an internal competency network and a specialized blog dedicated to resource developers.

What is User eXperience (UX) design in a life science context?

UX encompasses the end user’s interaction with a product, tool, or system. It involves deeply understanding users through research, organizing information, visual design and more, all with the goal of meeting user needs. This is especially important in the life sciences, where users are scientists in a complex environment and where more usable, efficient tools can accelerate innovation.

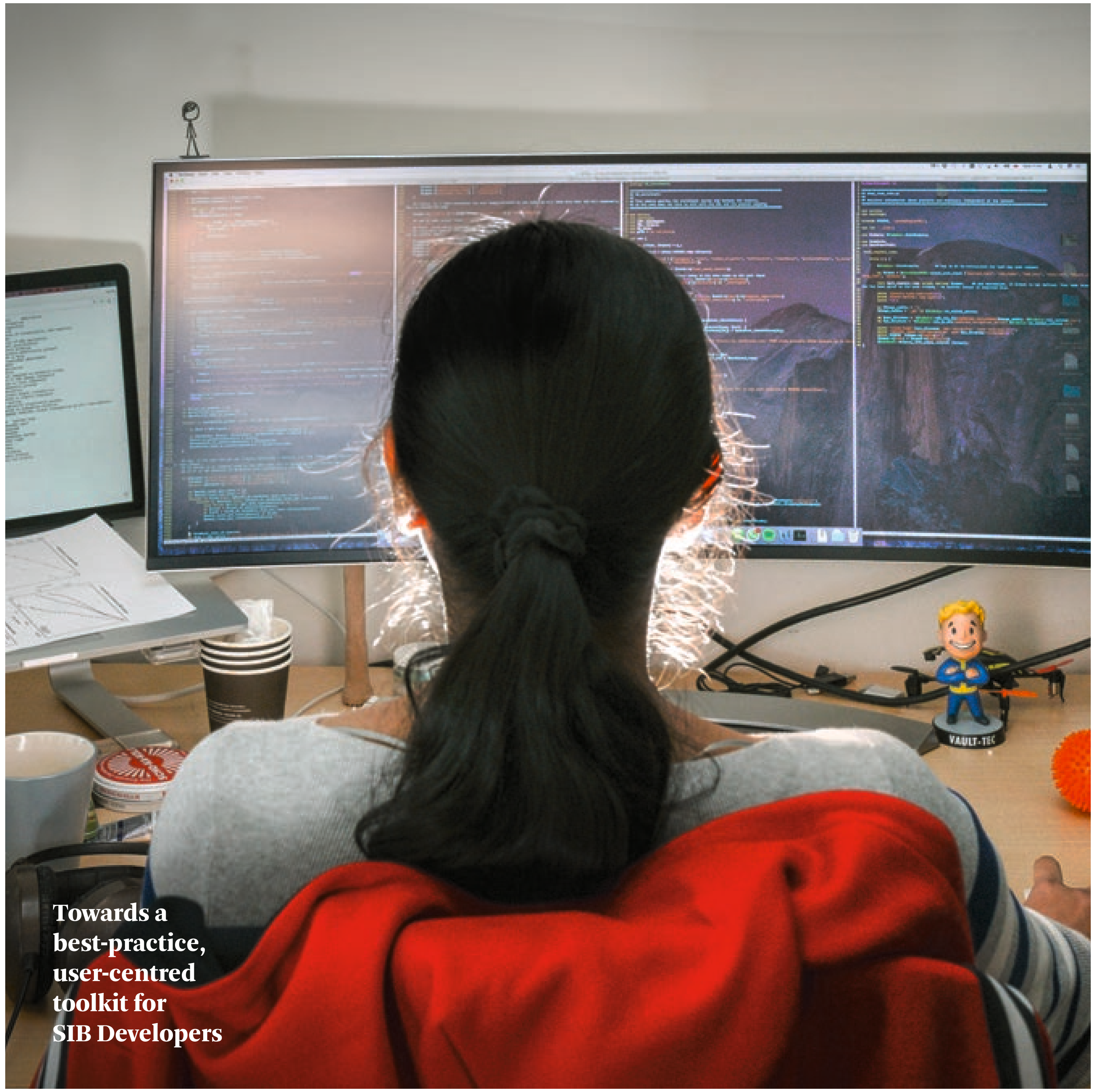
Source: Pistoia Alliance

> “Scientific excellence is no longer the sole criterion for a resource to be competitive in the long term”

CHIARA GABELLA
SIB Scientific Coordinator
and UX Architect

> “Ultimately, such a toolkit will help target development efforts and increase the return rate of users”

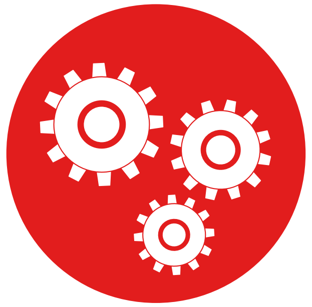
SÉVERINE DUVAUD
SIB Software Developer
and UX Designer



Towards a best-practice, user-centred toolkit for SIB Developers

SIB's 20th anniversary in images

Movies, events, a mobile game: this special year for bioinformatics was well celebrated! And these initiatives still continue to generate visibility and public interest for the field and its experts.

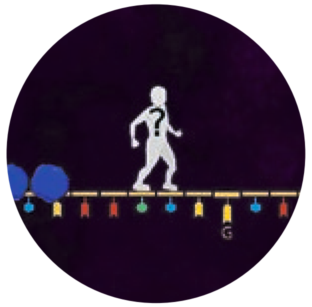


A BRIEF HISTORY OF SIB
An animated movie to retrace the steps of the Institute and its evolution over time.
goo.gl/yfM4UJ

SIB DAYS SPECIAL EVENING
Hundreds of SIB Members and eminent guests gathered on the shores of Lake Biel, Switzerland, for our biannual conference. The special evening was placed under the patronage of the SERI (State Secretariat for Education, Research and Innovation) State Secretary.

SIB, THE MOVIE – SWISS BIOINFORMATICS IN ACTION
This short offbeat clip has enthused a broad audience, from scientists to the general public: if you haven't already, meet Sophie the bioinformatician, Jürgen the software developer and Fred the biocurator.
goo.gl/RdckA2

SEQUENCING SIB'S BIRTHDAY CAKE
Hundreds of visitors at the Open Days of the University of Lausanne and at the Nuit de la Science used metagenomic approaches to analyse the ingredients of the SIB birthday cake, designed for the occasion.

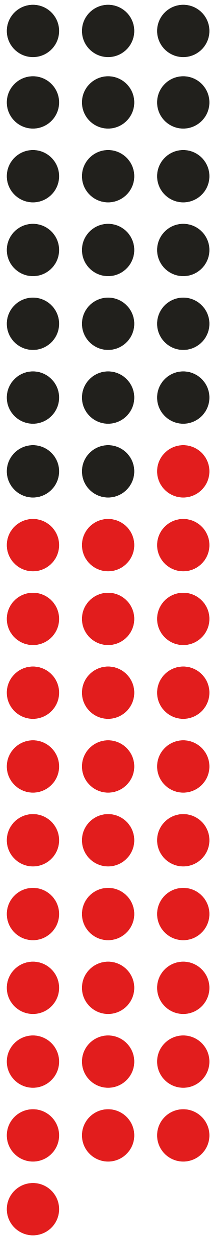


GENOME JUMPER, A MOBILE GAME TO EXPLORE HUMAN DIVERSITY
Why is it some people don't have fingerprints? Or some are more at risk of gluten intolerance? And what does bioinformatics have to do with it? Over 2,500 downloads already for this educational game:
genome-jumper.sib.swiss

SCIENCE FICTION BOOK
Combining the portraits of our Group Leaders and an overview of significant advances made possible by bioinformatics, this book is an aesthetic and fantasy-like journey to the land of biological data science.

49
articles in the press in 2018

including
29
linked to SIB's anniversary



SIB's anniversary projects and events were fully financed thanks to our sponsors and partners

Gold sponsors
Academic work
Leenaards Foundation

Silver sponsors
EPFL – Blue Brain Project
ETH Zurich
Hasler Stiftung
ISREC Foundation
Swiss Academy of Medical Sciences (SAMW/ASSM)
Swisscom
SWITCH
University of Lausanne
University of Zurich

Bronze sponsors
BioAlps
Geneva School of Business Administration (HEG)
Institute of Oncology Research (IOR)
Lonza
Precision for medicine
Publex
Sophia Genetics
Swiss Integrative Center for Human Health (SICHH)
Università della Svizzera italiana
University of Geneva – Faculty of Medicine
University of Geneva – Faculty of Sciences
Zurich University of Applied Sciences (ZHAW)

Find all our 20 years projects on sib.swiss/20years and on Twitter #SIB20years

CHAPTER 3

Activities 2018

Discover our latest developments and achievements on the bioinformatics infrastructure and community fronts.

Infrastructure

I.
DATABASES AND SOFTWARE TOOLS

...a large range of bioinformatics resources

II.
CORE FACILITIES AND COMPETENCE CENTRES

...easily accessible, in-depth expertise and support in bioinformatics

III.
PERSONALIZED HEALTH

...dedicated know-how, resources and secure data infrastructure for researchers and clinicians

Community

IV.
TRAINING IN BIOINFORMATICS

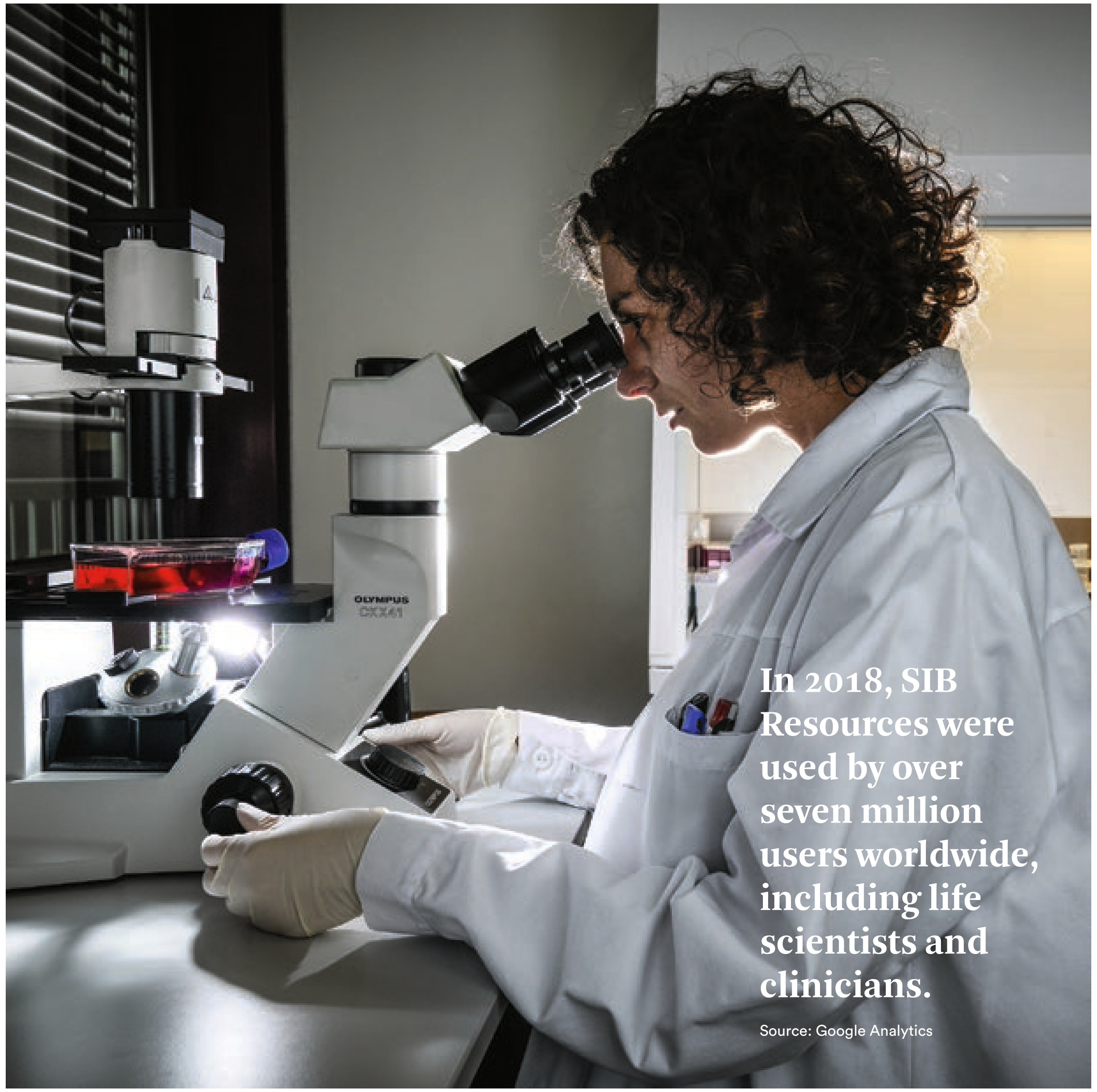
...a diverse portfolio of courses and workshops for life scientists and clinicians

V.
SCIENTIFIC COLLABORATION

...a unique ecosystem for bioinformatics groups from Swiss universities and research institutes

A complete bioinformatics infrastructure

From research tools and computing power to data science expertise and dedicated infrastructure for personalized health, SIB's portfolio addresses the bioinformatics needs of life scientists and clinicians, in academia or the private sector.



In 2018, SIB Resources were used by over seven million users worldwide, including life scientists and clinicians.

Source: Google Analytics

I.

DATABASES AND SOFTWARE TOOLS

Among the 150 databases and software tools developed and maintained by SIB Groups, the so-called SIB Resources are deemed of particular importance to the life-science community, and thus benefit from the Institute's specific support.

Life scientists often rely on dozens of databases or software tools to reconstruct DNA sequences, interpret the function of genes, or predict the 3D structure of proteins. It is therefore crucial that key resources are maintained in the long term while taking into account the fast-changing needs of researchers. It is SIB's mission to support such goals, and it has set up a process to identify and support such key resources.

Every four years, the Institute's external Scientific Advisory Board (SAB) is in charge of providing recommendations on the portfolio of SIB-supported Resources. By taking into account competitive criteria – such as the impact and scientific return on investment of candidate resources – together with SIB's global portfolio management and strategic orientation, the SAB identifies best-in-class resources, and provides guidelines for their continuous development.

Based on the these recommendations, the Board of Directors decides the allocation of funding, and SIB Management then assists the groups developing the resources in implementing their action plan and exchanging best practices. For instance, since 2018 a dedicated team has assisted developers in increasing the visibility and impact of their resource (SEE P. 24).

Swiss-Prot: a world flagship of biocuration

Thirty-three years after its creation and with some 50 employees, SIB's Swiss-Prot Group continues to be internationally recognized for its expertise in biocuration, ranging from annotating protein sequences, their functions, and their sequence variants (e.g. UniProtKB/Swiss-Prot, ViralZone, HAMAP, PROSITE), to metabolites and metabolic reactions – also featuring prominently in many human diseases – with SwissLipids.

ExPASy, SIB's bioinformatics resource portal, turned 25 years old in 2018: to mark the occasion, SIB initiated its in-depth redesign to better meet the community's recent needs and add value to the scientific resources it annotates and connects. The launch of ExPASy 3.0 is expected in 2019.

SWISS-MODEL celebrates its 25th anniversary

Once the first automated homology modelling service on the internet, SWISS-MODEL today processes over one million model requests and generates about 2,000 literature citations each year. It is one of the most widely used structure modelling servers worldwide. Although it is 25 years old, it has lost nothing of its original philosophy: to make protein modelling accessible to non-experts, thanks to an intuitive and interactive interface. It celebrated its anniversary with an international symposium in Basel, held in October 2018, which explored the state of the art of computational structural biology and its latest exciting developments, with new players such as Google DeepMind entering the field.

Sustainability of knowledgebases

In 2018, SIB pursued its efforts to find solutions to the long-term sustainability of knowledgebases with an F1000 Research blog on the study led by SIB and supported by ELIXIR on that topic. SIB's Chiara Gabella, first author on the study, and reviewer Helen Berman, Director Emerita and Associate Director of the RCSB Protein Data Bank (Rutgers, The State University of New Jersey) discussed funding models, open access data and the future of biocuration.

Proteins & proteomes

UniProtKB/Swiss-Prot SEE P. 56

Protein knowledgebase

TYPE Knowledgebase with manual expert curation
DESCRIPTION Hundreds of thousands of protein descriptions, including function, domain structure, subcellular location, post-translational modifications and functionally characterized variants.
HIGHLIGHT Expert-curated part of UniProt, the most widely used protein information resource in the world, with over six million page views per month. Recognized as an ELIXIR Core Data Resource.
NEW IN 2018 Includes the resource Rhea, a comprehensive expert-curated knowledgebase of biochemical reactions, to annotate and represent enzyme-catalysed reactions.

SWISS-MODEL SEE P. 64

Protein structure homology-modelling

TYPE Software tools and repository
DESCRIPTION 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.
HIGHLIGHT Easy-to-use web-based platform processing over one million model requests per year, providing model information for experts and non-specialists.
NEW IN 2018 Celebrated its 25th birthday with a scientific symposium in Basel.

STRING SEE P. 57

Protein-protein interactions

TYPE Knowledgebase
DESCRIPTION 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.
HIGHLIGHT Over 25,900,000 protein interactions at the highest confidence level. Recognized as an ELIXIR Core Data Resource.
NEW IN 2018 Deemed overall best-performing network resource according to a Californian study (Huang J K et al. Cell Systems 2018)
COMING UP Next major release (version 11.0), doubling the organism coverage. New features for improved detection of functional enrichments in user-provided inputs.

Glycomics@ExPASy SEE P.57

Glyco(proteo)mics resource portal

TYPE Web tools and knowledgebases with manual expert curation
DESCRIPTION Extensive set of resources on sugar structures (UniCarb-DB), their relationships with the proteins that carry them (GlyConnect) and the proteins that bind them, with a special focus on interactions between pathogen proteins and mammalian glycans (SugarBind).
HIGHLIGHT Unique portal for glyco(proteo)mics data analysis and interpretation.
NEW IN 2018 Release of UniLectin, a collaborative platform integrating information on non-enzymatic carbohydrate-binding proteins, aka lectins.

neXtProt SEE P. 56

Human protein knowledgebase

TYPE Knowledgebase with manual expert curation
DESCRIPTION Information on human protein biology such as function, involvement in diseases, mRNA/protein expression, protein-protein interactions, post-translational modifications, protein variations and their phenotypic effects.
HIGHLIGHT Advanced search options combining data at genomic, transcriptomic, proteomic and phenotypic levels. Tools specifically designed for the proteomics community.
NEW IN 2018 The neXtProt Cancer Variants portal contains over 6,300 annotated observations at the molecular and/or cellular level for BRCA1 variants.

Lipids

SwissLipids SEE P. 56

A knowledge resource for lipids

TYPE Knowledgebase
DESCRIPTION Information about known lipids, including knowledge of lipid structures, metabolism, interactions, and subcellular and tissular localization, providing a framework for the integration of lipid and lipidomic data with biological knowledge and models.
HIGHLIGHT Contains information on over 550,000 lipid structures from over 300 major lipid classes.
NEW IN 2018 Structures for 150 new lipid classes.
COMING UP Coverage of chemical space expected to double with structures for 300 new lipid classes, including many complex glycosphingolipids.

Genes and genomes

Bgee SEE P. 61

Gene expression expertise

TYPE Knowledgebase with manual expert curation and software tool
DESCRIPTION Gene expression data (including all types of transcriptomes), allowing retrieval and comparison of expression patterns between animals, human, model organisms and diverse species of evolutionary or agronomical relevance.
HIGHLIGHT Only resource to provide homologous gene expression between species.
NEW IN 2018 Publication of the R package BgeeDB for data access and TopAnat analyses.

EPD SEE P. 50

Eukaryotic Promoter Database

TYPE Knowledgebase with manual expert curation and software tools
DESCRIPTION Quality-controlled information on experimentally defined promoters of higher organisms, as well as web-based tools for promoter analysis.
HIGHLIGHT Over 180,000 promoters downloadable, analysable over a web interface and viewable in the UCSC genome browser.
NEW IN 2018 Promoter collections for four new model organisms: rhesus macaque, rat, dog and chicken.
COMING UP Extension of EPD to the first human parasite genome: *Plasmodium falciparum*.

SwissOrthology

Under this umbrella, SIB provides leading resources around orthology, i.e. corresponding genes across different species. In particular, it includes the two leading resources – OMA and OrthoDB – and two quality assessment tools – BUSCO for genome completeness and the Quest for Orthologs benchmark service. A SwissOrthology web portal is currently being developed and should go live in 2019.

OrthoDB SEE P. 55

The hierarchical catalogue of orthologs

TYPE Automated phylogenomic database and software tool
DESCRIPTION Comprehensive catalogue of orthologs with functional and evolutionary annotations.
HIGHLIGHT Largest orthology resource, enabling users to generate publication-quality comparative genomics charts, as well as to upload and analyse their own data.
COMING UP The OrthoDB v.10 update will feature the sampling of 1,271 eukaryotes, 6,013 prokaryotes and 6,488 viruses.

OMA SEE P. 58

Orthology MAtrix

TYPE Automated phylogenomic database and software tool
DESCRIPTION High-quality orthology predictions among complete genomes.
HIGHLIGHT Broad scope and size, feature-rich web interface, availability in a wide range of formats and interfaces, frequent update schedule.
NEW IN 2018 Bioconductor/R and Python APIs
COMING UP A redesign of the OMA browser.

SwissRegulon Portal SEE P. 54

Tools and data for regulatory genomics

TYPE Software tools and knowledgebases
DESCRIPTION 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.
HIGHLIGHT User can upload raw microarray, RNA-seq or ChIP-seq data to automatically infer the core regulatory networks acting in the system of interest.

V-Pipe SEE P. 58

Viral genomics pipeline

TYPE Software tool
DESCRIPTION Pipeline integrating various open-source software packages for assessing viral genetic diversity from next-generation sequencing data.
HIGHLIGHT Enabling reliable and comparable viral genomics and epidemiological studies and facilitating clinical diagnostics of viruses.
NEW IN 2018 A modular and extensible architecture that enables testing of new tools and supports establishing best practices for virus research and clinical diagnostics.

Structural biology

SwissDrugDesign SEE P.63

Drug design

TYPE Software tools
DESCRIPTION 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).
HIGHLIGHT Comprehensive and integrated web-based drug design environment.
NEW IN 2018 Passed the one million jobs mark.

II.

CORE FACILITIES AND COMPETENCE CENTRES

Swiss life scientists with computational needs can count on a tight mesh of expert hubs in Switzerland, offering high-performance computing, software development, bioinformatics expertise and specialized training – or all these at once.

SIB counts 13 core facilities and competence centres, distributed in major Swiss academic institutions.

Composed of teams of scientists and technical staff, they offer multidisciplinary expert services to researchers and clinicians from academia and the Swiss private sector, and are often key partners in international projects. The collaborations they engage in cover a wide array of domains and disciplines, from environmental sciences to medicine. Among them, SIB is strongly involved in four: two SIB Internal Groups (Core-IT and Vital-IT) and two SIB Affiliated Groups (BCF and sciCORE), which are respectively either entirely or partly composed of SIB Employees.

Core-IT: a new SIB Internal Group

Created in 2019 and headed by Heinz Stockinger, this Internal Group, based in Lausanne and Geneva, is in charge of coordinating, operating and maintaining information technology infrastructure and services, with a specific focus on data protection, quality assurance and information security.

Focus on Vital-IT, SIB's first core facility

Fifteen years after its creation, Vital-IT – one of SIB's largest groups and Switzerland's earliest bioinformatics core facility – continues to evolve. The group is now consolidating its position as a reference in bioinformatics for biomedical data in particular, and has had a new director since 2018: Mark Ibberson.

FROM A SMALL-SCALE HPC CLUSTER TO AN INTERDISCIPLINARY BIOINFORMATICS CENTRE OF REFERENCE

In 2003, SIB launched the first Swiss High-Performance Computing centre to cover the growing computational needs of its various partner institutions.

Fifteen years down the line, over 40 scientists and experts in data analysis, software development, database management, information technology or statistics, backed up by SIB's powerful computational power and storage, are involved in a wide diversity of research projects on the national and international scene.

Recognized for its ability to deliver on a wide variety of projects, the team's activities have long revolved around three main areas: HPC and bioinformatics services for SIB's partner institutions; developments for SIB's activities in the domain of biomedical data; and research collaborations.

Over
40
scientists and experts in data analysis, software development, database management, information technology or statistics

BOLSTERING THE GROUP'S ACTIVITIES FOR THE CLINICAL WORLD

Over the past decade, infrastructure maintenance costs and computational needs have evolved dramatically. Nowadays, the trend for Swiss universities is to set up their own computing infrastructure – for life science but also for other scientific needs. In parallel to this shift in the computational environment, SIB has developed a strong expertise in health-related domains, whether through its Clinical Bioinformatics activities (SEE P. 38), the BioMedIT project or its central role in the Swiss Personalized Health Network (SPHN) as Data Coordination Centre (SEE P. 40).

Therefore, acknowledging the growing needs of biomedical data management and analysis triggered by the development of personalized health research, Vital-IT is expanding in these strategic fields. Its main activities are focused on supporting research for health and personalized medicine, thanks to its recognized skills in software development and bioinformatics analysis in particular. As of 2019, HPC infrastructure activities – including those dedicated to clinical data management – are run through the Core-IT Group (SEE OPPOSITE), in close collaboration with Vital-IT.

Finally, the Vital-IT Group also manages over 1,000 software packages, which are highly used by researchers in its partner institutions.

Altogether, these developments, conducted on top of the group's cross-border research collaborations such as the IMI projects (SEE RIGHT BOX), contribute to reinforce its position as a partner of choice for international health projects.

Fighting diabetes at the European scale

Since 2010, Vital-IT has played a central role in a suite of pan-European projects from the Innovative Medicines Initiative (IMI) in the field of diabetes research. These public-private partnerships share the overarching aim of better understanding, predicting and ultimately preventing type-2 diabetes and its complications, a condition affecting over 6 % of people worldwide. Vital-IT acts both as Data Coordination Centre for data providers across Europe and as data analyst, with over 12 of its experts involved, whether on the database development, software engineering or bioinformatics analysis front. So far, the combined IMI endeavours have notably led to the identification of candidate biomarkers that would allow a detection of diabetes up to nine years before its onset, as well as to the identification of genes involved in the demise of diabetic beta cells.

Mark Ibberson, a new director from within

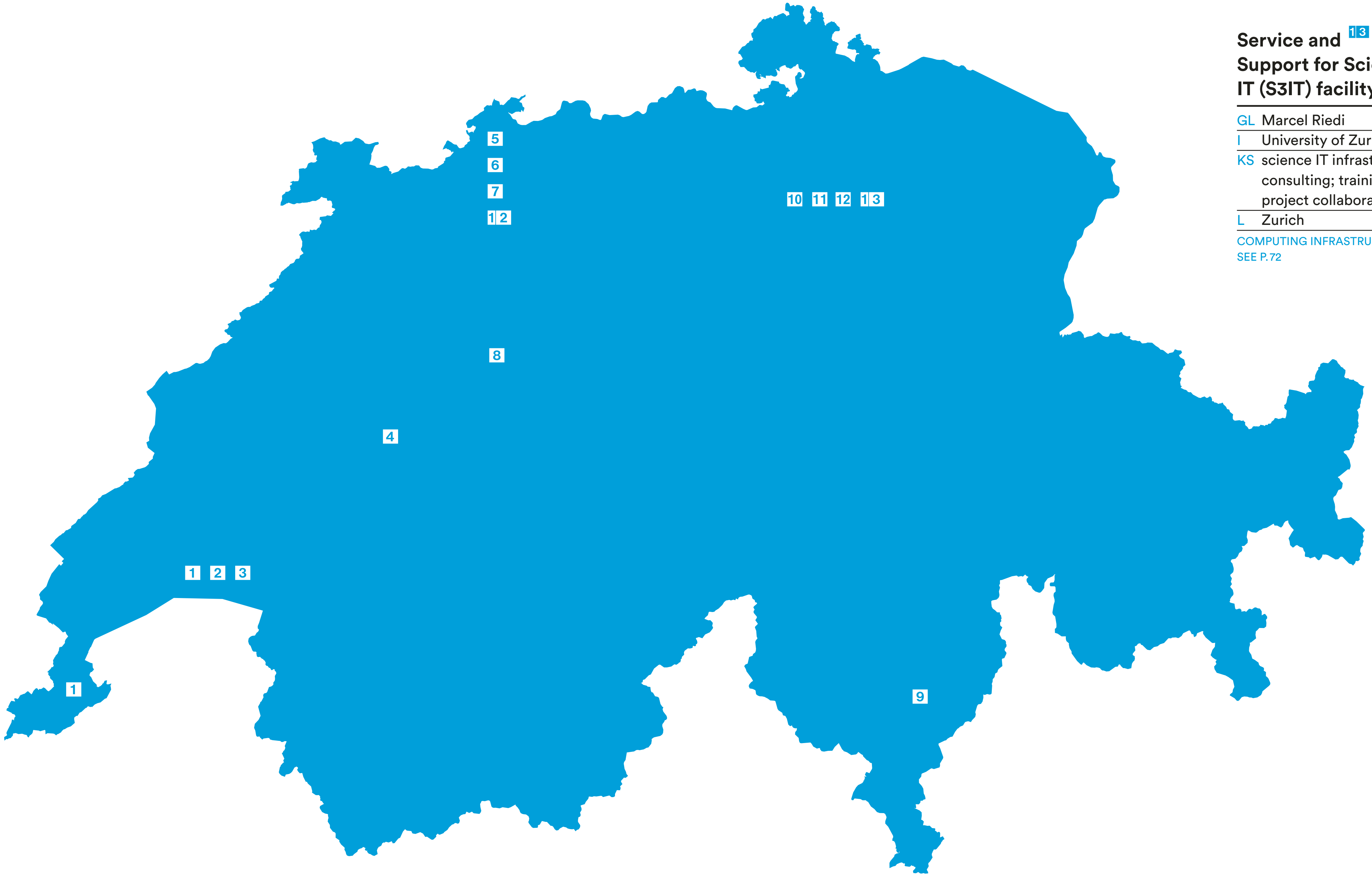
Mark, one of the group's senior scientists, was appointed as Vital-IT's director in November. Mark's experience in the private sector as well as his recognized expertise as co-PI for several international research projects – including the pan-European IMI projects mentioned above – will be invaluable assets for the group. After a PhD in medical genetics at the University of London, Mark has spent 10 years as bioinformatician in a major pharmaceutical company before joining SIB in 2010.

A national network to support the data science needs of life scientists

Thirteen core facilities and competence centres are distributed in all the major Swiss academic institutions, and managed by SIB Group Leaders.

GL = Group leader(s)
I = Institution(s)
KS = Key services provided
L = Location(s)

BioMedIT nodes are located in Basel (sciCORE), Lausanne (Core-IT / Vital-IT) and Zurich (SIS)



Core-IT [1]

HEAD	Heinz Stockinger
I	SIB
KS	secure HPC infrastructure; data protection (incl. training and documentation); quality assurance; IT services and support
L	Lausanne, Geneva

COMPUTING INFRASTRUCTURE
SEE P. 34

Vital-IT [2]

GL	Mark Ibberson
I	SIB
KS	bioinformatics/ biostatistical analyses; data integration and management; software and web development
L	Lausanne

SEE P. 71

Bioinformatics Core Facility (BCF) [3]

GL	Mauro Delorenzi and Frédéric Schütz
I	SIB, University of Lausanne
KS	bioinformatics/ biostatistical analyses; training and consulting; single-cell applications; profiling and sub-typing of tumours
L	Lausanne

SEE P. 71

Bioinformatics Unravelling Group (BUGFri) [4]

GL	Laurent Falquet
I	University of Fribourg
KS	project planning and grant writing; software testing and development; data management plan (DMP) and deposition in OpenData repositories; data analysis; training
L	Fribourg

SEE P. 51

FMI Computational Biology Group (FMICBG) [5]

GL	Michael Stadler
I	Friedrich Miescher Institute
KS	high-throughput data analysis, from experiment design to result interpretation; training; bioinformatics helpdesk; tool development
L	Basel

COMPUTING INFRASTRUCTURE
SEE P. 53

DBM Bioinformatics Core Facility [6]

GL	Robert Ivanek
I	University of Basel, University Hospital Basel
KS	consultancy on experimental design of genomics experiments; data analysis and visualization with focus on large-scale genomics and proteomics; consulting in statistics and programming in R
L	Basel

SEE P. 66

Centre for Scientific Computing (sciCORE) [7]

GL	Torsten Schwede and Thierry Sengstag
I	SIB, University of Basel
KS	scientific services hosting; bioinformatics analysis and training; data management support
L	Basel

COMPUTING INFRASTRUCTURE
SEE P. 72

Interfaculty Bioinformatics Unit (IBU) [8]

GL	Rémy Bruggmann
I	University of Bern
KS	next-generation data analyses; biostatistical analyses; training
L	Bern

COMPUTING INFRASTRUCTURE
SEE P. 70

Service and Support for Science IT (S3IT) facility [13]

GL	Marcel Riedi
I	University of Zurich
KS	science IT infrastructure; consulting; training; project collaborations
L	Zurich

COMPUTING INFRASTRUCTURE
SEE P. 72

Scientific Information Services (SIS) [12]

GL	Bernd Rinn
I	ETH Zurich
KS	research data management; scientific software engineering; data management plan (DMP); research data pipelines; data co-analysis
L	Zurich, Basel

COMPUTING INFRASTRUCTURE
SEE P. 72

NEXUS Personalized Health Technologies [11]

GL	Daniel Stekhoven
I	ETH Zurich
KS	bioinformatics/ biostatistical data analysis; translational research support; clinical decision support; software development
L	Zurich

COMPUTING INFRASTRUCTURE
SEE P. 73

FGCZ Genome Informatics (FGCZ-GI) [10]

GL	Hubert Rehrauer
I	ETH Zurich, University of Zurich
KS	next-generation data analysis; analysis of long-reads; bioinformatics training; clinical genomics
L	Zurich

COMPUTING INFRASTRUCTURE
SEE P. 71

Bioinformatics Core Unit (BCU) [9]

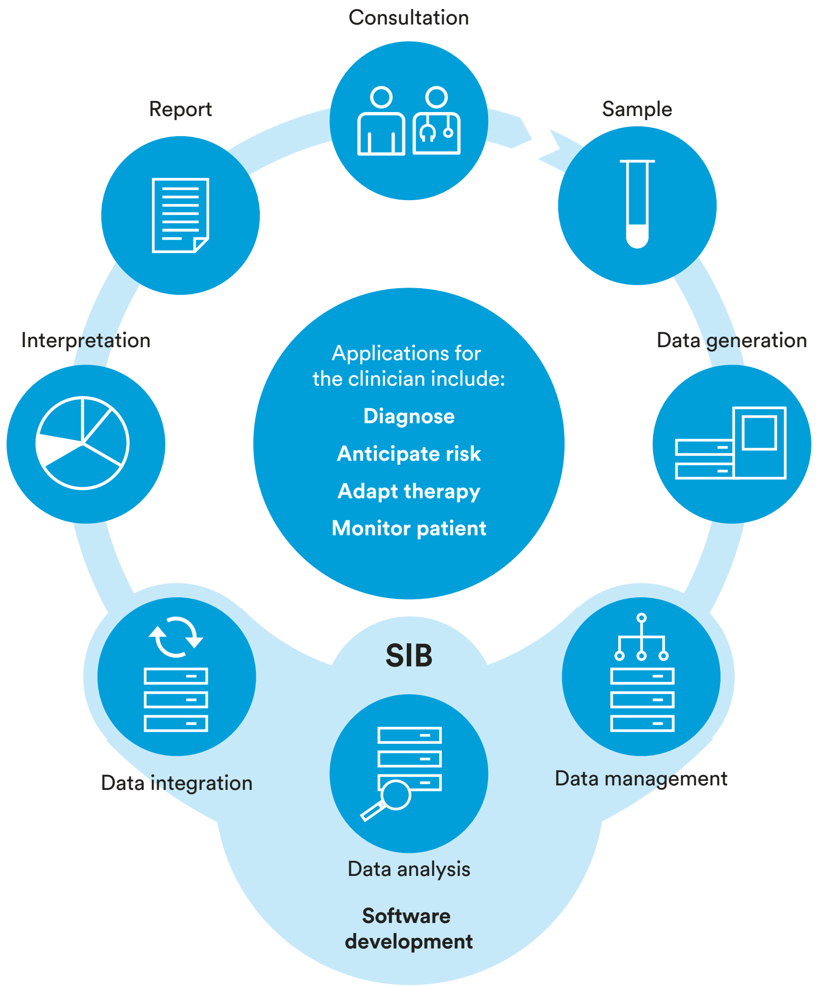
GL	Luciano Cascione
I	Institute of Oncology Research
KS	biostatistical analyses; data mining; training
L	Bellinzona

COMPUTING INFRASTRUCTURE
SEE P. 70

III.
PERSONALIZED HEALTH

SIB is playing a leading role in the Swiss personalized health landscape, by sharing its expertise with hospitals and by building the nationwide infrastructure that will enable patients' data to be used for research.

Clinical bioinformatics for medical care



NAME **CLINICAL BIOINFORMATICS GROUP**

GROUP LEADER Valérie Barbié

TAGS human genetics; infectious disease; interoperability; medicine and health; personalized medicine; oncology; outreach; training

The Group provides expertise and support for the organization, analysis and interpretation of patient-related data (e.g. omics data), converting them into clinically useful information for health professionals. Notably, it:

Establishes consensus and good practices for diagnostic omics analyses across Swiss hospitals;

Partners with Swiss hospitals to develop, implement and sustain state-of-the-art diagnostic tools;

Provides clinical bioinformatics training across Switzerland;

Facilitates interactions between SIB's research groups and the medical realm for clinical research projects.

VALÉRIE BARBIÉ
On 21 June 2018, Valérie Barbié was appointed Group Leader by SIB's Foundation Council

2018 highlights

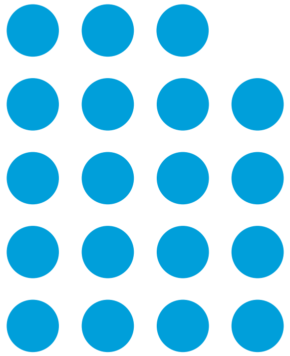
WORKING GROUPS

With each hospital using different bioinformatics methods to analyse their molecular data, such as DNA from tumour tissues, results are likely to vary significantly from one to another. Several benchmarking projects led by the Clinical Bioinformatics Group thus aim to identify the most appropriate methods and harmonize practices among hospitals, leading to faster diagnosis and better targeting of treatments:

THE SOMATIC MUTATION CALLING WORKING GROUP has benchmarked Swiss hospitals' next-generation sequencing analysis and interpretation practices in cancer diagnosis;

THE MICROBE TYPING AND CHARACTERIZATION WORKING GROUP has benchmarked clinical next-generation sequencing practices in the diagnosis of infectious diseases, both in virology and bacteriology.

The results of these two benchmarking studies will be published in 2019.



19
collaborating institutions (universities, hospitals, etc.)

COLLABORATIVE CLINICAL PLATFORMS

The Group is co-leading the **SWISS VARIANT INTERPRETATION PLATFORM** project (SVIP), funded by SPHN and BiomedIT, which aims to provide a central resource for clinicians and researchers to annotate and interpret clinically verified variants in cancer. The tool will be available in 2019. SVIP is a partner in the ELIXIR Annotation and Curation of Human Genomic Variations implementation study, which aims to understand the existing infrastructure, resources and protocols for human genome variation annotation and curation.
svip.ch

The group is also actively involved in a multi-site National Research Programme "Antimicrobial Resistance" (NRP72) project led by the University Hospital of Basel, on the development of a **SWISS PATHOGEN SURVEILLANCE PLATFORM** – a nationwide database that will make it possible to track inter-cantonal outbreaks of multiresistant pathogens in near real-time. Version 1 will be released in early 2019.
spsp.ch

PARTNER HOSPITALS

At the Geneva University Hospitals, OncoBench® – a tool for analysing and storing genomic data in the field of oncology – has been running since 2016, and version 3 will be released in early 2019. A tailored management and analysis platform for medical genetic data is currently under development for the Institute of Medical Genetics of the University of Zurich.

TRAINING

Organized jointly with the University Hospitals of Basel and Lausanne, the first Certificate of Advanced Studies in **PERSONALIZED MOLECULAR ONCOLOGY** has started in the fall of 2018 and continues over 2019 with 20 participants.
pmo.unibas.ch

Other training activities included the second edition of a course specifically designed for hospital staff on next-generation sequencing quality control and annotation, and a contribution to a MOOC on Precision Medicine, both in collaboration with the Geneva University Hospitals. On the international level, the Group also ran a training event in Bahrain, and is co-organizing a European technical workshop on microbial genomics with the Lausanne University Hospital.

In 2018, the Group continued to take part in outreach-related events such as the Planète Santé health fair, and a series of videos about personalized health, funded by the Leenaards Foundation.
goo.gl/f1y4KY

What should family doctors know about personalized health?

Two articles in the Swiss Medical Newsletter (DE/FR) highlighted the relevance of the Group's activities for general practitioners in 2018

NAME **PERSONALIZED HEALTH INFORMATICS GROUP (PHI)**

DIRECTOR Katrin Crameri

TAGS data management; data security; interoperability; medicine and health; personalized medicine

The Group establishes mechanisms for nationwide interoperability and sharing of health-related data, in particular in the context of the Swiss Personalized Health Network (SPHN). It lays the foundations needed to facilitate research projects in this area. PHI manages the SPHN Data Coordination Centre and the BioMedIT project (SEE BELOW), thereby covering three focus areas:

Technical interoperability standards: facilitating processes for data exchange and management between hospitals, analytics platforms and researchers; managing the national network of secure bioinformatics core facilities (BioMedIT project); coordinating expert working groups.

Semantic interoperability standards: defining, implementing and monitoring processes for national interoperability data standards; coordinating expert working groups and SPHN Driver Projects, while aligning with international standardization efforts.

Information hub & training: providing platforms for efficient information exchange and competence building within the heterogeneous SPHN network.

KATRIN CRAMERI
As of 1 March 2019, Katrin Crameri is heading SIB's Personalized Health Informatics Group

2018 highlights

Focus on the BioMedIT project

To promote the development of personalized health research, the Swiss government launched the Swiss Personalized Health Network (SPHN) in 2017. In parallel, the BioMedIT project was funded to establish a coordinated network of core facilities at Swiss universities for the secure processing of biomedical data.

BIOMEDIT IN A NUTSHELL

The ability of life scientists to advance our knowledge of diseases and to identify novel and personalized therapies hinges on their access to large amounts of high-quality patient data. However, such data are associated with stringent levels of protection and information security, which impose strict requirements on the underlying information technology infrastructure, processes and expertise.

BioMedIT will establish a coordinated nationwide network of secure infrastructures (the 'nodes') to support biomedical research, by building on existing expertise and research infrastructure at the partner institutions. At this stage, three BioMedIT nodes are being set up, in Basel (operated by SIB's and the University of Basel's sciCORE Group), Lausanne (operated by SIB) and Zurich (operated by ETH Zurich's SIS Group).

BioMedIT will thus provide the researchers with a secure high-performance computing environment – and the accompanying technical expertise – to leverage knowledge from patient data. Discover "The researcher's new path to nationwide clinical data" on the opposite page.

WHERE ARE WE TODAY?

The first successful steps towards a coordinated encryption/decryption workflow were made in September 2018. They included a computational analysis workflow in the form of a container to demonstrate the feasibility of robust, distributed data analysis across the BioMedIT network. An "SPHN information security policy" clarifying the roles and responsibilities of the various parties was also rolled out in September, and a Security Awareness training workshop was held in Zurich at the end of last year.

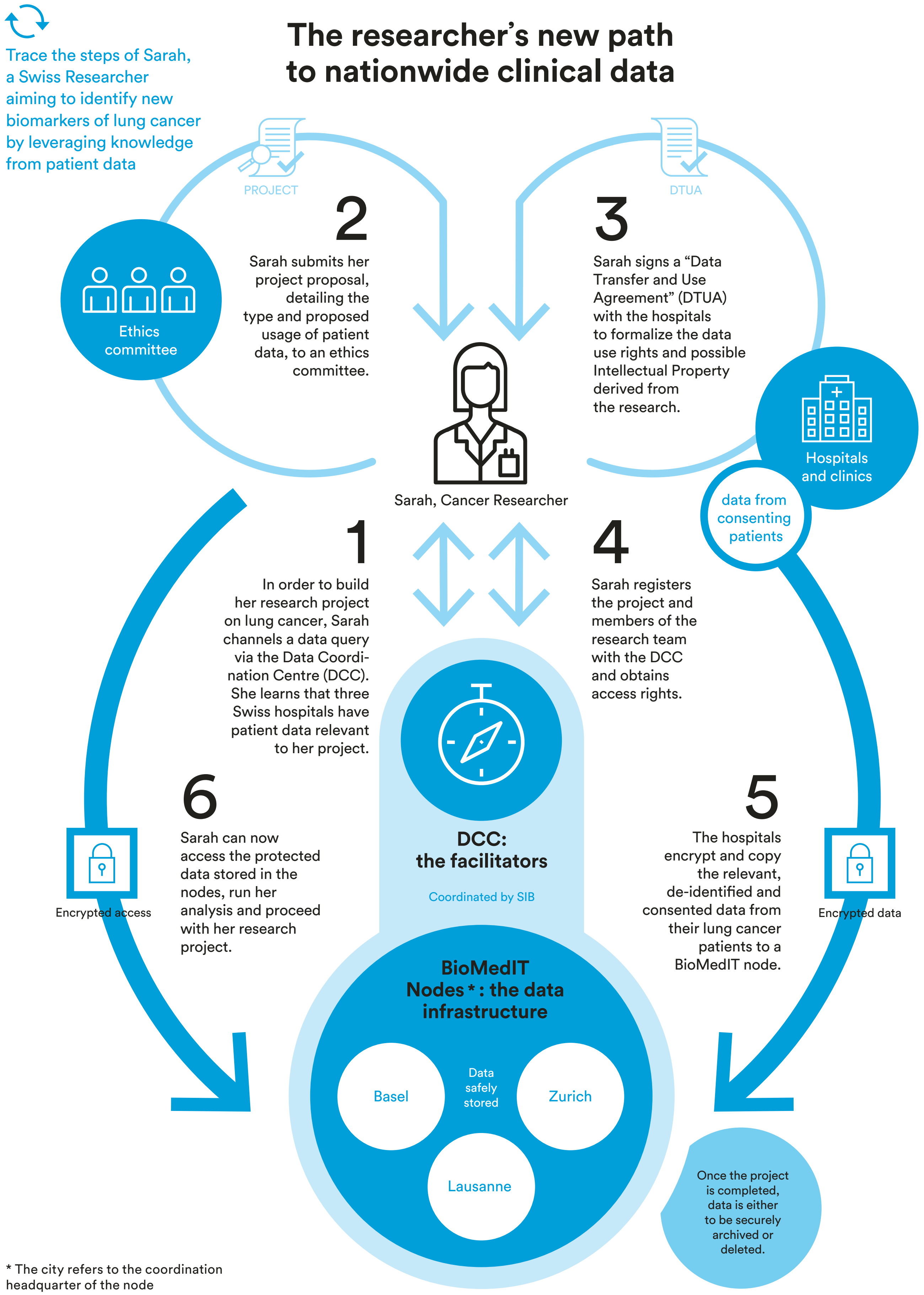
Today, the high-performance computing infrastructure (including both storage and computing) dedicated to BioMedIT is already operational in Zurich. It is currently being deployed in Basel, and will follow in Lausanne in the coming months.

Martin Fox, BioMedIT Project Manager at PHI: "Not all BioMedIT nodes evolve at the same pace, each location having its own technical constraints. While it is one thing to have the individual nodes ready to operate and welcoming users, it is another to have the full network interoperable and securely connected to the hospital data warehouses. I am confident that by the end of 2019 we will have three BioMedIT nodes on line, and the frameworks in place to efficiently process data for personalized health research."

WHY IS SIB LEADING THIS PROJECT?

SIB has been mandated by the Swiss Government to lead the Data Coordination Centre of the SPHN initiative as well as the BioMedIT project because of its unique position as a non-profit, nationwide foundation with no interest other than the advancement of life science and health. Thanks to 20 years of experience in research infrastructure, its involvement as Data Coordination Centre for several public-private Europe-wide health projects and its established partnerships with university hospitals for clinical bioinformatics activities, SIB is indeed well positioned to bridge the gap between research and the clinical world.

The researcher's new path to nationwide clinical data



Trace the steps of Sarah, a Swiss Researcher aiming to identify new biomarkers of lung cancer by leveraging knowledge from patient data

1 In order to build her research project on lung cancer, Sarah channels a data query via the Data Coordination Centre (DCC). She learns that three Swiss hospitals have patient data relevant to her project.

2 Sarah submits her project proposal, detailing the type and proposed usage of patient data, to an ethics committee.

3 Sarah signs a "Data Transfer and Use Agreement" (DTUA) with the hospitals to formalize the data use rights and possible Intellectual Property derived from the research.

4 Sarah registers the project and members of the research team with the DCC and obtains access rights.

5 The hospitals encrypt and copy the relevant, de-identified and consented data from their lung cancer patients to a BioMedIT node.

6 Sarah can now access the protected data stored in the nodes, run her analysis and proceed with her research project.

Once the project is completed, data is either to be securely archived or deleted.

* The city refers to the coordination headquarter of the node

A strong bioinformatics community

IV.
TRAINING IN BIOINFORMATICS

One of SIB's missions is to train the next generation of bioinformaticians and to ensure that scientists make the most out of bioinformatics resources. The SIB Training Group promotes and coordinates training in bioinformatics, both in Switzerland and internationally.

23
SIB groups engaged in teaching activities

81
experts and trainers

56
courses and workshops

105
training days

204
students part of the SIB PhD Training Network

1,265
trainees

Working closely with Patricia Palagi, Team Leader of the Training Group, a dedicated committee composed of SIB Group Leaders as well as external scientists and chaired by Marc Robinson-Rechavi, provides strategic outlook on the course content and audience reach.

Metrics for training quality and impact

Assessing the quality and the long-term impact of training on the work of participants is of great importance to SIB, which has been systematically evaluating its courses since 2013. In 2018, SIB's metrics were integrated into a pan-European analysis led by ELIXIR. The final outcome of this study will be released in 2019, and the preliminary results already show the benefits European and SIB's bioinformatics training have in the researchers' work, such as an improvement in their ability to handle and analyse scientific data.

The SIB Training portfolio is constantly evolving to meet the scientific community's needs: find the full list of courses at

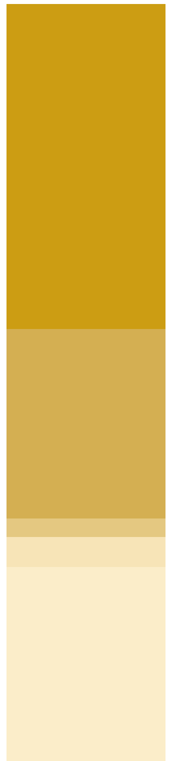
sib.swiss/training

FAIR training

Embracing FAIR principles will directly benefit SIB courses, trainers and course participants, by making training events and materials "Findable, Accessible, Interoperable and Reusable". In order to tackle the challenges and developments required in this context, SIB Training co-led an ELIXIR workshop in 2018. The practical aspects of FAIR training implementation are currently being set up and should be integrated into SIB courses in the next years.

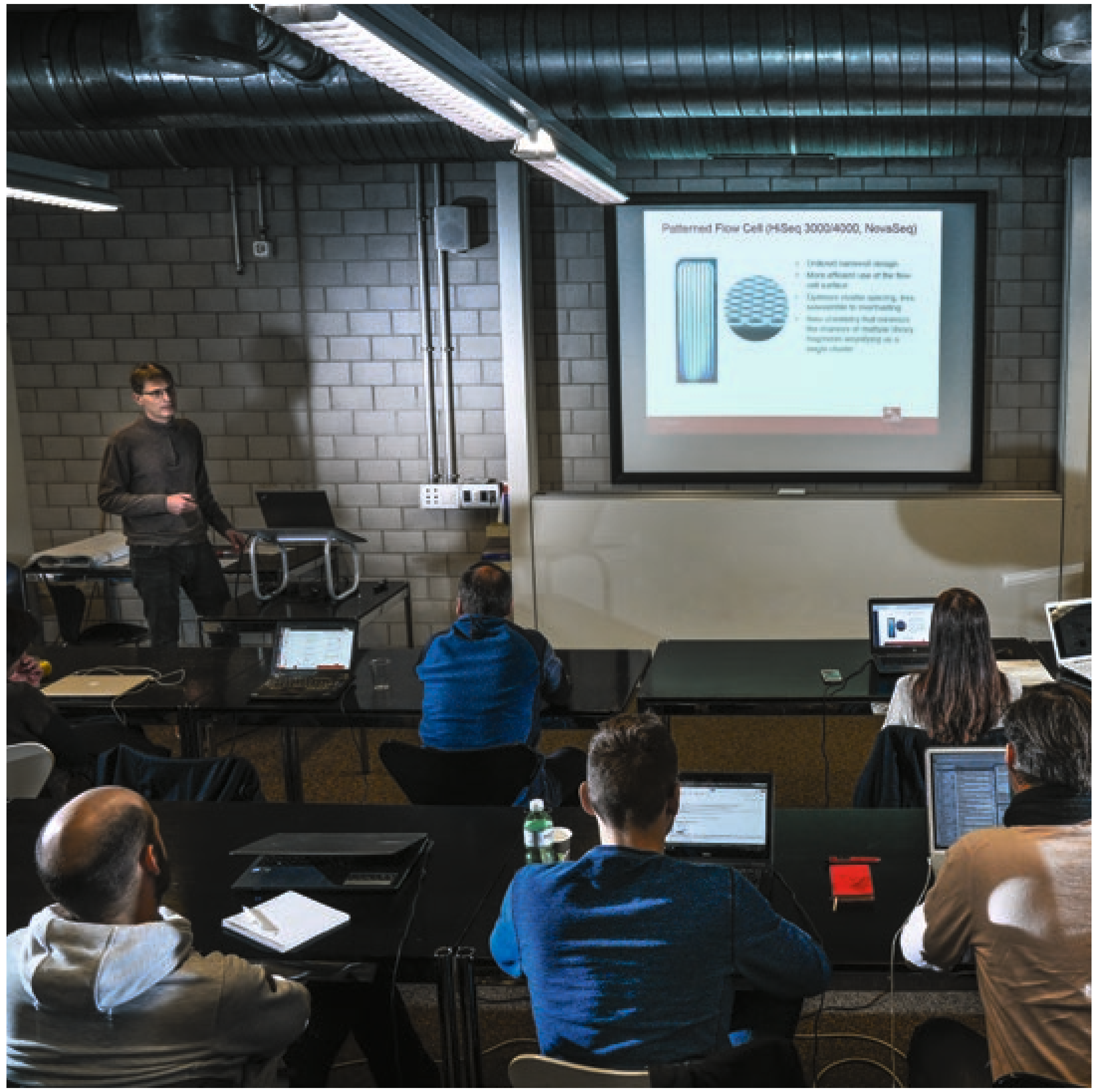
Reaching out to Europe and the world

While the vast majority of attendees at SIB courses originate from Switzerland, in 2018 participants also came from Belgium, Brazil, Canada, Czech Republic, France, Germany, Italy, the Netherlands, Poland, Qatar, Spain and Tunisia.



Who takes part in SIB courses?

- ■ 43% PhD students
- ■ 25% Postdocs
- ■ 2.5% Principal Investigators
- ■ 4% Laboratory technicians
- ■ 25.5% Other scientists



Some of the most taught skills in 2018

Machine learning Sensitive data management
Data Management Plan R software
Workflows Variant analysis Computer-aided drug design
Unix Next-generation sequencing HPC
Metabolomics Protein databases RNA-Seq
Best coding practices Computational biology
Enrichment analysis Statistics
Python Single-cell sequencing Chip-seq
Long read sequencing Big data

Bioinformatics capacity building in South Africa

The development of bioinformatics expertise among future African scientists is key for finding solutions to local health problems but often relies on powerful and complex infrastructure. The Swiss Tropical and Public Health Institute (Swiss TPH) joined forces with SIB's Training and sciCORE Groups to create a novel virtual educational infrastructure to support a bioinformatics course in South Africa. It is now being adapted to enable its use in additional courses.

Creating opportunities to exchange know-how and promoting scientific results broadly

V.

SCIENTIFIC COLLABORATION

SIB is a unique ecosystem, in which Swiss bioinformatics service providers rub shoulders with users – researchers working at the interface between experimental and computational biology – on a daily basis.

Key scientific events in 2018

At SIB, fostering scientific collaboration means creating opportunities for our scientists to exchange know-how, and to promote their work. Some of the actions implemented during the year are listed below, followed by a digest of some of the key research results on the next page.

SIB DAYS
KEY FIGURES

10	20-minute oral presentations
84	posters
7	workshops
10	lightning talks by young researchers
3	awardees

SIB Days – special 20th anniversary edition, Biel

SIB's two-day internal conference, allowing the Swiss bioinformatics community to get together;

Keynote speakers included Nick Goldman (EMBL-EBI) on connected DNA and data storage and Fabian Theis (Institute of Computational Biology at the Helmholtz Zentrum) on single-cell RNA sequencing;

Lakeside celebrations on the first evening provided the hundreds of SIB Members and guests with a relaxed and informal moment to celebrate SIB's anniversary.

Congratulations to the winners of the SIB Days Awards 2018

Best Lightning Talk Award

EMMA RICART, PhD student from the Proteome Informatics Group, Geneva (Group Leader: Frédérique Lisacek)

ADITHI VARADARAJAN, PhD student from the Bioinformatics and Proteogenomics Group, Wädenswil (Group Leader: Christian Ahrens)

Best Poster Award

DAVID DYLUS, Postdoctoral researcher from the Laboratory of Computational Evolutionary Biology, Lausanne (Group Leader: Christophe Dessimoz)

GA4GH 6th plenary meeting, Basel

Annual meeting of the Global Alliance for Genomics and Health (GA4GH), a policy-framing and technical standards-setting organization, seeking to enable responsible genomic data sharing within a human rights framework;

SIB and ELIXIR co-hosted the event;

The first set of deliverables developed under the five-year strategic plan to enable real-world genomic data sharing by 2022 was presented at the event.

LS2 Annual meeting, Lausanne

International scientists from all backgrounds gathered to explore the diverse range of life science topics;

SIB held a symposium named "Bioinformatics and computational biology: an evolving field".

Biology18, Neuchâtel

The Swiss conference on organismic biology and joint congress of the Swiss Zoological Society, the Swiss Botanical Society, and the Swiss Systematics Society;

For the second time, SIB organized workshops presenting a selection of resources and tools: Bgee, ATLAS and a multiple sequence alignment reconstruction tool.

Key research outcomes in 2018

DNA loops and smoke rings have a lot in common

Why do DNA loops forming different knot types sediment at different speeds? Researchers from SIB, the University of Lausanne and EPFL, together with Polish colleagues, observed that this property resembled another, fascinating physics phenomenon, namely the way smoke rings behave in the air. Their results could improve our understanding of the shape and function of biomolecules at large.

Published in *Physical Review Letters*

SIB Group involved

DNA and Chromosome Modelling P. 54

How the African elephant cracked its skin to cool off

An intricate network of minuscule crevices adorns the skin surface of the African bush elephant. By retaining water and mud, these micrometer-wide channels greatly help elephants in regulating their body temperature and protecting their skin against parasites and intense solar radiation. SIB Researchers at the University of Geneva found that African elephant skin channels are true fractures of the animal's skin outermost layer.

Published in *Nature Communications*

SIB Group involved

Artificial & Natural Evolutionary Development of Complexity P. 66

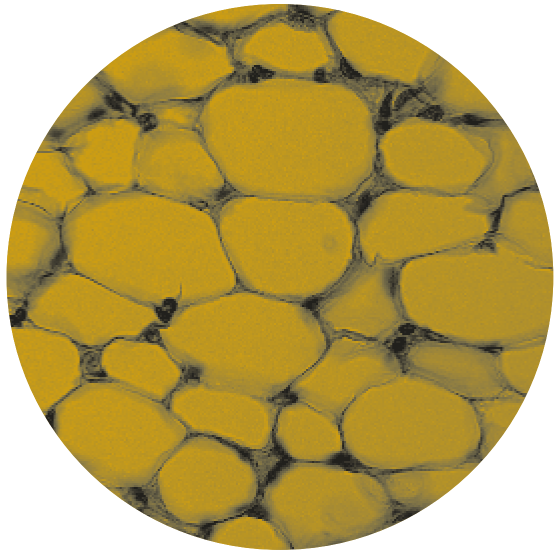
The genetic burden of pioneers

Using a unique combination of genomic and historical data, researchers from SIB at the University of Bern, in collaboration with Canadian colleagues, showed that having a pioneer ancestor could explain the excess of harmful mutations observed in recently colonized areas of Quebec. The findings have broad implications for disease incidence in expanding human populations.

Published in *Genetics*

SIB Group involved

Computational and Molecular Population Genetics P. 59



Fat, a personal affair

Thirteen percent of the world's population suffers from obesity, a particularly complex multi-factor disease. Understanding the different factors that bring it about, and defining adiposity subgroups, would offer the possibility of more personalized interventions. Together with colleagues from the University of Regensburg, SIB Scientists at the Lausanne University Hospital describe a first important step towards this goal.

Published in *Nature Communications*

SIB Group involved

Statistical Genetics P. 52

Decoding the genome of the wheat stem sawfly, a major agricultural pest

What makes the wheat stem sawfly such a major pest in the grasslands of North America? An international study co-led by SIB Researchers at the University of Lausanne unraveled its genome and found clues on possible mechanisms it employs to select its host plants.

Published in *Genome Biology and Evolution*

SIB Group involved

Evolutionary-Functional Genomics P. 55

Zooming into our metabolism

Every minute, each one of our cells carries out countless chemical reactions to convert food into energy, build proteins or eliminate toxic waste. Taken as a whole, these reactions make up our metabolism. Researchers from SIB and the Lausanne University Hospital found a new method to understand how metabolism is regulated at a far higher resolution.

Published in *Bioinformatics*

SIB Group involved

Vital-IT P. 71

Cells are able to detect more shades of information than previously thought

Hacking a cell's communication system is a common feat of information diseases, such as cancer. Researchers at SIB and the University of Lausanne provided the first direct experimental evidence that one of the most common cell receptors – until then thought to function only as on/off switches – can also reliably convey intermediate levels of signal concentrations, thus acting rather like a 'dimmer'.

Published in *Nature Communications*

SIB Group involved

Computational Biology P. 50

Cracking the 'dark matter' in bacterial genomes

Bacterial genomes are a treasure trove of information, be it for the development of novel antibiotics or the protection of crops against pathogens. SIB Scientists at the Agroscope in Wädenswil provided hints as to how to resolve the highly complex regions these genomes contain and revealed that they may contain functionally important features.

Published in *Nucleic Acids Research*

SIB Group involved

Bioinformatics and Proteogenomics P. 55

Fields of activity of SIB Groups

Bioinformatics is an interdisciplinary field, where the combination of genetics, physiology, chemistry and physics leads to many fields of activities and applications.

Genes and genomes

Life's instruction manual

A genome is the sum of genetic material of an organism, including all of its genes. It is composed of DNA and contains all the information needed to create and maintain an organism, as well as the instructions on how this information should be expressed.

Bioinformatics develops tools able to read genomes, and store, analyse and interpret the resulting data.

SEE P. 50

Proteins and proteomes

More than meets the eye

A proteome is the sum of proteins expressed by a cell, a tissue or an organism, at a given time. Proteins are the products of genes, and are involved in nearly every task carried out within an organism – from carrying oxygen to fighting off pathogens.

Bioinformatics develops tools to understand the role of proteins.

SEE P. 55

Evolution and phylogeny

Splitting ends

Changes that occur in genomes tell life scientists how an organism has evolved over time. Comparisons made between genomes from different species or populations tell them how species are related to one another – this is the field of phylogenetics.

Bioinformatics develops tools able to compare the genomes of organisms, as well as computing methods to reconstruct their past and build their 'family' trees.

SEE P. 58

Structural biology

The third dimension

Macromolecules such as DNA and proteins have specific 3D structures that are dictated by their sequence. A protein's function is defined by its 3D structure, or architecture, which in turn defines the way it reacts with other molecules.

Bioinformatics develops software to create 3D models of proteins to study their interactions with other molecules, such as drugs.

SEE P. 62

Systems biology

Never alone

Life occurs and is sustained by a mesh of interactions within and between cells, tissues, organisms, and their environment. Understanding how these complex systems function allows scientists to predict what happens if one of the components changes or the conditions are altered.

Bioinformatics develops models to delineate metabolic pathways.

SEE P. 64

Text mining and machine learning

Rise of the machines

Text-mining algorithms are designed to recognize patterns within text so that computers can extract the information of interest, such as biomedical terms. Using machine learning, computers can also acquire the ability to learn without explicit instructions.

Bioinformatics develops text-mining tools that complement expert biocuration, as well as machine-learning techniques that can improve the prediction of protein structures.

SEE P. 68

Core facilities and competence centres

The means to an end

The quantity of data generated by the life sciences has grown exponentially over the years, and needs to be stored and processed. Researchers also need help in interpreting their data. Core facilities centralize research resources, and provide tools, technologies, services and expert consultation to this end.

SIB coordinates 13 core facilities and competence centres located in the major Swiss academic institutions.

SEE P. 70

Many areas of application, including…

…AGRICULTURE
from predicting the spread of bird flu outbreaks and understanding the lifecycle of agricultural pests, to improving crop productivity.

…BASIC RESEARCH
from unravelling the evolutionary processes that have shaped today's biodiversity, to solving the equation behind a lizard's scale colour pattern.

…MEDICINE AND HEALTH
from designing optimized proteins in cancer immunotherapy, to creating biomedical decision-support tools using text mining.

…ENVIRONMENTAL SCIENCES
from understanding how organisms adapt to climate change, to how microbial communities can be used to break down pollutants in oil spills.

Michael BAUDIS

GROUP NAME **COMPUTATIONAL ONCOGENOMICS GROUP**
INSTITUTION UNIVERSITY OF ZURICH
CITY ZURICH

The Computational Oncogenomics Group works on the computational analysis of structural genome variations in cancer, centred around our arrayMap and Progenetix resources of curated molecular-cytogenetic and sequencing data. Specific projects explore computational methods, genomics of selected tumour entities and genomic variation patterns across malignancies and populations. As members of the Global Alliance for Genomics and Health, the group is developing standards in biocuration and data sharing for genomic variants and phenotypic data. Other projects are related to genome data epistemology, e.g. geographic and diagnostic sampling bias.

MAIN DOMAIN OF ACTIVITY
Genes and genomes

SECONDARY DOMAINS OF ACTIVITY
API; Data mining; Data visualization; Database curation; DNA microarrays; Human genetics; Oncology; Ontology; Personalized medicine; Population genetics

DOMAIN OF APPLICATION

KEY RESOURCES
Progenetix
Arraymap
Schemablocks

KEY PAPERS IN 2018

Gao B *et al.*
segment_liftover: a Python tool to convert segments between genome assemblies.
F1000 Res.

Carrio-Cordo P and Baudis M.
Mountains and chasms – Surveying the oncogenomic publication landscape.
Oncology

Dyke SOM *et al.*
Registered access: authorizing data access.
Eur J Hum Genet.

Sven BERGMANN

GROUP NAME **COMPUTATIONAL BIOLOGY GROUP**
INSTITUTION UNIVERSITY OF LAUSANNE
CITY LAUSANNE

We develop concepts and algorithmic tools for the analysis of large-scale biological and clinical data. We participate in many genome-wide association studies (GWAS) for human traits and have a particular interest in the integration of genotypic and complex phenotypic datasets (such as gene expression or metabolomics). A key approach is the reduction of complexity through modular and network analysis. A complementary direction of our research pertains to relatively small genetic networks, whose components are well known.

MAIN DOMAIN OF ACTIVITY
Genes and genomes

SECONDARY DOMAINS OF ACTIVITY
Biostatistics; Epigenetics; Genetic networks; GWAS; Metabolomics; Microarrays; Protein networks; Software development; Systems biology; Transcriptomics

DOMAINS OF APPLICATION

KEY RESOURCES
pascal
metabomatching
Iterative signature algorithm

Philipp BUCHER

GROUP NAME **COMPUTATIONAL CANCER GENOMICS GROUP**
INSTITUTION EPFL
CITY LAUSANNE

In the Computational Cancer Genomics Group, we are interested in gene regulation in both healthy and diseased cells. Breakthroughs in genomics technologies have led to the production of large volumes of data that could potentially tell us something about how gene regulatory instructions are encoded in our DNA. Our group develops new algorithms, computer programs, web services and databases that will help us and others to extract knowledge and understanding from such data. We also collaborate with experimental biologists on interdisciplinary projects and organize postgraduate courses in genomic data analysis.

MAIN DOMAIN OF ACTIVITY
Genes and genomes

SECONDARY DOMAINS OF ACTIVITY
Benchmarking; Comparative genomics; Database curation; Epigenetics; Functional genomics; Gene regulation; Transcriptomics

DOMAIN OF APPLICATION

KEY RESOURCES
EPD
ChIP-seq tools and web server
SNP2TFBS

KEY PAPERS IN 2018

Ambrosini G *et al.*
PWMScan: a fast tool for scanning entire genomes with a position-specific weight matrix.
Bioinformatics

Dreos R *et al.*
MGA repository: a curated data resource for ChIP-seq and other genome annotated data.
Nucleic Acids Res.

Fourati S *et al.*
A crowdsourced analysis to identify *ab initio* molecular signatures predictive of susceptibility to viral infection.
Nat Commun.

Bart DEPLANCKE

GROUP NAME **LABORATORY OF SYSTEMS BIOLOGY AND GENETICS**
INSTITUTION EPFL
CITY LAUSANNE

In the Laboratory of Systems Biology and Genetics (LSBG), we use high-throughput sequencing, single-cell genomics, microfluidics, and computational approaches: 1) to decipher the regulatory code in Drosophila and mammals with a specific focus on mesenchymal stem cell function, adipose biology and gut immunity, and 2) to examine how variations in this code affect molecular and organismal diversity. In addition to our research interests, we are actively pursuing the development of new research tools and computational pipelines that enable a better characterization of gene regulatory networks.

MAIN DOMAIN OF ACTIVITY
Genes and genomes

SECONDARY DOMAINS OF ACTIVITY
Functional genomics; Gene regulation; Gene regulatory network analysis; GWAS; Human genetics; Machine learning; Microfluidics; Personalized medicine; Single-cell biology; Transcriptomics

DOMAIN OF APPLICATION

KEY RESOURCES
ASAP
GETPrime

KEY PAPERS IN 2018

Schwalie PC *et al.*
A stromal cell population that inhibits adipogenesis in mammalian fat depots.
Nature

Wang QA *et al.*
Reversible de-differentiation of mammary adipocytes into adipose precursor-like cells during lactation.
Cell Metab.

Emmanouil DERMITZAKIS

GROUP NAME **GENOMICS OF COMPLEX TRAITS GROUP**
INSTITUTION UNIVERSITY OF GENEVA
CITY GENEVA

Our group has a strong interest in population genomics and genetics of complex traits. We use various methodologies to understand the role of genetic variation in phenotypic variation and what fraction of genetic variation is harboured within functional elements of the human genome. Our main focus is on genome-wide analysis of gene expression and cellular phenotypes and their association with nucleotide variation. We attempt to detect functional genetic variation in regulatory elements and then use functional variation and accurately measured gene expression variation to bridge the genotype with disease phenotypes in association studies.

MAIN DOMAIN OF ACTIVITY
Genes and genomes

SECONDARY DOMAINS OF ACTIVITY
Comparative genomics; Population genetics; Epigenetics; Functional genomics; Genome structure; GWAS; Next-generation sequencing; Oncology; Systems biology; Transcriptomics

DOMAIN OF APPLICATION

KEY RESOURCE
QTLtools

Laurent FALQUET

GROUP NAME **BIOINFORMATICS UNRAVELLING GROUP (BUGFRI)**
INSTITUTION UNIVERSITY OF FRIBOURG
CITY FRIBOURG

In the Bioinformatics Unravelling Group of the University of Fribourg (BUGFri), we support life science researchers by providing expertise in data analysis of next-generation sequencing experiments, or any large-scale biological experiment requiring bioinformatics resources. We focus on genome assembly, annotation and comparison, as well as on mutant and structure variant identification by resequencing. We also perform metagenomics, RNAseq and ChIPseq data analysis, proteome clustering and ortholog/paralog classification, as well as pathway and gene set enrichment analysis.

MAIN DOMAINS OF ACTIVITY
Genes and genomes
Core facilities and competence centres

SECONDARY DOMAIN OF ACTIVITY
Comparative genomics; Core facilities and competence centres; Data management; Data mining; Data visualization; Epigenetics; Metagenomics; Microbiology; Transcriptomics; Workflows

DOMAINS OF APPLICATION

KEY RESOURCES
TASmania
PACMAN

KEY PAPERS IN 2018

Aguilar-Bultet L *et al.*
Genetic separation of *Listeria monocytogenes* causing central nervous system infections in animals.
Front Cell Infect Microbiol.

Thoms M *et al.*
Suppressor mutations in Rpf2-Rrs1 or Rpl5 bypass the Cgr1 function for pre-ribosomal 5S RNP-rotation.
Nat Commun.

Babatunde KA *et al.*
Malaria infected red blood cells release small regulatory RNAs through extracellular vesicles.
Sci Rep.

Jacques FELLAY

GROUP NAME **HOST-PATHOGEN GENOMICS GROUP**
INSTITUTION EPFL
CITY LAUSANNE

In the Host-Pathogen Genomics Group, we explore the genetic roots of inter-individual differences in response to infections, with a particular focus on the genomic interactions between pathogens and their human hosts. At the crossroads between basic science and the clinical world, we are committed to translational genomic research, aimed at identifying, validating and bringing to clinical use genetic markers of susceptibility to infectious diseases – e.g. host genomics of HIV infection, joint analyses of interactions between human and viral genomes, and exome sequencing on patients with extreme infectious disease phenotypes.

MAIN DOMAIN OF ACTIVITY
Genes and genomes

SECONDARY DOMAINS OF ACTIVITY
Human genetics;
Immunology;
Infectious diseases;
Personalized medicine

DOMAIN OF APPLICATION

KEY PAPERS IN 2018

Scepanovic P *et al.* Human genetic variants and age are the strongest predictors of humoral immune responses to common pathogens and vaccines. Genome Med.

Patin E *et al.* Natural variation in the parameters of innate immune cells is preferentially driven by genetic factors. Nat Immunol.

Naret O *et al.* Correcting for population stratification reduces false positive and false negative results in joint analyses of host and pathogen genomes. Front Genet.

Zoltán KUTALIK

GROUP NAME **STATISTICAL GENETICS GROUP**
INSTITUTION CHUV / UNIVERSITY OF LAUSANNE
CITY LAUSANNE

In the Statistical Genetics Group, we are interested in the development of statistical methodologies in order to decipher the genetic architecture of complex human traits related to obesity. To do this, we efficiently combine large-scale genome-wide association studies (GWAS) with various -omics data. Our methods improve genetic fine-mapping, reveal gene-environment interactions, dissect genetic subtypes of obesity, enhance causal effect estimation and detect various statistical artefacts. Furthermore, we are involved in large consortia researching the genetic basis of anthropometric traits (GIANT) and longevity (LifeGen).

MAIN DOMAIN OF ACTIVITY
Genes and genomes

SSECONDARY DOMAINS OF ACTIVITY
Biostatistics; Epigenetics;
GWAS; Human genetics;
Mathematical modelling;
Metabolomics;
Personalized medicine;
Population genetics;
Systems biology;
Transcriptomics

DOMAIN OF APPLICATION

KEY RESOURCES
Quicktest
SSimp

KEY PAPERS IN 2018

Winkler T W *et al.* A joint view on genetic variants for adiposity differentiates subtypes with distinct metabolic implications. Nat Commun.

Rüeger S *et al.* Evaluation and application of summary statistic imputation to discover new height-associated loci. PLoS Gen.

Tin A *et al.* Large-scale whole-exome sequencing association studies identify rare functional variants influencing serum urate levels. Nat Commun.

Anna-Sapfo MALASPINAS

GROUP NAME **EVOLUTIONARY GENOMICS GROUP**
INSTITUTION UNIVERSITY OF LAUSANNE
CITY LAUSANNE

Ancient and modern DNA research have both entered the genomics era. At the Evolutionary Genomics Group (EGG), we aim to characterize evolutionary processes (genetic drift, natural selection, migration and mutation) while relying on genomics data from both modern and ancient samples. We develop analytical and computational methods to analyse and interpret time-sampled data and we apply those methods to novel ancient DNA datasets via collaborative projects. Our work should allow us to quantify and time adaptive and migration events, notably related to the human colonization of the world.

MAIN DOMAIN OF ACTIVITY
Genes and genomes

SECONDARY DOMAINS OF ACTIVITY
Population genetics;
Paleogenomics;
Evolution

DOMAIN OF APPLICATION

Mark ROBINSON

GROUP NAME **STATISTICAL BIOINFORMATICS GROUP**
INSTITUTION UNIVERSITY OF ZURICH
CITY ZURICH

In the Statistical Bioinformatics Group, we develop robust data analysis solutions, including new or improved methods, for the analysis of genome-scale data. We develop statistical methods for interpreting data from high-throughput sequencing and other technologies in the context of genome sequencing, gene expression and regulation and analysis of epigenomes. We are largely data- and problem-driven, and ultimately the methods we develop are geared to the characteristics of the technology platform generating the data. We develop publicly available open-source software tools, generally through the Bioconductor project.

MAIN DOMAIN OF ACTIVITY
Genes and genomes

SECONDARY DOMAINS OF ACTIVITY
Benchmarking;
Biostatistics; Deep sequencing data;
Functional genomics;
Immunology; Next generation sequencing;
Single-cell biology;
Training; Transcriptomics;
Workflows

DOMAINS OF APPLICATION

KEY RESOURCES
conquer
CrispRVariantsLite
DuoClustering2018

Matthew ROBINSON

GROUP NAME **COMPLEX TRAIT GENETICS GROUP**
INSTITUTION UNIVERSITY OF LAUSANNE
CITY LAUSANNE

We develop and apply statistical methodologies to large human phenotype-genotype datasets, to address long-standing questions in population and quantitative genetics. These have improved the prediction accuracy of disease risk in personalized medicine, quantified the contribution of genotype-environment interaction effects to obesity risk, and provided frameworks to examine the genetic basis of phenotypic differences among people. Our current work focuses on improved testing for sex-, age-, or environment-specific genetic effects and investigating the role of interactions between microbial and host genotypes in shaping human phenotypes.

MAIN DOMAIN OF ACTIVITY
Genes and genomes

SECONDARY DOMAINS OF ACTIVITY
Data mining; GWAS;
Machine learning;
Phylogenetic tree analysis (natural selection);
Population genetics;
Software engineering

DOMAIN OF APPLICATION

Michael STADLER

GROUP NAME **FMI COMPUTATIONAL BIOLOGY GROUP**
INSTITUTION FRIEDRICH MIESCHER INSTITUTE
CITY BASEL

We study gene regulation through the analysis and modelling of genome-wide datasets. We collaborate closely with experimental researchers on various biological topics, including cancer progression, chromatin biology and cellular differentiation, using data from various aspects of gene expression, such as DNA methylation and RNA transcription in single-cells, measured by high-throughput sequencing. Our aim is to better understand how the different layers of epigenetic, transcriptional and post-transcriptional regulation interact and contribute to the control of gene expression.

MAIN DOMAINS OF ACTIVITY
Genes and genomes
Core facilities and competence centres

SECONDARY DOMAINS OF ACTIVITY
DNA microarrays;
Epigenetics; Functional genomics; Gene regulation; Oncology;
Single-cell biology;
Software engineering;
Systems biology;
Transcriptomics

DOMAINS OF APPLICATION

KEY RESOURCES
QuasR
MethylSeekR
griph

KEY PAPERS IN 2018

Ostapcuk V *et al.* Activity-dependent neuroprotective protein recruits HP1 and CHD4 to control lineage-specifying genes. Nature

Kumari P *et al.* Evolutionary plasticity of the NHL domain underlies distinct solutions to RNA recognition. Nat Commun.

Fassnacht C *et al.* The CSR-1 endogenous RNAi pathway ensures accurate transcriptional reprogramming during the oocyte-to-embryo transition in *C. elegans*. PLoS Genet.

Andrzej STASIAK

GROUP NAME **DNA AND CHROMOSOME MODELLING GROUP**
INSTITUTION UNIVERSITY OF LAUSANNE
CITY LAUSANNE

Using Brownian dynamics simulations, we model chromosome structure in interphase nuclei of mammalian cells testing the possibility that transcription-induced supercoiling is involved in the formation of TADs. In 2018, we proposed a new model of chromatin loop extrusion, in which cohesin rings are passively pushed by formation of chromatin plectonemes resulting from transcription-induced supercoiling. We also showed by numerical simulations that chromatin loop extrusion is capable of directing DNA knots for unknotting by DNA topoisomerases.

MAIN DOMAIN OF ACTIVITY
Genes and genomes

SECONDARY DOMAINS OF ACTIVITY
Biophysics; Functional genomics; Gene regulation; Human genetics; Mathematical modelling; Structural biology

DOMAIN OF APPLICATION

KEY RESOURCES
KnotProt
Knoto-ID

KEY PAPERS IN 2018

Racko D *et al.*
Transcription-induced supercoiling as the driving force of chromatin loop extrusion during formation of TADs in interphase chromosomes.
Nucleic Acids Res.

Dorier J *et al.*
Knoto-ID: a tool to study the entanglement of open protein chains using the concept of knotoids.
Bioinformatics

Racko D *et al.*
Are TADs supercoiled?
Nucleic Acids Res.

Shinichi SUNAGAWA NEW

GROUP NAME **MICROBIOME RESEARCH GROUP**
INSTITUTION ETH ZURICH
CITY ZURICH

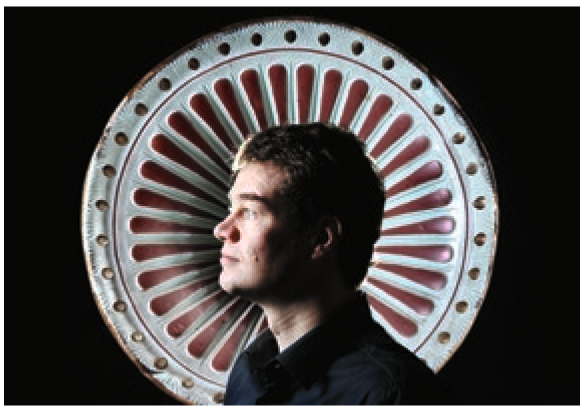
We are interested in studying ecological and evolutionary factors that determine the structure, function and diversity of microbial communities – with a focus on the ocean ecosystem and the gastrointestinal tract of animals and humans. To this end, we develop and combine bioinformatic and experimental approaches to integrate quantitative 'meta-omics' readouts with contextual information, with the goal of better understanding and predicting the role of environmental microorganisms and the underlying mechanisms of host-microbial homeostasis.

MAIN DOMAIN OF ACTIVITY
Genes and genomes

SECONDARY DOMAINS OF ACTIVITY
Evolutionary biology; Experimental biology; Metagenomics; Microbiology; Next-generation sequencing; Personalized medicine; Phylogeny; Population genetics; Software engineering; Systems biology

DOMAINS OF APPLICATION

KEY RESOURCES
mOTU-tool



Erik VAN NIMWEGEN

GROUP NAME **GENOME SYSTEMS BIOLOGY GROUP**
INSTITUTION UNIVERSITY OF BASEL
CITY BASEL

Our main research interest in the Genome Systems Biology (GSB) Group is the study of genome-wide regulatory systems, in order to reconstruct them from high-throughput molecular data, understand and model how they have evolved, and search for design principles in their construction. In particular, we are developing and applying new algorithmic tools for the automated reconstruction of genome-wide regulatory networks from comparative genomic, deep sequencing, and other high-throughput data. In addition, methods are being developed for studying genome evolution, and the evolution of regulatory networks, in particular.

MAIN DOMAIN OF ACTIVITY
Genes and genomes

SECONDARY DOMAINS OF ACTIVITY
Comparative genomics; Population genetics; Epigenetics; Evolution; Genetic networks; Next-generation sequencing; Phylogeny; Systems biology; Systems modelling; Transcriptomics
Gene regulation; Single-cell biology

DOMAINS OF APPLICATION

KEY RESOURCES
SwissRegulon
ISMARA
CRUNCH

Robert WATERHOUSE

GROUP NAME **EVOLUTIONARY-FUNCTIONAL GENOMICS GROUP**
INSTITUTION UNIVERSITY OF LAUSANNE
CITY LAUSANNE

The group's research is focused on elucidating interactions between gene evolution and gene function through developing computational approaches to interrogate evolutionary and functional genomics data. We are developing quantifications of gene evolutionary dynamics and functional properties to build models that link the underlying genetics to observable biological features. Our research projects focus on insects and other arthropods as their countless adaptations to exploit ecological niches mean that they are ideal for investigating how conservation or divergence of functional genomic elements give rise to the splendour of animal biology.

MAIN DOMAIN OF ACTIVITY
Genes and genomes

SECONDARY DOMAINS OF ACTIVITY
Comparative genomics; Database curation; Evolutionary biology; Functional genomics; Phylogenetic tree analysis natural selection; Phylogeny; Population genetics; Software engineering; Systems biology; Transcriptomics

DOMAINS OF APPLICATION

KEY PAPERS IN 2018

Ruzzante L *et al.*
Of genes and genomes: mosquito evolution and diversity.
Trends parasitol.

Robertson H M *et al.*
Genome sequence of the wheat stem sawfly representing an early-branching lineage of the Hymenoptera, illuminates (...)
Genome Bio Evol.

Johnson K P *et al.*
Phylogenomics and the evolution of hemipteroid insects.
PNAS

Evgeny ZDOBNOV

GROUP NAME **COMPUTATIONAL EVOLUTIONARY GENOMICS GROUP**
INSTITUTION UNIVERSITY OF GENEVA
CITY GENEVA

The group is active in the fields of comparative genomics and shotgun metagenomics. We study molecular evolution, develop approaches to genomics data analyses, and implement computational pipelines. We apply evolutionary models to digest sequencing data, and revise these models using the novel data. We study functional genomic elements on the basis of sequence variability among different species and within populations. Our interests range from arthropod genomics, including invertebrate vectors of human pathogens, to the evolution of viruses and clinical microbiology.

MAIN DOMAIN OF ACTIVITY
Genes and genomes

SECONDARY DOMAINS OF ACTIVITY
Benchmarking; Comparative genomics; Evolutionary biology; Functional genomics; Genome reconstruction; Infectious diseases; Metagenomics; Microbiology; Next-generation sequencing; Phylogeny

DOMAINS OF APPLICATION

KEY RESOURCES
OrthoDB
(part of SwissOrthology)
BUSCO
LEMMI

KEY PAPERS IN 2018

Kriventseva EV *et al.*
OrthoDB v10: sampling the diversity of animal, plant, fungal, protist, bacterial and viral genomes for evolutionary and functional (...)
Nucleic Acids Res.

Seppey M *et al.*
BUSCO: Assessing Genome Assembly and Annotation Completeness.
Meth Mol Bio: Eukaryotic Gene Prediction

Rands CM *et al.*
ACI-1 beta-lactamase is widespread across human gut microbiomes in Negativicutes due to transposons harboured by tailed prophages.
Environ Microbiol.

Christian AHRENS

GROUP NAME **BIOINFORMATICS AND PROTEOGENOMICS GROUP**
INSTITUTION AGROSCOPE
CITY WÄDENSWIL

Our research revolves around the bioinformatic integration and analysis of datasets from state-of-the-art omics technologies, which we obtain through close collaboration with experimental biologists. These datasets include genome sequences, gene and protein expression, as well as metabolomics data. One focus is to exploit the unique advantages of proteomics data, including strategies to identify all proteins encoded in a genome (proteogenomics). Another focus is to study the role of microbiomes – e.g. for plant protection – by applying metagenomic, comparative genomic and transcriptomic approaches and integration of the resulting datasets.

MAIN DOMAIN OF ACTIVITY
Proteins and proteomes

SECONDARY DOMAINS OF ACTIVITY
Comparative genomics; Data mining; Drug resistance; Functional genomics; Metagenomics; Microbiology; Next generation sequencing; Proteomics; Software engineering; Transcriptomics

DOMAINS OF APPLICATION

KEY RESOURCES
iPtgxDB
PeptideRank
PeptideClassifier

KEY PAPERS IN 2018

Schmid M *et al.*
Pushing the limits of *de novo* genome assembly for complex prokaryotic genomes harboring very long, near identical repeats.
Nucleic Acids Res.

Schmid M *et al.*
Comparative genomics of completely sequenced *Lactobacillus helveticus* genomes provides insights into strain-specific genes (...)
Front Microbiol.

Katja BAERENFALLER

GROUP NAME **MOLECULAR ALLERGOLOGY GROUP**
INSTITUTION SWISS INSTITUTE OF ALLERGY AND ASTHMA RESEARCH (SIAF)
CITY DAVOS

In the Molecular Allergology group we are aiming to unravel the molecular basis of allergic diseases, antigen presentation, immunotherapies, immunotolerance and tolerance breakage with a combination of functional genomics techniques, including proteomics, peptidomics and immunopeptidomics. The acquired large-scale molecular data are analysed with the aim of identifying molecular markers with which risk factors, different disease phenotypes, their underlying pathogenic mechanisms, and different patient response groups can be distinguished.

MAIN DOMAIN OF ACTIVITY
Proteins and proteomes

SECONDARY DOMAINS OF ACTIVITY
Biomarkers; Functional genomics; Gene regulation; Immunology; Peptidomics; Proteomics; Systems biology

DOMAIN OF APPLICATION

KEY RESOURCES
pep
pro database

KEY PAPER IN 2018

Seaton DD *et al.*
Photoperiodic control of the *Arabidopsis* proteome reveals a translational (...)
Mol Syst Biol.

Amos BAIROCH
Lydie LANE

GROUP NAME **COMPUTER AND LABORATORY INVESTIGATION OF PROTEINS OF HUMAN ORIGINS (CALIPHO)**
INSTITUTION UNIVERSITY OF GENEVA
CITY GENEVA

We aim to use a combination of bioinformatics and experimental methodologies to increase knowledge about the function of the 20,000 protein-coding genes in the human genome. Our main mission is the development of neXtProt, a human protein knowledge resource. We also annotate the effects of human protein variations in the context of cancers and genetic diseases. We are part of the HUPO Human Protein Project, which aims to validate the existence of all predicted human proteins in biological samples by mass spectrometry. We are active in the development of ontologies/standardization resources such as the Cellosaurus and ICEPO.

MAIN DOMAIN OF ACTIVITY
Proteins and proteomes

SECONDARY DOMAINS OF ACTIVITY
API; Data management; Data mining; Data visualization; Database curation; Ontology; Proteomics; Semantic web format

DOMAIN OF APPLICATION

KEY RESOURCES
neXtProt
Cellosaurus

KEY PAPERS IN 2018

Duek P *et al.*
Exploring the uncharacterized human proteome using neXtProt.
J Proteome Res.

Cusin I *et al.*
A new bioinformatics tool to help assess the significance of BRCA1 variants.
Hum Genomics

Bairoch A.
The Cellosaurus, a cell-line knowledge resource.
J Biomol Tech.

Alan BRIDGE NEW

GROUP NAME **SWISS-PROT**
INSTITUTION SIB GENEVA
CITY GENEVA

In the Swiss-Prot Group, we develop, annotate and maintain the UniProtKB/Swiss-Prot protein sequence database, the most widely used protein information resource in the world (and an ELIXIR Core Data Resource), as well as a number of other internationally renowned expert-curated knowledge resources. These include the HAMAP and PROSITE databases of protein families and domains, the ENZYME database of enzyme nomenclature, the Rhea database of biochemical reactions, the SwissLipids database of lipid structures and biological knowledge, and the ViralZone portal. The group also participates in the development and maintenance of the ExPASy website.

MAIN DOMAIN OF ACTIVITY
Proteins and proteomes

SECONDARY DOMAINS OF ACTIVITY
Biochemistry; Core facilities and competence centres; Database curation; Lipids; Metabolomics; Ontology; Protein interactions; Proteomics; Semantic web format; Systems biology

DOMAIN OF APPLICATION

KEY RESOURCES
UniProtKB/Swiss-Prot
Rhea
ViralZone

KEY PAPERS IN 2018

Lombardot T *et al.*
Updates in Rhea: SPARQLing biochemical reaction data.
Nucleic Acids Res.

UniProt Consortium
UniProt: a worldwide hub of protein knowledge.
Nucleic Acids Res.

Lee K *et al.*
Scaling up data curation using deep learning: An application to literature triage in genomic (...)
PLoS Comput Biol.

David GFELLER

GROUP NAME **COMPUTATIONAL CANCER BIOLOGY GROUP**
INSTITUTION UNIVERSITY OF LAUSANNE
CITY LAUSANNE

In the Computational Cancer Biology Group, our aim is to study interactions between cancer and immune cells. We are focusing on molecular and cellular aspects of cancer immune cell interactions. At the molecular level, we have developed tools to predict (neo-) antigen presentation from HLA peptidomics datasets and are currently working on TCR-peptide-HLA interaction predictions. At the cellular level, we are developing novel approaches to characterizing immune infiltrations and the different states of immune cells from (single-cell) gene expression data.

MAIN DOMAIN OF ACTIVITY
Proteins and proteomes

SECONDARY DOMAINS OF ACTIVITY
Immunology; Oncology; Protein interactions

DOMAIN OF APPLICATION

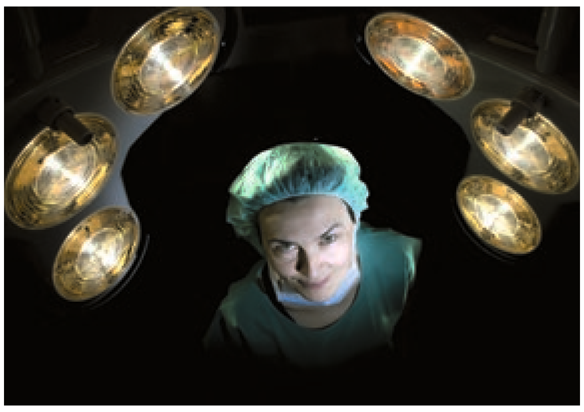
KEY RESOURCES
EPIC
MixMHCp
TILAtlas

KEY PAPERS IN 2018

Guillaume P *et al.*
The C-terminal extension landscape of naturally presented HLA-I ligands.
PNAS

Gfeller D *et al.*
The length distribution and multiple specificity of naturally presented HLA-I ligands.
J Immunol.

Dib L *et al.*
Polymorphic sites preferentially avoid co-evolving residues in MHC class I proteins.
PLoS Comp Bio.



Frédérique LISACEK

GROUP NAME **PROTEOME INFORMATICS GROUP**
INSTITUTION UNIVERSITY OF GENEVA
CITY GENEVA

In the Proteome Informatics Group (PIG), we are involved in software and database development for the benefit of the proteomics and glycomics communities. These resources are made available through the ExPASy server. Software tools support experimental mass spectrometry data analysis, mainly for the detection of posttranslational modifications, with a strong focus on glycosylation. We also develop integrated tools and databases to query knowledge of carbohydrates attached to proteins, carbohydrate structures and protein-carbohydrate interactions.

MAIN DOMAIN OF ACTIVITY
Proteins and proteomes

SECONDARY DOMAINS OF ACTIVITY
Data visualization; Database curation; Glycomics; Protein interactions; Proteomics

DOMAIN OF APPLICATION

KEY RESOURCES
GlyConnect
Glynsight
UniLectin

KEY PAPERS IN 2018

Mariethoz J *et al.*
Glycomics@ExPASy: Bridging the gap
Mol Cell Proteomics

Alocci D *et al.*
Glyconnect: glycoproteomics goes visual, interactive and analytical.
J Prot Res.

Bonnardel F *et al.*
UniLectin3D, a database of carbohydrate binding proteins with curated information on 3D structures and interacting ligands.
Nucleic Acids Res.

Christian VON MERING

GROUP NAME **BIOINFORMATICS / SYSTEMS BIOLOGY GROUP**
INSTITUTION UNIVERSITY OF ZURICH
CITY ZURICH

In the Bioinformatics / Systems Biology Group, we study the dynamics of entire biological systems, both at evolutionary time-scales and at shorter time-scales – down to a few minutes. We often work in close collaboration with laboratory scientists, focusing on the computational aspects of studying such systems, in fields ranging from genetics to genomics and proteomics. In addition, we produce and maintain several online resources for the life science community, including STRING-db (protein networks), EGGNOG-db (gene orthology relations), and PAX-db (protein abundances).

MAIN DOMAIN OF ACTIVITY
Proteins and proteomes

SECONDARY DOMAINS OF ACTIVITY
Comparative genomics; Deep sequencing data; Evolutionary biology; Metagenomics; Protein interactions; Proteomics; Text mining

DOMAIN OF APPLICATION

KEY RESOURCES
STRING-db
PAX-db
microbeatlas

KEY PAPERS IN 2018

Tackmann J *et al.*
Ecologically informed microbial biomarkers and accurate classification of mixed and unmixed samples in an extensive cross-study of human body sites.
Microbiome

Daga N *et al.*
Growth-restricting effects of siRNA transfections: a largely deterministic combination of off-target binding and hybridization-independent competition.
Nucleic Acids Res.

Huerta-Cepas J *et al.*
eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5,090 organisms and 2,502 viruses.
Nucleic Acids Res.

Maria ANISIMOVA

GROUP NAME **APPLIED COMPUTATIONAL GENOMICS TEAM**
INSTITUTION ZHAW
CITY WÄDENSWIL

The Applied Computational Genomics Team focuses on theoretical and computational aspects of modelling the process of genome evolution and adaptive change. With the growing size and complexity of molecular data, we strive to keep pace by providing accurate, scalable and practical computational solutions that enable a wide range of scientists to analyse patterns of evolution and natural selection in large genomic and omics data. We develop bioinformatic methods for real applications, ranging from biotechnology to biomedical research, ecology and agriculture.

MAIN DOMAIN OF ACTIVITY
Evolution and phylogeny

SECONDARY DOMAINS OF ACTIVITY
Benchmarking; Biomarkers; Comparative genomics; Data mining; Evolutionary biology; Genome reconstruction; Machine learning; Phylogenetic tree analysis natural selection; Phylogeny; Semantic web format

DOMAINS OF APPLICATION

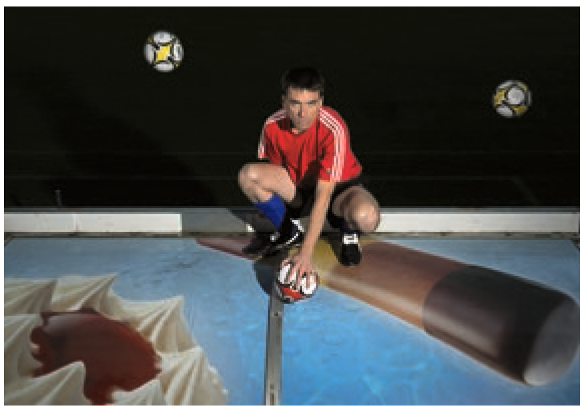
KEY RESOURCES
TRAL
CodonPhyML and Castor
ProGraph and ProPIP

KEY PAPERS IN 2018

Maiolo M *et al.* Progressive multiple sequence alignment with indel evolution. BMC Bioinfo.

Mattle-Greminger M P *et al.* Genomes reveal marked differences in the adaptive evolution between orangutan species. Genome Biol.

Garcia V *et al.* Accounting for programmed ribosomal frameshifting in the computation of codon usage bias indices. G3 (Bethesda)



Niko BEERENWINKEL

GROUP NAME **COMPUTATIONAL BIOLOGY GROUP**
INSTITUTION ETH ZURICH, D-BSSE
CITY ZURICH & BASEL

The Computational Biology Group is involved in research and teaching in the areas of computational biology, biostatistics, and systems biology. We aim to support the rational design of medical interventions in complex and rapidly evolving systems. To achieve this goal, we develop models and algorithms for the statistical analysis of high-throughput molecular data, we analyse biological networks and predict the effect of perturbations, and we design evolutionary models of rapidly adapting disease-causing agents. We are engaged in several personalized medicine efforts, particularly in oncology and virology.

MAIN DOMAIN OF ACTIVITY
Evolution and phylogeny

SECONDARY DOMAINS OF ACTIVITY
Biostatistics; Deep sequencing data; Infectious diseases; Next-generation sequencing; Oncology; Personalized medicine; Phylogeny; Systems biology; Workflows

DOMAIN OF APPLICATION

KEY RESOURCE
V-pipe

KEY PAPERS IN 2018

Kuipers J *et al.* Mutational interactions define novel cancer subgroups. Nat Commun.

Singer J *et al.* SCIΦ: Single-cell mutation identification via phylogenetic inference. Nat Commun.

Pirkl M and Beerenwinkel N Single-cell network analysis with a mixture of Nested Effects Models.Bioinformatics

Christophe DESSIMOZ

GROUP NAME **LABORATORY OF COMPUTATIONAL EVOLUTIONARY BIOLOGY**
INSTITUTION UNIVERSITY OF LAUSANNE
CITY LAUSANNE

At the interface of biology and computer science, we seek to better understand evolutionary and functional relationships between genes, genomes and species. Key underlying questions are: 1) How can we extrapolate current biological knowledge, concentrated in a few model organisms, to the rest of life? 2) Conversely, how can we exploit the wealth and diversity of life to better grasp an organism of interest? 3) Can we summarize the evolutionary history of species as a sparse mixture of tree topologies? Our activities are divided between bioinformatics methods and resource development, and their application – typically with experimentalists.

MAIN DOMAIN OF ACTIVITY
Evolution and phylogeny

SECONDARY DOMAINS OF ACTIVITY
Benchmarking; Comparative genomics; Evolutionary biology; Genome reconstruction; Machine learning; Mathematical modelling; Ontology; Phylogenetic tree analysis natural selection; Phylogeny; Semantic web format

DOMAINS OF APPLICATION

KEY RESOURCES
OMA
(part of SwissOrthology)
Phylo.io
ALF

KEY PAPERS IN 2018

Altenhoff A M *et al.* The OMA orthology database in 2018: retrieving evolutionary relationships among all domains of life through (...). Nucleic Acids Res.

wWarwick Vesztrocy A *et al.* Prioritising candidate genes causing QTL using hierarchical orthologous groups. Bioinformatics

Piližota I *et al.* Phylogenetic approaches to identifying fragments of the same gene, with application to the wheat genome. Bioinformatics

Laurent EXCOFFIER

GROUP NAME **COMPUTATIONAL AND MOLECULAR POPULATION GENETICS GROUP**
INSTITUTION UNIVERSITY OF BERN
CITY BERN

In the Computational and Molecular Population Genetics (CMPG) Group, we develop new methodologies for the simulation and analysis of molecular polymorphisms within species, with a particular focus on humans. We also maintain the Arlequin software, a popular package for the analysis of multi-locus genetic diversity within and between populations, as well as statistical methods to reconstruct and infer evolutionary processes from genomic data (i.e. using the fastsimcoal2 program). The team focuses on the effect of range expansions on genomic and functional diversity, and the detection of signatures of adaptation and selection at the molecular level.

MAIN DOMAIN OF ACTIVITY
Evolution and phylogeny

SECONDARY DOMAINS OF ACTIVITY
Comparative genomics; Evolutionary biology; Human genetics; Mathematical modelling; Paleogenomics; Population genetics

DOMAINS OF APPLICATION

KEY RESOURCES
Arlequin
fastsimcoal
PGDSpider

KEY PAPERS IN 2018

Pouyet F *et al.* Background selection and biased gene conversion affect more than 95% of the human genome and bias (...). eLife

Gilbert KJ *et al.* Mutation load dynamics during environmentally-driven range shifts. PLoS Genetics

Peischl S *et al.* Relaxed selection during a recent human expansion. Genetics

Gaston GONNET

GROUP NAME **COMPUTATIONAL BIOCHEMISTRY RESEARCH GROUP**
INSTITUTION ETH ZURICH
CITY ZURICH

In the Computational Biochemistry Research Group, we are interested in the modelling and analysis of biological problems at the molecular level. In particular, our expertise lies in searching algorithms, optimizing algorithms, mathematical modelling, and computational systems. Most of our research efforts are concentrated on the Orthologous MAtrix (OMA) project. This particular project aims to produce, automatically, reliable orthologous groups of proteins that are derived from entire genomes.

MAIN DOMAIN OF ACTIVITY
Evolution and phylogeny

SECONDARY DOMAINS OF ACTIVITY
API; Benchmarking; Comparative genomics; Machine learning; Mathematical modelling; Phylogeny; Semantic web format; Software engineering

DOMAINS OF APPLICATION

KEY RESOURCES
OMA
(part of SwissOrthology)
Orthology Benchmarking Service

KEY PAPERS IN 2018

Altenhoff A M *et al.* The OMA orthology database in 2018: retrieving evolutionary relationships among all domains of life (...) Nucleic Acids Res.

Train C M *et al.* iHam & pyHAM: visualizing and processing hierarchical orthologous groups. Bioinformatics

Antelo V *et al.* First release of the bacterial biobank of the urban environment (BBUE). Microbiol Res Announc.

Jérôme GOUDET

GROUP NAME **POPULATION GENETICS AND GENOMICS GROUP**
INSTITUTION UNIVERSITY OF LAUSANNE
CITY LAUSANNE

Our group's interest is focused on understanding how the interplay of population structure, trait architecture and selection can be disentangled. To this end, we use different approaches, from theory and the development of statistical tools to field observations. The main biological models currently used are humans and the barn owl. On the theoretical side, we investigate the dynamics of multilocus genetic systems under the influence of selection, migration and drift, and develop comprehensive individual-based models as well as statistical methods to infer selection, mating systems and population structure.

MAIN DOMAIN OF ACTIVITY
Evolution and phylogeny

SECONDARY DOMAINS OF ACTIVITY
Biostatistics; Data mining; Evolutionary biology; Mathematical modelling; Next-generation sequencing; Population genetics; Training

DOMAIN OF APPLICATION

KEY RESOURCES
quantiNemo
Hierfstat

KEY PAPERS IN 2018

Goudet J *et al.* How to estimate kinship. Mol Ecol.

Kanitz R *et al.* Complex genetic patterns in human arise from a simple range-expansion model over continental landmasses. Plos One

Neuenschwander S *et al.* QuantiNemo 2: a Swiss knife to simulate complex demographic and genetic scenarios, forward and backward in time. Bioinformatics

Sara MITRI

GROUP NAME **EVOLUTIONARY MICROBIOLOGY GROUP**
INSTITUTION UNIVERSITY OF LAUSANNE
CITY LAUSANNE

Our group focuses on studying social interactions in microbial communities and how they evolve. We follow an interdisciplinary approach, where we quantify and characterize interactions within small bacterial communities in the lab, and then use computational and mathematical models to explain and predict how these interactions will change over ecological and evolutionary time-scales.

MAIN DOMAIN OF ACTIVITY
Evolution and phylogeny

SECONDARY DOMAINS OF ACTIVITY
Deep sequencing data; Evolutionary biology; Experimental biology; Mathematical modelling; Microbiology; Synthetic biology; Systems biology

DOMAINS OF APPLICATION

KEY PAPER IN 2018

Frost I *et al.*
Cooperation, competition and antibiotic resistance in bacterial colonies.
ISME J

Richard NEHER

GROUP NAME **MICROBIAL EVOLUTION GROUP**
INSTITUTION UNIVERSITY OF BASEL
CITY BASEL

Our goal is to understand how evolution works at the molecular level and how organisms adapt to changing conditions through random mutations and recombination. Pathogens provide excellent model systems for studying these processes. We apply modern sequencing techniques to decipher the genome of thousands of HIV particles and develop new algorithms to elucidate the interactions between HIV and the immune system. We have developed a method that can predict the evolution of influenza viruses based on their phylogenetic tree (nextflu.org). Such predictions can help to ensure that the seasonal influenza vaccine matches the viruses in circulation.

MAIN DOMAIN OF ACTIVITY
Evolution and phylogeny

SECONDARY DOMAINS OF ACTIVITY
Data visualization; Deep sequencing data; Drug resistance; Evolutionary biology; Infectious diseases; Microbiology; Phylogeny; Population genetics

DOMAIN OF APPLICATION

KEY RESOURCES
nexstrain
pangenome
treetime

KEY PAPERS IN 2018

Hadfield J *et al.*
Nextstrain: real-time tracking of pathogen evolution.
Bioinformatics

Neher R A and Bedford T.
Real-time analysis and visualization of pathogen sequence data.
J Clin Mic.

Sagulenko P *et al.*
TreeTime: Maximum-likelihood phylodynamic analysis.
Virus Evol.

Joshua PAYNE

GROUP NAME **COMPUTATIONAL BIOLOGY GROUP**
INSTITUTION ETH ZURICH
CITY ZURICH

We are interested in understanding the design constraints, robustness, and evolution of gene regulatory systems, particularly at the level of transcription, using both modeling and data-driven approaches.

MAIN DOMAIN OF ACTIVITY
Evolution and phylogeny

SECONDARY DOMAINS OF ACTIVITY
Evolutionary biology; Functional genomics; Gene regulation; Gene regulatory network analysis; Systems biology

DOMAIN OF APPLICATION

KEY RESOURCE
Genonets

KEY PAPERS IN 2018

Payne J L *et al.*
RNA-mediated gene regulation is less evolvable than transcriptional regulation.
PNAS

Aguilar-Rodriguez J *et al.*
The architecture of an empirical genotype-phenotype map.
Evolution

Marc ROBINSON-RECHAVI

GROUP NAME **EVOLUTIONARY BIOINFORMATICS GROUP**
INSTITUTION UNIVERSITY OF LAUSANNE
CITY LAUSANNE

In the Evolutionary Bioinformatics Group, we are mainly concerned with determining the role of evolutionary innovation and constraint in animals. For this, we develop methods and databases to extract reliable information from genome and transcriptome data. These include Bgee, a database for gene expression evolution, and Selectome, a database of positive selection. While developing these resources, we also conduct research on ontologies, biocuration, and high-performance computing. Our biological focus is to link Evo-Devo with phylogenomics. Notably, we study the role of gene duplication in the divergence between genes and between species.

MAIN DOMAIN OF ACTIVITY
Evolution and phylogeny

SECONDARY DOMAINS OF ACTIVITY
Comparative genomics; Database curation; Evo-Devo; Evolutionary biology; Gene regulation; Next-generation sequencing; Ontology; Phylogenetic tree analysis (natural selection); Training; Transcriptomics

DOMAIN OF APPLICATION

KEY RESOURCES
Bgee
Selectome

KEY PAPERS IN 2018

Komljenovic A *et al.*
BgeeDB, an R package for retrieval of curated expression datasets and for gene list expression localization enrichment tests.
F1000Res.

Liu J *et al.*
Adaptive evolution of animal proteins over development: support for the Darwin selection opportunity hypothesis of Evo-Devo.
Mol Biol Evol.

Liu J and Robinson-Rechavi M.
Developmental constraints on genomeevolution in four bilaterian model species.
Genome Biol Evol.

Nicolas SALAMIN

GROUP NAME **COMPUTATIONAL PHYLOGENETICS GROUP**
INSTITUTION UNIVERSITY OF LAUSANNE
CITY LAUSANNE

We develop models and software to better understand the evolution of organisms and to test macroevolutionary hypotheses. We are looking at the ecological, genomic and morphological factors that constrain speciation and adaptation. We focus on phylogenetic methods, clownfish and plant genomics, the estimation of positive selection, modelling the evolution of DNA sequences and phenotypes, the mode and tempo of species evolution and the spatially explicit evolution of diversity. Our aim is to develop better models to analyse sequence data and quantitative models to estimate macroevolutionary patterns and processes.

MAIN DOMAIN OF ACTIVITY
Evolution and phylogeny

SECONDARY DOMAINS OF ACTIVITY
Comparative genomics; Evolution; Functional genomics; Machine learning; Next-generation sequencing; Phylogeny; Selection; Software development; Transcriptomics

DOMAINS OF APPLICATION

KEY RESOURCE
coevweb

Tanja STADLER

GROUP NAME **COMPUTATIONAL EVOLUTION GROUP**
INSTITUTION ETH ZURICH, D-BSSE
CITY ZURICH, BASEL

In the Computational Evolution Group, we develop phylogenetic tools in order to understand evolutionary processes. Using our phylogenetic methods, we aim to improve our understanding of past evolutionary and population dynamic processes on different scales. We address questions in a number of fields, focusing on epidemiology, public health and medicine, ecology and evolution, and language evolution. In our daily work, we define and analyse stochastic models, implement computational methods, analyse empirical data, and discuss our new insights with clinicians, and public-health policy makers, as well as ecologists and palaeontologists.

MAIN DOMAIN OF ACTIVITY
Evolution and phylogeny

SECONDARY DOMAINS OF ACTIVITY
Biostatistics; Evolutionary biology; Infectious diseases; Mathematical modelling; Next-generation sequencing; Phylogenetic tree analysis (natural selection); Phylogeny; Single-cell biology; Software engineering

DOMAINS OF APPLICATION

KEY RESOURCE
BEAST2

KEY PAPERS IN 2018

Moeller S *et al.*
The impact of the tree prior on estimating clock rates during epidemic outbreaks.
PNAS

Stadler T *et al.*
The fossilized birth-death model for the analysis of stratigraphic range data under different speciation concepts.
J Theo Biol.

Silvestro D *et al.*
Closing the gap between palaeontological and neontological speciation and extinction rate estimates.
Nat Commun.

Andreas WAGNER

GROUP NAME **EVOLUTIONARY SYSTEMS BIOLOGY GROUP**
INSTITUTION UNIVERSITY OF ZURICH
CITY ZURICH

In the Evolutionary Systems Biology Group, we study the evolution and evolvability of biological systems at all levels of biological organization, from genes and genomes to biological networks and whole organisms. We develop bioinformatics tools to integrate data from a variety of sources, including comparative whole-genome sequence data, microarray expression data, and high-throughput protein interaction data. Our work uses comparative analysis of genomic data, laboratory evolution experiments and mathematical modelling.

MAIN DOMAIN OF ACTIVITY
Evolution and phylogeny

SECONDARY DOMAINS OF ACTIVITY
Comparative genomics; Deep sequencing data; Experimental biology; Gene regulatory network analysis; Population genetics; Systems biology

DOMAIN OF APPLICATION

KEY PAPERS IN 2018

Schaerli Y *et al.* Synthetic circuits reveal how mechanisms of gene regulatory networks constrain evolution. Mol Syst Biol.

Hosseini S R and Wagner A. Genomic organization underlying deletional robustness in bacterial metabolic systems. PNAS

Sprouffsky K *et al.* High mutation rates limit evolutionary adaptation in *Escherichia coli.* PLoS Genet.

Daniel WEGMANN

GROUP NAME **STATISTICAL AND COMPUTATIONAL EVOLUTIONARY BIOLOGY GROUP**
INSTITUTION UNIVERSITY OF FRIBOURG
CITY FRIBOURG

When observing nature, it is easy to be impressed by the huge diversity seen on any biological scale. Our primary aim is to better understand the underlying evolutionary and ecological processes that have been shaping this diversity over the course of evolution on our planet. To achieve this, we design and evaluate new statistical and computational approaches to infer complex evolutionary histories. For this, we develop and apply machine learning algorithms, with a particular focus on likelihood-free methods. We then apply these approaches to the wealth of data currently being generated.

MAIN DOMAIN OF ACTIVITY
Evolution and phylogeny

SECONDARY DOMAINS OF ACTIVITY
Biostatistics; Evolutionary biology; GWAS; Human genetics; Machine learning; Mathematical modelling; Next-generation sequencing; Paleogenomics; Population genetics

DOMAINS OF APPLICATION

KEY RESOURCES
ABCtoolbox
Atlas
RASPberry

KEY PAPERS IN 2018

Scheib C *et al.* Ancient human parallel lineages within North America contributed to a coastal expansion. Science

Veeramah K *et al.* Population genomic analysis of elongated skulls reveals extensive female-biased immigration in early medieval Bavaria. PNAS

Kousathanas A *et al.* A guide to general-purpose ABC software. Handbook of Approximate Bayesian Computation

Andrea CAVALLI NEW

GROUP NAME **COMPUTATIONAL STRUCTURAL BIOLOGY**
INSTITUTION UNIVERSITÀ DELLA SVIZZERA ITALIANA
CITY BELLINZONA

The overall objective of our research is to understand the role that structure and dynamics play in the definition of the function of biomolecules. In order to perform their function, proteins and other biological molecules undergo a series of conformational changes that requires a precise balance between flexibility and stability. Changes in this equilibrium induced by genetic mutations are often at the origin of diseases. Our aim is to develop methods to incorporate data on structure an dynamics into computer simulations to study of complex processes such as molecular recognition, protein misfolding and aggregation at an atomistic level.

MAIN DOMAIN OF ACTIVITY
Structural biology

SECONDARY DOMAINS OF ACTIVITY
Drug design; Molecular dynamics; Structural biology

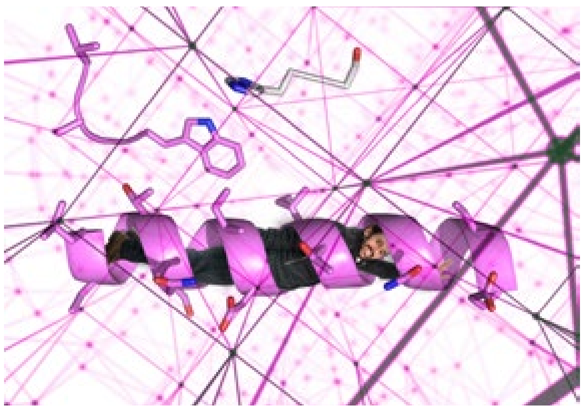
DOMAIN OF APPLICATION

KEY RESOURCE
ALMOST

KEY PAPERS IN 2018

Kassem M M *et al.* Enhancing coevolution-based contact prediction by imposing structural self-consistency of the contacts. Sci Rep.

Sgrignani J *et al.* How phosphorylation influences E1 subunit pyruvate dehydrogenase: A computational study. Sci Rep.



Bruno CORREIA

GROUP NAME **LABORATORY OF PROTEIN DESIGN AND IMMUNOENGINEERING**
INSTITUTION EPFL
CITY LAUSANNE

The Laboratory of Protein Design & Immunoengineering started in March 2015. We have established both the computational and experimental branches of our research.We currently perform our computational calculations at the High Performance Computing facility at the EPFL. The experimental arm of the laboratory is prepared to perform molecular biology, protein expression and purification, and protein biophysical characterization.

MAIN DOMAIN OF ACTIVITY
Structural biology

SECONDARY DOMAINS OF ACTIVITY
Biochemistry; Immunology; Infectious diseases; Vaccines

DOMAIN OF APPLICATION

KEY RESOURCES
Rosetta Fold From Loops
TopoBuilder
SEEDER

Matteo DAL PERARO

GROUP NAME **LABORATORY FOR BIOMOLECULAR MODELLING**
INSTITUTION EPFL
CITY LAUSANNE

Our main goal at the Laboratory for Biomolecular Modelling is to understand the physical and chemical properties of complex biological systems, in particular their function with regard to structure and dynamics. To this end, we use and develop a broad spectrum of computational tools, fully integrated with experimental data. Multiscale simulations and dynamic integrative modelling are used to investigate the function of molecular assemblies, mimicking the conditions of the native cellular environment.

MAIN DOMAIN OF ACTIVITY
Structural biology

SECONDARY DOMAINS OF ACTIVITY
Biochemistry; Drug resistance; Lipids; Molecular dynamics; Software development

DOMAINS OF APPLICATION

KEY RESOURCES
pow^{er}
LipidBuilder

Olivier MICHIELIN
Vincent ZOETE

GROUP NAME **MOLECULAR MODELLING GROUP**
INSTITUTION UNIVERSITY OF LAUSANNE
CITY LAUSANNE

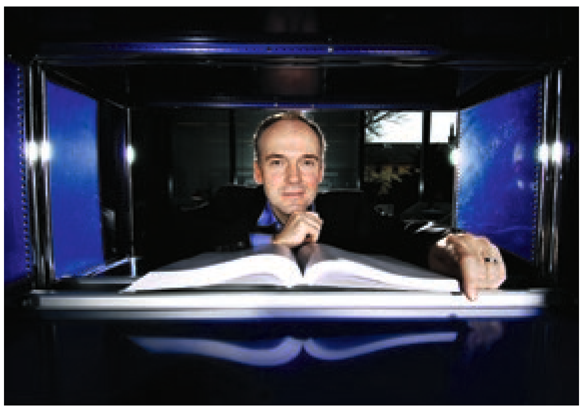
The Molecular Modelling Group studies molecular interactions using modelling techniques such as homology modelling, molecular dynamics, docking and free energy simulations. Our main activity consists in developing and applying methods for computer-aided protein engineering and drug design. Notably, we develop novel inhibitors of important targets for cancer therapy and design optimized proteins, such as T-cell receptors, for cancer immunotherapy. We develop and maintain web tools for drug design, such as SwissDock, SwissTargetPrediction and SwissADME. We also run the Protein Modelling Facility of the University of Lausanne.

MAIN DOMAIN OF ACTIVITY
Structural biology

SECONDARY DOMAINS OF ACTIVITY
Drug design; Immunology; Molecular dynamics; Oncology; Personalized medicine; Software development; Systems biology

DOMAIN OF APPLICATION

KEY RESOURCES
SwissDock
SwissTargetPrediction
SwissADME



Torsten SCHWEDE

GROUP NAME **COMPUTATIONAL STRUCTURAL BIOLOGY GROUP**
INSTITUTION UNIVERSITY OF BASEL
CITY BASEL

In the Computational Structural Biology (CSB) Group, we focus on the development of methods and algorithms to model, simulate and analyse three-dimensional protein structures and their molecular properties in order to apply these techniques to the understanding of biological processes at a molecular level. Our main emphasis is on homology modelling – using evolutionary information to model protein tertiary and quaternary structures. Applications in biomedical research include the study of protein-ligand interactions, drug discovery, structure-guided protein engineering, and the interpretation of disease-causing mutations.

MAIN DOMAIN OF ACTIVITY
Structural biology

SECONDARY DOMAINS OF ACTIVITY
Benchmarking; Biophysics; Data management; Drug design; Drug resistance; Homology modelling; Protein engineering; Structural biology; Workflows

DOMAINS OF APPLICATION

KEY RESOURCES
SWISS-MODEL
CAMEO
Model Archive

KEY PAPERS IN 2018

Waterhouse A *et al.*
SWISS-MODEL: homology modelling of protein structures and complexes.
Nucleic acids res.

Haas J *et al.*
Continuous Automated Model Evaluation (CAMEO) complementing the Critical Assessment of Structure Prediction in CASP12.
Proteins

Studer G *et al.*
Modeling of protein tertiary and quaternary structures based on evolutionary information.
Methods Mol Biol.

Vincent ZOETE NEW

GROUP NAME **COMPUTER-AIDED MOLECULAR ENGINEERING GROUP**
INSTITUTION UNIVERSITY OF LAUSANNE
CITY LAUSANNE

The laboratory specializes in the development of computer-aided algorithms, programs and databases for protein engineering and drug design, with applications in oncology, notably in immunotherapy of cancer. We also provide support in Molecular Modelling for the Molecular Tumour Board of the Lausanne University Hospital.

MAIN DOMAIN OF ACTIVITY
Structural biology

SECONDARY DOMAINS OF ACTIVITY
Drug design; Immunology; Molecular dynamics; Oncology

DOMAIN OF APPLICATION

Bastien CHOPARD

GROUP NAME **SCIENTIFIC AND PARALLEL COMPUTING GROUP**
INSTITUTION UNIVERSITY OF GENEVA
CITY GENEVA

In the Scientific and Parallel Computing (SPC) Group, we develop new algorithms and methods to better understand and/or predict various phenomena in biology. We focus on multiscale modelling and computing, high-performance computing, cellular automata, lattice Botzmann methods, multi-agent systems, and optimizing techniques and machine learning. A core activity of our group is the modelling and simulation of complex systems. In bioinformatics, the objective is to develop advanced numerical methodologies to model biological processes.

MAIN DOMAIN OF ACTIVITY
Systems biology

SECONDARY DOMAINS OF ACTIVITY
Data mining; Evo-Devo; Proteomics; Software engineering

DOMAINS OF APPLICATION

KEY RESOURCE
Terminus

Giovanni CIRIELLO

GROUP NAME **COMPUTATIONAL SYSTEMS ONCOLOGY GROUP**
INSTITUTION UNIVERSITY OF LAUSANNE
CITY LAUSANNE

The Computational Systems Oncology lab integrates algorithmic design, numerical modelling, and molecular biology approaches to address relevant questions in cancer biology and therapeutics. We explore single and combinations of genetic and epigenetic alterations that emerge and are selected during cancer evolution, to understand cancer driving mechanisms and inform precision medicine approaches.

MAIN DOMAIN OF ACTIVITY
Systems biology

SECONDARY DOMAINS OF ACTIVITY
Biostatistics; Epigenetics; Evolutionary biology; Mathematical modelling; Oncology; Personalized medicine

DOMAINS OF APPLICATION

KEY RESOURCES
SELECT
RESET

KEY PAPERS IN 2018

Sanchez-Vega F *et al.*
Oncogenic Signalling pathways in The Cancer Genome Atlas.
Cell

Raynaud F *et al.*
Pan-cancer inference of intra-tumour heterogeneity reveals associations with different forms of genomic instability.
PLoS Genet.

Saghafinia S *et al.*
Pan-cancer landscape of aberrant DNA methylation across human tumours.
Cell Rep.

Manfred CLAASSEN

GROUP NAME **COMPUTATIONAL SINGLE CELL BIOLOGY GROUP**
INSTITUTION ETH ZURICH
CITY ZURICH

Our research in the Computational Single Cell Biology Group aims to elucidate the composition of heterogeneous cell populations and how these implement function in the context of cancer and immune biology. To accomplish this task, we build on concepts from statistics, machine learning and mathematical optimization to develop probabilistic approaches to describe biological systems, learn these descriptions from data, and design experiments to validate hypotheses following computational analyses. Our research can be used to pinpoint therapeutic targets with a view to designing drugs.

MAIN DOMAIN OF ACTIVITY
Systems biology

SECONDARY DOMAINS OF ACTIVITY
Gene regulatory network analysis; Immunology; Machine learning; Mathematical modelling; Oncology; Personalized medicine; Signalling pathways; Single-cell biology; Transcriptomics

DOMAINS OF APPLICATION

KEY RESOURCES
CellCnn
Reactionet lasso
TreeTop

KEY PAPER IN 2018

Arvaniti E *et al.*
Automated Gleason grading of prostate cancer tissue microarrays via deep learning.
Sci Rep.

Dagmar IBER

GROUP NAME **COMPUTATIONAL BIOLOGY GROUP**
INSTITUTION ETH ZURICH, D-BSSE
CITY ZURICH, BASEL

The Computational Biology Group (CoBi) develops computational models of developmental processes. We place a particular focus on mechanistic 4D image-based *in silico* models of organogenesis (mouse lung, kidney, pancreas, brain, Drosophila wing and eye) and on the delineation of fundamental mechanisms, such as those that restrict the size of organs and those that maintain the proportions of structures in different-sized embryos. The group runs a wet lab to obtain image data, and collaborates with tissue engineers to build spatially-organized tissue from stem cells, as well as with clinicians to apply its techniques to disease models.

MAIN DOMAIN OF ACTIVITY
Systems biology

SECONDARY DOMAINS OF ACTIVITY
Biophysics; Cell mechanobiology; Deep sequencing data; Evo-Devo; Experimental biology; Gene regulatory network analysis; *In silico* organogenesis; Mathematical modelling; Signalling pathways; Single-cell biology

DOMAINS OF APPLICATION

KEY RESOURCE
LBIBCell

KEY PAPERS IN 2018

Boareto M *et al.*
Positional information encoded in the dynamic differences between neighbouring oscillators during vertebrate segmentation.
BioRxiv

Menshykau D *et al.*
Image-based modeling of kidney branching morphogenesis reveals GDNF-RET based turing-type mechanism and pattern-modulating (...)
Nat Commun.

Robert IVANEK

GROUP NAME **DBM BIOINFORMATICS CORE FACILITY**
INSTITUTION UNIVERSITY OF BASEL AND UNIVERSITY HOSPITAL BASEL
CITY BASEL

Our group is located in the Department of Biomedicine (DBM) at the University of Basel. We collaborate with scientists from DBM on projects covering a broad spectrum of research topics, from cellular differentiation and evolutionary biology to the identification of a molecular basis for various human diseases. To do this, we focus on the analysis, interpretation and visualization of expression, epigenetic and genomic data, which are mainly derived from next-generation sequencing experiments.

MAIN DOMAINS OF ACTIVITY
Systems biology
Core facilities and competence centres

SECONDARY DOMAINS OF ACTIVITY
Biostatistics; Epigenetics; Functional genomics; Single-cell biology

DOMAIN OF APPLICATION

KEY RESOURCE
Gviz

KEY PAPERS IN 2018

Rodrigues P F *et al.*
Distinct progenitor lineages contribute to the heterogeneity of plasmacytoid dendritic cells.
Nature immunol.

Drobek A *et al.*
Strong homeostatic TCR signals induce formation of self-tolerant virtual memory CD8 T cells.
EMBO J.

Christian MAZZA

GROUP NAME **BIOMATHEMATICS AND COMPUTATIONAL BIOLOGY GROUP**
INSTITUTION UNIVERSITY OF FRIBOURG
CITY FRIBOURG

Mathematical modelling is becoming more and more instrumental in life sciences; the data complexity and the high number of interacting components, from molecules to animals, render intuitive reasoning very difficult. The idea consists in formulating mathematically how certain biological units affect each other and how these interactions affect the whole system. Such quantitative approaches produce mathematical models, which are then implemented using computer software. We study phyllotaxis by trying, for example, to explain the formation of geometrically regular patterns in plants, such as spirals in sunflowers.

MAIN DOMAIN OF ACTIVITY
Systems biology

SECONDARY DOMAINS OF ACTIVITY
Biostatistics; Evolutionary biology; Gene regulatory network analysis; Signalling pathways; Synthetic biology

DOMAINS OF APPLICATION

Michel MILINKOVITCH

GROUP NAME **ARTIFICIAL & NATURAL EVOLUTIONARY DEVELOPMENT OF COMPLEXITY GROUP**
INSTITUTION UNIVERSITY OF GENEVA
CITY GENEVA

We combine Evolutionary and Developmental Biology as well as genomic approaches with the study of physical processes to understand the mechanisms generating life's complexity and diversity. The self-organizational capabilities of tissues are pertinent to Evo-Devo at multiple spatial scales. We aim to understand the interactions between physical (e.g., mechanics, reaction diffusion) and biological (e.g., cell signalling, proliferation, migration) parameters, which generate patterns and shapes during development. Our multidisciplinary team includes Evolutionary and Developmental Biologists, Computer Scientists, Engineers, and Physicists.

MAIN DOMAIN OF ACTIVITY
Systems biology

SECONDARY DOMAINS OF ACTIVITY
Biophysics; Comparative genomics; Evo-Devo; Evolutionary biology; Experimental biology; Machine learning; Next-generation sequencing; Phylogeny; Software engineering; Transcriptomics

DOMAIN OF APPLICATION

KEY RESOURCES
MetaPIGA
Reptilomics

Félix NAEF

GROUP NAME **COMPUTATIONAL SYSTEMS BIOLOGY GROUP**
INSTITUTION EPFL
CITY LAUSANNE

Systems biology is aimed at a quantitative and dynamic understanding of cellular networks by combining experimental data with theoretical and computational methodologies. Our interest lies in the regulatory and cellular networks involved in oncogenic signalling, cell-cycle regulation, and molecular oscillators. Data obtained from technologies such as microarrays, chromatin-immunoprecipitation (ChIP) and genome sequencing are brought together to discover regulatory dependencies between genes and regulatory proteins involved in cell proliferation. One thematic focus is the study of biomolecular oscillators, in particular the circadian clock.

MAIN DOMAIN OF ACTIVITY
Systems biology

SECONDARY DOMAINS OF ACTIVITY
Functional genomics; Single-cell biology; Transcriptomics

DOMAIN OF APPLICATION

Igor PIVKIN

GROUP NAME **SCIENTIFIC COMPUTING GROUP**
INSTITUTION UNIVERSITÀ DELLA SVIZZERA ITALIANA
CITY LUGANO

Our research interests in the Scientific Computing Group lie in the area of multiscale/multiphysics modelling and parallel large-scale simulations of biological systems. We focus on the development of new computational models and corresponding numerical methods suitable for the next generation of super computers. We are working on stochastic multiscale modelling of motion, the interaction, deformation and aggregation of cells under physiological flow conditions, biofilm growth, and coarse-grained molecular dynamics simulations, as well as the modelling of transport processes in healthy and tumour-induced microcirculation.

MAIN DOMAIN OF ACTIVITY
Systems biology

SECONDARY DOMAINS OF ACTIVITY
Biophysics; Cell mechanobiology; Mathematical modelling; Particle-based methods

DOMAIN OF APPLICATION

KEY PAPERS IN 2018

Christel S *et al.*
Weak iron oxidation by *Sulfobacillus thermosulfidooxidans* maintains a favorable redox potential (...)
Frontiers Microbiol.

Bellenberg S *et al.*
Automated microscopic analysis of metal sulfide colonization by acidophilic microorganisms
Appl Environ Microbiol.

Jörg STELLING

GROUP NAME **COMPUTATIONAL SYSTEMS BIOLOGY GROUP**
INSTITUTION ETH ZURICH, D-BSSE
CITY ZURICH, BASEL

We develop and apply computational and – most recently – experimental methods to analyse and design complex cellular networks, with a focus on large-scale mechanistic approaches. The group comprises biologists, computer scientists, engineers, and mathematicians, who perform interdisciplinary research in systems and synthetic biology. We focus on developing and applying computational methods and mechanistic mathematical models to study complex cellular networks and their operating principles, and to enable their rational re-design. Applications rely on our experimental lab, which uses budding yeast as a model organism, and on collaboration.

MAIN DOMAIN OF ACTIVITY
Systems biology

SECONDARY DOMAINS OF ACTIVITY
Mathematical modelling; Signalling pathways; Single-cell biology; Synthetic biology

DOMAINS OF APPLICATION

KEY RESOURCES
CellX
MetaNetX
TopoFilter

KEY PAPERS IN 2018

Yordanov P and Stelling J.
Steady-state differential dose response in biological systems.
Biophys J.

Whitehead E *et al.*
Automated planning enables complex protocols on liquid-handling robots.
ACS Synt Biol.

Widmer LA and Stelling J.
Bridging intracellular scales by mechanistic computational models.
Curr Opin Biotechnol.

Mihaela ZAVOLAN

GROUP NAME **RNA REGULATORY NETWORKS GROUP**
INSTITUTION UNIVERSITY OF BASEL
CITY BASEL

The individual cells of a body exhibit a stunning diversity of phenotypes, despite carrying a largely identical genetic makeup. This is due to the distinct ways in which the same genetic information can be read, interpreted and translated into function. At the RNA Regulatory Networks (RRN) Group at the Biozentrum in Basel, we combine computational modelling with big data and experimental analysis to discover and understand the regulatory networks governing the interpretation of genetic information at the level of tissues and single cells.

MAIN DOMAIN OF ACTIVITY
Systems biology

SECONDARY DOMAINS OF ACTIVITY
Data mining; Deep sequencing data; Experimental biology; Gene regulatory network analysis; Machine learning; Mathematical modelling; Single-cell biology; Transcriptomics; Workflows

DOMAIN OF APPLICATION

KEY RESOURCES
MIRZA
polyAsite

KEY PAPERS IN 2018

Gruber AJ *et al.* Terminal exon characterization with TECtool reveals an abundance of cell-specific isoforms. Nat Methods

Rzepiela A *et al.* Single-cell mRNA profiling reveals the hierarchical response of miRNA targets to miRNA induction. Mol Syst Biol.

Gruber AJ *et al.* Discovery of physiological and cancer-related regulators of 3′ UTR processing with KAPAC. Genome Biol.

Karsten BORGWARDT

GROUP NAME **MACHINE LEARNING AND COMPUTATIONAL BIOLOGY LAB**
INSTITUTION ETH ZURICH
CITY ZURICH

Our lab acts as the bridge between big data analysis and biomedical research. We develop novel data-mining algorithms to detect patterns and statistical dependencies in large datasets from the fields of biology and medicine. Our major goals are twofold: 1) to enable the automatic generation of new knowledge from big data through machine learning, and 2) to gain an understanding of the relationship between biological systems and their molecular properties. Such an understanding is of fundamental importance for personalized medicine, which tailors medical treatment to the molecular properties of a person.

MAIN DOMAIN OF ACTIVITY
Text mining and machine learning

SECONDARY DOMAINS OF ACTIVITY
Biomarkers; Biostatistics; Data mining; GWAS; Machine learning; Personalized medicine; Software engineering; Systems biology

DOMAIN OF APPLICATION

KEY RESOURCES
easyGWAS
AraPheno
graphkernels

KEY PAPERS IN 2018

Bock C *et al.* Association mapping in biomedical time series via statistically significant shapelet mining. Bioinformatics

Togninalli M *et al.* Accurate and adaptive imputation of summary statistics in mixed-ethnicity cohorts. Bioinformatics

Llinares-Lopez F *et al.* CASMAP: Detection of statistically significant combinations of SNPs in association mapping. Bioinformatics

Carlos PEÑA

GROUP NAME **COMPUTATIONAL INTELLIGENCE FOR COMPUTATIONAL BIOLOGY (CI4CB) GROUP**
INSTITUTION HEIG-VD
CITY YVERDON-LES-BAINS

With the advent of high-throughput technologies and clinical information systems, the life sciences and clinical sciences now produce very large amounts of data (big data). Our goal is to uncover hidden patterns in these data, as well as build data-driven models as tools to discover biomarkers and assist clinicians in their decisions. Our projects encompass the fields of transcriptomics, systems biology, and clinical bioinformatics & analytics.

MAIN DOMAIN OF ACTIVITY
Text mining and machine learning

SECONDARY DOMAINS OF ACTIVITY
Biomarkers; Data mining; Machine learning; Metagenomics; Software engineering; Systems biology

DOMAINS OF APPLICATION

KEY RESOURCE
WeStBESel

KEY PAPERS IN 2018

Gomez S *et al.* Improving neural network interpretability via rule extraction. International Conference of Artificial Neural Networks (ICANN)

Despraz J *et al.* Exploring internal representations of deep neural networks. Studies in Comp Intel. Springer

Leite D *et al.* Computational prediction of inter-species relationships through omics data analysis and machine learning. BMC Bioinfo.

Gunnar RÄTSCH

GROUP NAME **BIOMEDICAL INFORMATICS GROUP**
INSTITUTION ETH ZURICH
CITY ZURICH

We are interested in modern machine-learning techniques suitable for the analysis of problems that arise in medicine and biology. In particular, we develop new learning techniques that are capable of dealing with large amounts of genomic data and medical data. These techniques aim to provide accurate predictions on the phenomenon at hand and to comprehensively provide reasons for their prognoses, and thereby assist in gaining new biomedical insights.

MAIN DOMAIN OF ACTIVITY
Text mining and machine learning

SECONDARY DOMAINS OF ACTIVITY
Deep sequencing data; Electronic health record; Infectious diseases; Machine learning; Metagenomics; Oncology

DOMAIN OF APPLICATION

KEY PAPERS IN 2018

Kahles A *et al.* Comprehensive analysis of alternative splicing across tumours from 8,705 patients. Cancer Cell

Mustafa H *et al.* Dynamic compression schemes for graph coloring. Bioinformatics

Locatello F *et al.* Boosting black box variational inference. NIPS spotlight

Fabio RINALDI

GROUP NAME **BIOMEXT: BIOMEDICAL INFORMATION EXTRACTION**
INSTITUTION UNIVERSITY OF ZURICH
CITY ZURICH

The BioMeXT group specializes in Information Extraction from biomedically relevant textual sources, such as the scientific literature, clinical records, or social media. We focus in particular on the extraction of domain-specific entities (such as genes, drugs and diseases), and their semantic relationships (e.g. gene-disease associations). Our tools are often evaluated through participation in community-run evaluation challenges (e.g. BioCreAtIvE). We also provide an environment for Assisted Curation (ODIN), which is used in the curation pipeline of the RegulonDB database, in an NIH-funded project.

MAIN DOMAIN OF ACTIVITY
Text mining and machine learning

SECONDARY DOMAINS OF ACTIVITY
API; Database curation; Electronic health record; Machine learning; Ontology; Semantic web format; Software engineering; Text mining

DOMAINS OF APPLICATION

KEY RESOURCES
BioTermHub
OGER++

KEY PAPERS IN 2018

Lauriola I *et al.* Learning preferences for large scale multi-label problems. International Conference of Artificial Neural Networks (ICANN)

Küker S *et al.* The value of necropsy reports for animal health surveillance. BMC Vet Res.

Patrick RUCH

GROUP NAME **TEXT MINING GROUP**
INSTITUTION HES-SO - GENEVA SCHOOL OF BUSINESS ADMINISTRATION (HEG)
CITY GENEVA

In the Text Mining Group, we carry out activities in semantic interoperability and text analytics applied to the health and life sciences. Previously hosted by the Radiology and Medical Informatics Department of the Geneva University Hospitals, our group moved to the University of Applied Sciences Geneva (HES-SO- HEG Geneva) in 2008. We develop text-mining solutions to support both the annotation of curated databases and the work of a wide range of biomedical professionals from drug designers to clinicians. We thus develop specific biomedical decision-support systems, in particular in oncology and molecular pathology.

MAIN DOMAIN OF ACTIVITY
Text mining and machine learning

SECONDARY DOMAINS OF ACTIVITY
Data management; Database curation; Electronic health record; Oncology; Software engineering

DOMAIN OF APPLICATION

KEY RESOURCES
GOCat
neXtA5
EAGLi

KEY PAPERS IN 2018

Britan A *et al.* Accelerating annotation of articles via automated approaches: evaluation of the neXtA5 curation-support tool by neXtProt. Database

Gobeill J *et al.* Overview of the BioCreative VI text-mining services for Kinome Curation Track. Database

Niveau G *et al.* Suicide notes: their utility in understanding the motivations behind suicide and preventing future ones. Arch Suicide Res.

Julia VOGT NEW

GROUP NAME **ADAPTIVE SYSTEMS AND MEDICAL DATA SCIENCE GROUP**
INSTITUTION UNIVERSITY OF BASEL
CITY BASEL

The primary research interest of our group lies at the intersection of machine learning and medicine. We work on advancing and developing novel machine learning techniques for precision medicine, the life sciences and clinical data analysis. The field of action comprises many areas such as prediction of response to treatment in personalized medicine, (sparse) biomarker detection, tumour classification or the understanding of interactions between genes or groups of genes. We take a data- and problem-driven approach to these challenges. In close collaboration with physicians we identify gaps where current technologies fail and develop tailored solutions.

MAIN DOMAIN OF ACTIVITY
Text mining and machine learning

SECONDARY DOMAIN OF ACTIVITY
Machine learning

DOMAIN OF APPLICATION

Rémy BRUGGMANN

GROUP NAME **INTERFACULTY BIOINFORMATICS UNIT (IBU)**
INSTITUTION UNIVERSITY OF BERN
CITY BERN

In the Interfaculty Bioinformatics Unit of the University of Bern (IBU), we provide services and expertise to assist researchers of the three “Life-Science” Faculties and Insel hospital in data analysis and project planning for large-scale experiments with focus on next-generation sequencing application. We have our own research programme and collaborate on large and complex projects. Furthermore, we develop methods to analyse high-throughput data and provide bioinformatics training. We have a high-performance computing cluster and a data storage system that we use for our own research, collaborations and service projects.

MAIN DOMAIN OF ACTIVITY
Core facilities and competence centres

SECONDARY DOMAINS OF ACTIVITY
Biostatistics; Data management; Next-generation sequencing; Personalized medicine; Single-cell biology; Systems biology; Training; Transcriptomics; Workflows

DOMAINS OF APPLICATION

KEY RESOURCES
SetRank

KEY PAPERS IN 2018

Widmer Y F *et al.* Regulators of long-term memory revealed by mushroom body-specific gene expression profiling in *Drosophila melanogaster*. Genetics

Schmid M W *et al.* Contribution of epigenetic variation to adaptation in Arabidopsis. Nat commun.

Wüthrich D *et al.* Conversion of methionine to cysteine in *Lactobacillus paracasei* depends on the highly mobile cysK-ctl-cysE gene cluster. Front Microbiol.

Luciano CASCIONE

GROUP NAME **BIOINFORMATICS CORE UNIT**
INSTITUTION INSTITUTE OF ONCOLOGY RESEARCH
CITY BELLINZONA

Our main task at the Bioinformatics Core Unit (BCU) is to support the research groups at the Institute of Oncology Research (IOR) with computational and statistical services. Our research interests are focused on the genetics and biology of cancer with a major emphasis on lymphomas and epithelial cancers, such as prostate, breast and ovarian cancer. Importantly, more than as just a supporting role, we proactively identify and develop novel bioinformatics projects that can complement and in many cases drive our biologic research.

MAIN DOMAIN OF ACTIVITY
Core facilities and competence centres

SECONDARY DOMAINS OF ACTIVITY
Biostatistics; Drug resistance; Machine learning; Systems biology; Transcriptomics

DOMAIN OF APPLICATION

KEY PAPERS IN 2018

Mensah AA *et al.* Bromodomain and extra-terminal domain inhibition modulates the expression of pathologically relevant microRNAs in diffuse large B-cell lymphoma. Haematologica

Bonnici V *et al.* cuRnet: an R package for graph traversing on GPU. BMC Bioinformatics

Ceriani L *et al.* Metabolic heterogeneity on baseline 18FDG-PET/CT scan is a predictor of outcome in primary mediastinal B-cell lymphoma. Blood

Mauro DELORENZI
Frédéric SCHÜTZ

GROUP NAME **BIOINFORMATICS CORE FACILITY**
INSTITUTION UNIVERSITY OF LAUSANNE
CITY LAUSANNE

In the Bioinformatics Core Facility (BCF), we promote trans-disciplinary collaborations between research teams in medicine, molecular biology, genetics, genomics, statistics, and bioinformatics. In particular, we carry out analyses of biomedical-genomics data with a focus on biomarker studies in cancer research, building on our specific expertise in statistical methods for genomics data analysis. Recently, we have concentrated on molecular heterogeneity in cancer, colorectal cancer, single cells transcriptomics, genetic association studies, metagenomics and other new technologies. We are open to any kind of new research direction.

MAIN DOMAIN OF ACTIVITY
Core facilities and competence centres

SECONDARY DOMAINS OF ACTIVITY
Biomarkers; Data mining; Deep sequencing data; Mathematical modelling; Metagenomics; Next generation sequencing; Oncology; Single-cell biology; Transcriptomics

DOMAINS OF APPLICATION

KEY PAPERS IN 2018

Schütz F and Zollinger A ABPS: An R package for calculating the abnormal blood profile score. Front Physiol.

Nisa L *et al.* Comprehensive genomic profiling of patient-matched head and neck cancer cells: A preclinical pipeline for metastatic (...). Mol Cancer Res.

Bady P *et al.* The DNA methylome of DDR genes and benefit from RT or TMZ in IDH mutant low-grade glioma treated in EORTC 22033. Acta Neuropathol.

Mark IBBERSON NEW

GROUP NAME **VITAL-IT**
INSTITUTION SIB LAUSANNE
CITY LAUSANNE

We are a competence centre in bioinformatics with a multidisciplinary team of scientists and technical staff. Our mission is to provide bioinformatics support to life science projects in Switzerland and abroad through scientific collaborations with academic and industry partners. The services and expertise we offer include algorithm development and optimization, software engineering, data management including database expertise (Oracle certified), data analysis (genomics, transcriptomics, proteomics, metabolomics, lipidomics), development of web-based data mining and visualization tools and *in silico* modelling.

MAIN DOMAIN OF ACTIVITY
Core facilities and competence centres

SECONDARY DOMAINS OF ACTIVITY
Data management; Data mining; Genome reconstruction; Machine learning; Mass spectrometry; Next-generation sequencing; Personalized medicine; Software engineering; Structural biology; Systems biology

DOMAINS OF APPLICATION

KEY RESOURCES
SOURCEDATA
SwissLipids
MetaNetX

KEY PAPERS IN 2018

Tran VDT *et al.* Condition-specific series of metabolic sub-networks and its application for gene set enrichment analysis. Bioinformatics

Bekkar A *et al.* Expert curation for building network-based dynamical models: a case study on atherosclerotic plaque formation. Database

Tanyi JL *et al.* Personalized cancer vaccine effectively mobilizes antitumour T cell immunity in ovarian cancer. Sci Transl Med.

Hubert REHRAUER

GROUP NAME **FUNCTIONAL GENOMICS CENTRE ZURICH (FGCZ)**
INSTITUTION ETH ZURICH, UNIVERSITY OF ZURICH
CITY ZURICH

We are dedicated to the processing, analysis and interpretation of next-generation sequencing data. We interact closely with research groups, and provide tailored comprehensive bioinformatics solutions. Additionally, we provide standard analysis pipelines for the most frequent research questions. We train researchers and bioinformaticians on various aspects of data analysis, and provide access to our computing infrastructure for running analyses.

MAIN DOMAIN OF ACTIVITY
Core facilities and competence centres

SECONDARY DOMAINS OF ACTIVITY
Biomarkers; Biostatistics; Data management; Data mining; Deep sequencing data; Gene regulatory network analysis; Next-generation sequencing; Single-cell biology; Software engineering; Transcriptomics

DOMAINS OF APPLICATION

KEY RESOURCE
Sushi

KEY PAPERS IN 2018

Rehrauer H *et al.* How asbestos drives the tissue towards tumours: YAP activation, macrophage and mesothelial precursor recruitment, RNA editing, and somatic mutations. Oncogene

Degirmenci B *et al.* GLI1-expressing mesenchymal cells form the essential Wnt-secreting niche for colon stem cells. Nature

Marcel RIEDI

GROUP NAME **SERVICE AND SUPPORT FOR SCIENCE IT (S3IT)**
INSTITUTION UNIVERSITY OF ZURICH
CITY ZURICH

In the Service and Support for Science IT (S3IT) unit, we provide support for science in general, and for life science and medicine in particular. S3IT serves as a partner for both local and national projects to enable competitive research with the advanced use of computational methods and resources. Our team advises groups and projects about data management and data analysis, and cooperates to optimize their specific workflow. S3IT also takes part in national projects and cooperates with similar technology-oriented groups to ensure that its expertise is always up to date.

MAIN DOMAIN OF ACTIVITY
Core facilities and competence centres

SECONDARY DOMAINS OF ACTIVITY
Data management; Data mining; Infectious diseases; Machine learning; Next-generation sequencing; Protein interactions; Proteomics; Software engineering; Systems biology

DOMAIN OF APPLICATION

KEY RESOURCES
S3IT compute and storage infrastructure
ElastiCluster / GC3Pie
S3IT data manager

KEY PAPER IN 2018

Pedrioli D L *et al.*
Comprehensive ADP-ribosylome analysis identifies tyrosine as an ADP-ribose acceptor site.
EMBO Rep.

Bernd RINN

GROUP NAME **SCIENTIFIC IT SERVICES (SIS)**
INSTITUTION ETH ZURICH, D-BSSE
CITY ZURICH, BASEL

SIS is an interdisciplinary bioinformatics and scientific IT support group which builds up computational tools – from lab databases to reusable framework components for data analysis and data management – for life science research and beyond. We improve and port scientific software, integrate and operate data analysis pipelines, and provide training and consulting in databases, scientific software development, high-performance and cloud computing. We collaborate with Swiss national research initiatives, European research groups and the life science industry (e.g. SPHN, PHRT, BioMedIT, SystemsX.ch, HPC-CH, EnhanceR).

MAIN DOMAIN OF ACTIVITY
Core facilities and competence centres

SECONDARY DOMAINS OF ACTIVITY
Data management; Data security; Data visualization; Software engineering; Training; Workflows

DOMAINS OF APPLICATION

KEY RESOURCES
openBIS
emzed

KEY PAPERS IN 2018

Panyasantisuk J *et al.*
Mapping anisotropy improves QCT-based finite element estimation of hip strength in pooled stance and side-fall load configurations.
Med Eng & Phys.

Shi G *et al.*
Elucidating thermal behavior, native contacts, and folding funnels of simple lattice proteins using replica exchange Wang-Landau sampling.
J of Chem Phys.

Torsten SCHWEDE
Thierry SENGSTAG

GROUP NAME **SCICORE**
INSTITUTION UNIVERSITY OF BASEL
CITY BASEL

sciCORE is a competence centre in scientific computing located at the University of Basel. Over 170 research groups in the fields of bioinformatics, computational chemistry, physics, biology, medicine, and economics use our services. Our expertise ranges from high-performance computing infrastructure, large-scale storage resources, scientific software and databases to scientific consulting and education. We help scientists with computational needs achieve their research goals. We operate the IT infrastructure behind several key SIB bioinformatics resources and, as a BioMedIT node, we actively support Swiss biomedical research.

MAIN DOMAIN OF ACTIVITY
Core facilities and competence centres

SECONDARY DOMAINS OF ACTIVITY
Biostatistics; Data management; Data mining; Data security; Training; Workflows

Daniel STEKHOVEN

GROUP NAME **CLINICAL BIOINFORMATICS UNIT**
INSTITUTION ETH ZURICH
CITY ZURICH

We are the Clinical Bioinformatics Unit at the ETH Zurich core facility NEXUS Personalized Health Technologies. Our unit offers highly customizable bioinformatics and biostatistics services for projects in the field of biomedical research and development. We maintain close collaborations with hospitals in Switzerland, enabling and accelerating the use of top-notch data analysis, software development, and clinical decision support. Based on a fee-for-service or joint collaboration model, we aim for tailor-made full services for our clients, translating computational and analytical tools to meet their needs.

MAIN DOMAIN OF ACTIVITY
Core facilities and competence centres

SECONDARY DOMAINS OF ACTIVITY
API; Biostatistics; Data visualization; Drug resistance; Oncology; Personalized medicine; Proteomics; Single-cell biology; Software engineering; Transcriptomics

DOMAINS OF APPLICATION

KEY PAPERS IN 2018

Singer F *et al.*
SwissMTB: establishing comprehensive molecular cancer diagnostics in Swiss clinics.
BMD Med Inform Decis Mak.

Stekhoven D *et al.*
Swiss Variant Interpretation for Oncology (SVIP-O).
Swiss Med Info.

Kahles A *et al.*
Comprehensive analysis of alternative splicing across tumours from 8,705 patients.
Cancer Cell

INDEX OF SIB GROUP LEADERS

as of 31 December 2018

ACKNOWLEDGEMENTS

We gratefully acknowledge the following funders, sponsors and partners for their financial support and encouragement in helping us fulfil our mission in 2018.

- The Swiss government and in particular:
  - The State Secretariat for Education, Research and Innovation (SERI)
  - The Swiss National Science Foundation (SNSF)
  - Innosuisse
- Our institutional partners
- The European Commission
- The Leenaards Foundation
- The Medic Foundation
- The National Institutes of Health (NIH)
- The Research for Life Foundation
- SystemsX.ch

We are very thankful to our 20th anniversary's sponsors and partners (SEE P. 27) whose contribution fully financed the projects and events launched to mark the occasion.

The SIB Fellowship programme continues, thanks to the generous support of:

- The R. Geigy Foundation
- The University of Geneva
- The University of Lausanne

We also thank all industrial and academic partners who trust SIB's expertise.

IMPRESSUM

SIB GROUP LEADERS AND ACTIVITY PICTURES BY
Nicolas Righetti, www.lundi13.ch

DESIGN AND LAYOUT BY
Bogsch & Bacco, www.bogsch-bacco.ch

IMAGES CREDITS
(from top to bottom and from left to right)

P. 7 Franziska Gruhl - SIB. All rights reserved
Franziska Gruhl - SIB. All rights reserved
Sutthaburawonk / iStock
Fabio Rinaldi - SIB. All rights reserved
P. 27 Bogsch & Bacco
P. 46 f00sion / iStock
JohanSjolander / iStock
Interfoto / Alamy Stock Photo
P. 47 Jose Luis Calvo / Shutterstock
Alex Puddephatt / Shutterstock

SIB Swiss Institute
of Bioinformatics
Quartier Sorge
Bâtiment Amphipôle
CH – 1015 Lausanne
T. +41 21 692 40 50

www.sib.swiss