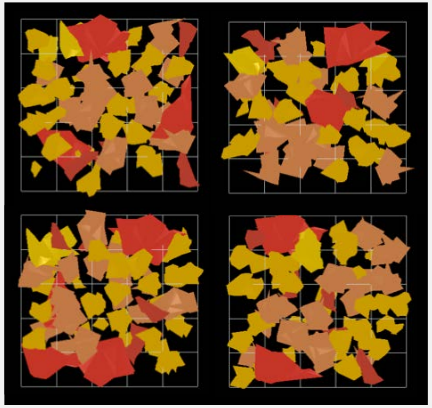
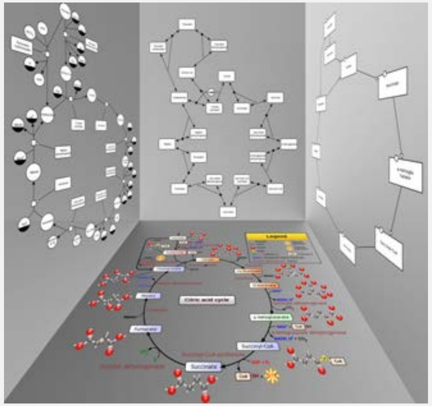




MAX-PLANCK-GESELLSCHAFT

Zukunft Reproducible Science

ZKI Frühjahrstagung Konstanz 2018

Björn Sommer, University of Konstanz

Talk for

- Kamran Safi
    - MPI Radolfzell
    - Research Scientist
    - Movebank/AniMove
- Falk Schreiber
    - University of Konstanz
    - Prof.
    - COMBINE coordinator

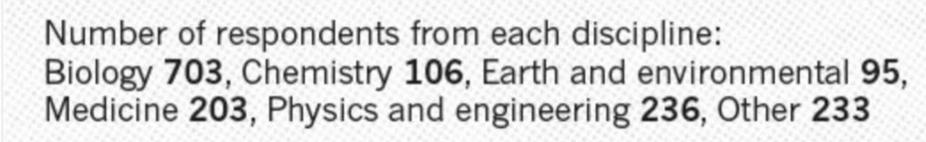
MAX-PLANCK-GESELLSCHAFT

Max-Planck-Institut
für Ornithologie

Universität
Konstanz

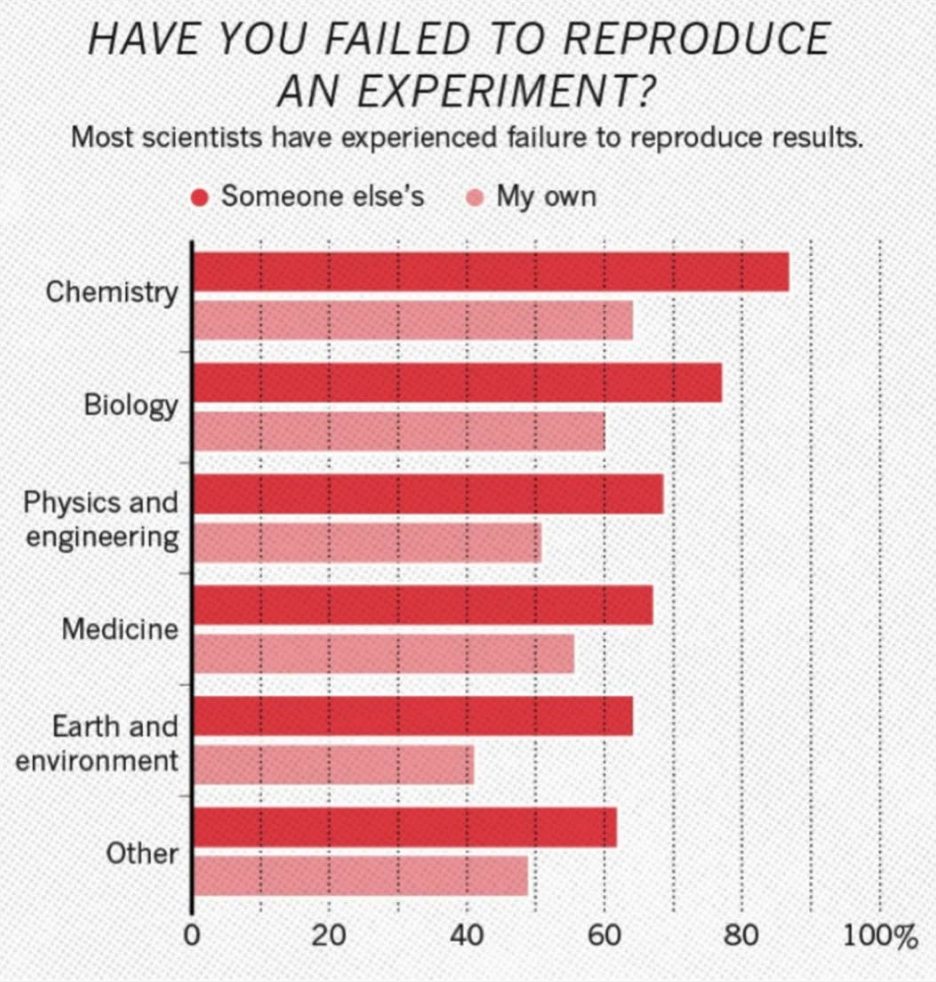
Reproducibility crisis is partly caused by:

- Misinterpretation,
- Misunderstanding,
- Incompleteness, and/or
- Unreadability.



Number of respondents from each discipline:
Biology **703**, Chemistry **106**, Earth and environmental **95**, Medicine **203**, Physics and engineering **236**, Other **233**

Baker, Monya. "1,500 scientists lift the lid on reproducibility." Nature News 533.7604 (2016): 452.



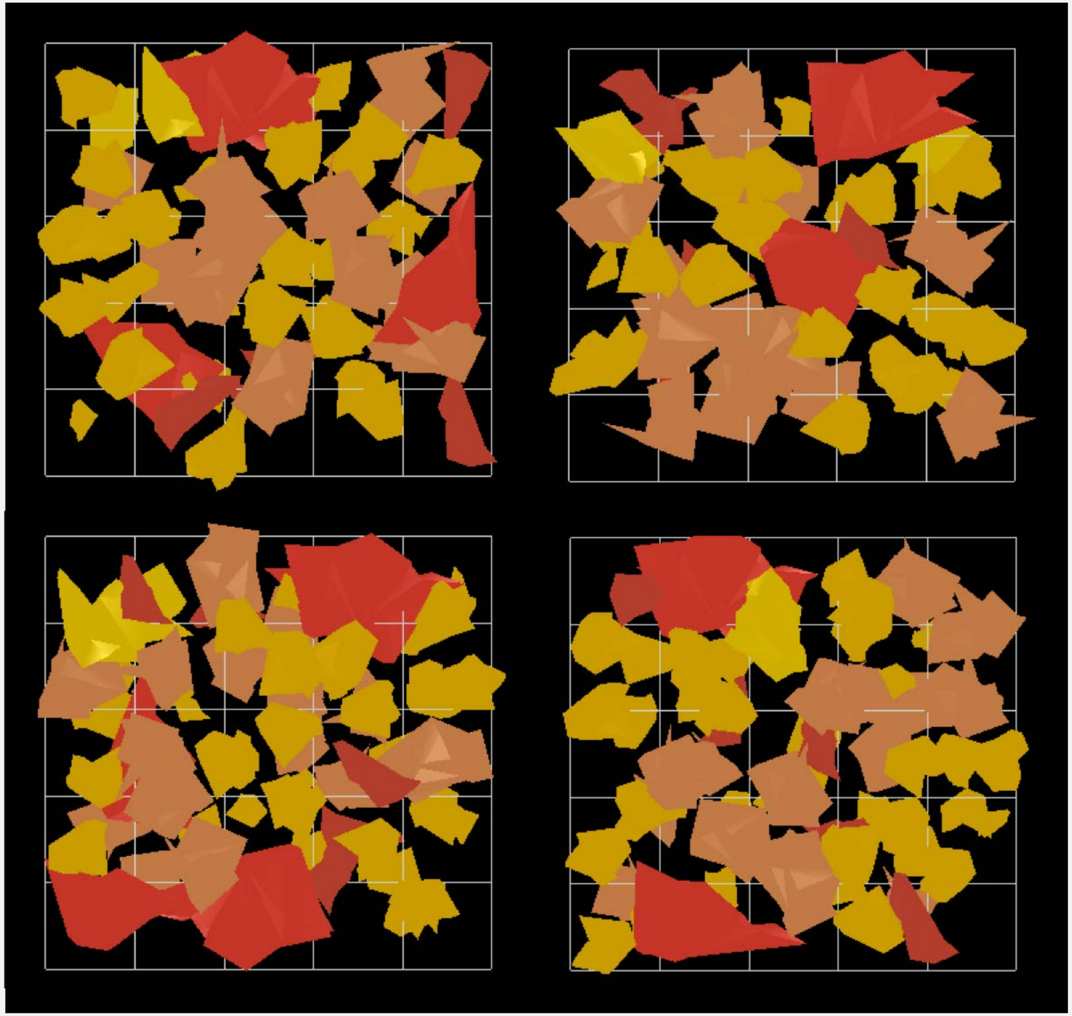
Part of the Reproducibility Problem

- *“**An article** about computational science in a scientific publication **is not the scholarship itself**, **it is merely advertising of the scholarship**.”*
- *“The actual **scholarship is the complete** ... **set of instructions which generated the figures**.”*
  - *David Donoho, 1998*

3 Data Science Applications of Bio Research

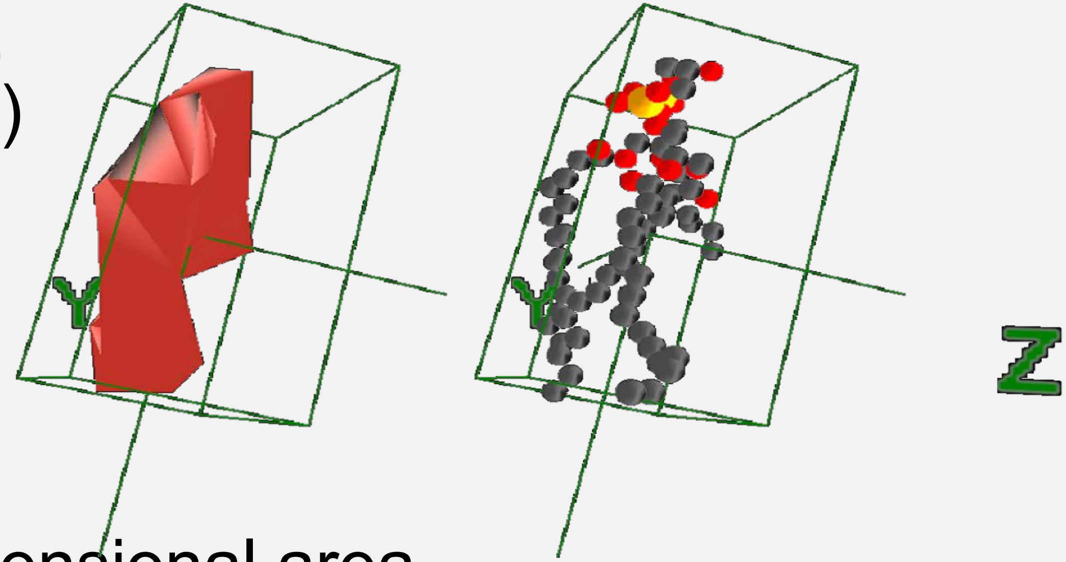
- 3 Topics Today
  - Molecular Biology
  - Systems Biology
  - Collective Behaviour

2. Molecular Biology



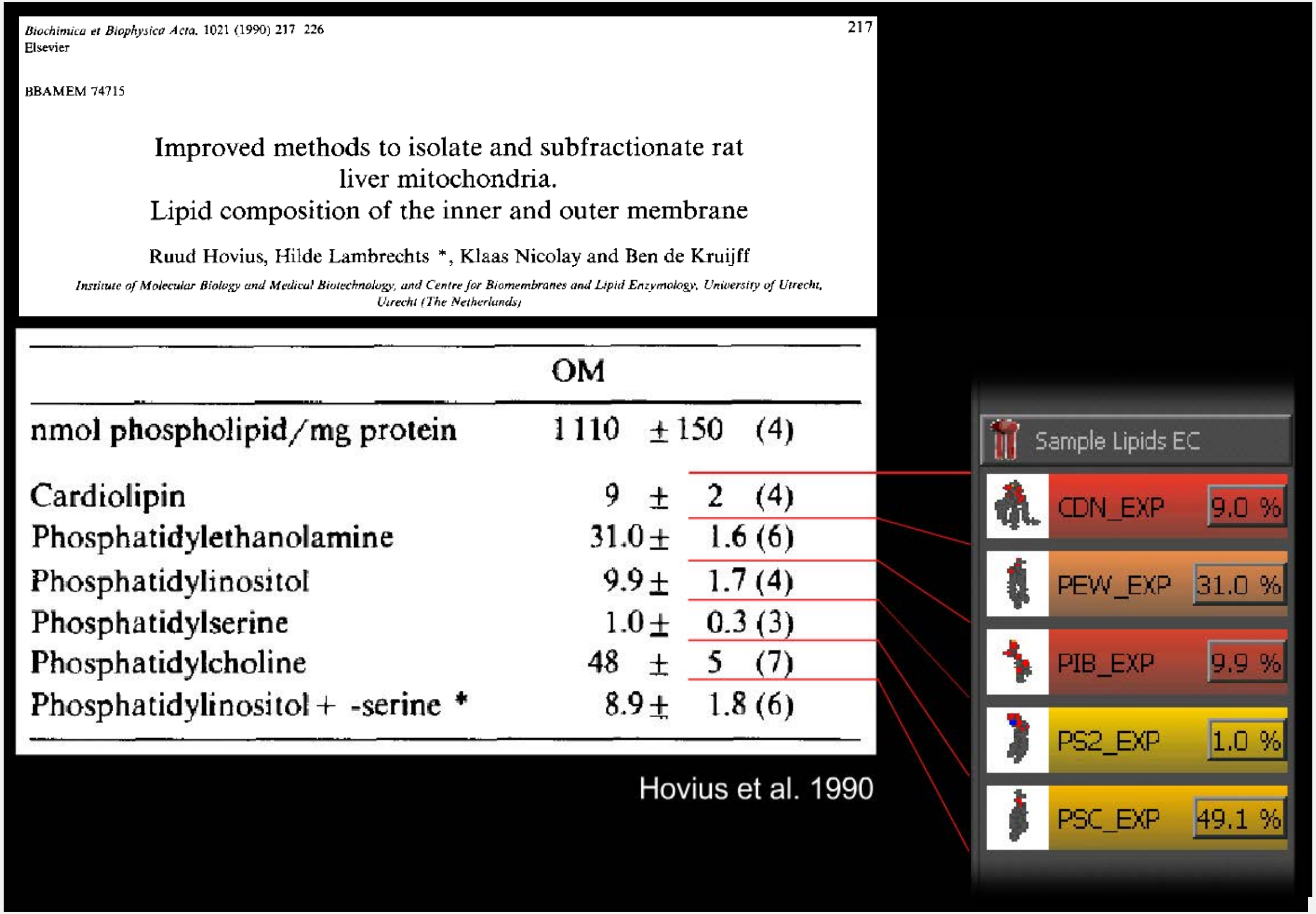
Membrane Generation and Simulation

- in CELLmicrocosmos 2.2 MembraneEditor (CmME)



- solving the Lipid Packing Problem
    - movement along 2-dimensional area
    - shape-based collision detection operates in 3 dimensions
    - all packing settings are stored in a protocol

Defining a Membrane Composition in CmME



Biochimica et Biophysica Acta, 1021 (1990) 217 226
Elsevier

217

BBAMEM 74715

Improved methods to isolate and subfractionate rat liver mitochondria.
Lipid composition of the inner and outer membrane

Ruud Hovius, Hilde Lambrechts *, Klaas Nicolay and Ben de Kruijff

Institute of Molecular Biology and Medical Biotechnology, and Centre for Biomembranes and Lipid Enzymology, University of Utrecht, Utrecht (The Netherlands)

	OM
nmol phospholipid/mg protein	1 110 ± 150 (4)
Cardiolipin	9 ± 2 (4)
Phosphatidylethanolamine	31.0 ± 1.6 (6)
Phosphatidylinositol	9.9 ± 1.7 (4)
Phosphatidylserine	1.0 ± 0.3 (3)
Phosphatidylcholine	48 ± 5 (7)
Phosphatidylinositol + -serine *	8.9 ± 1.8 (6)

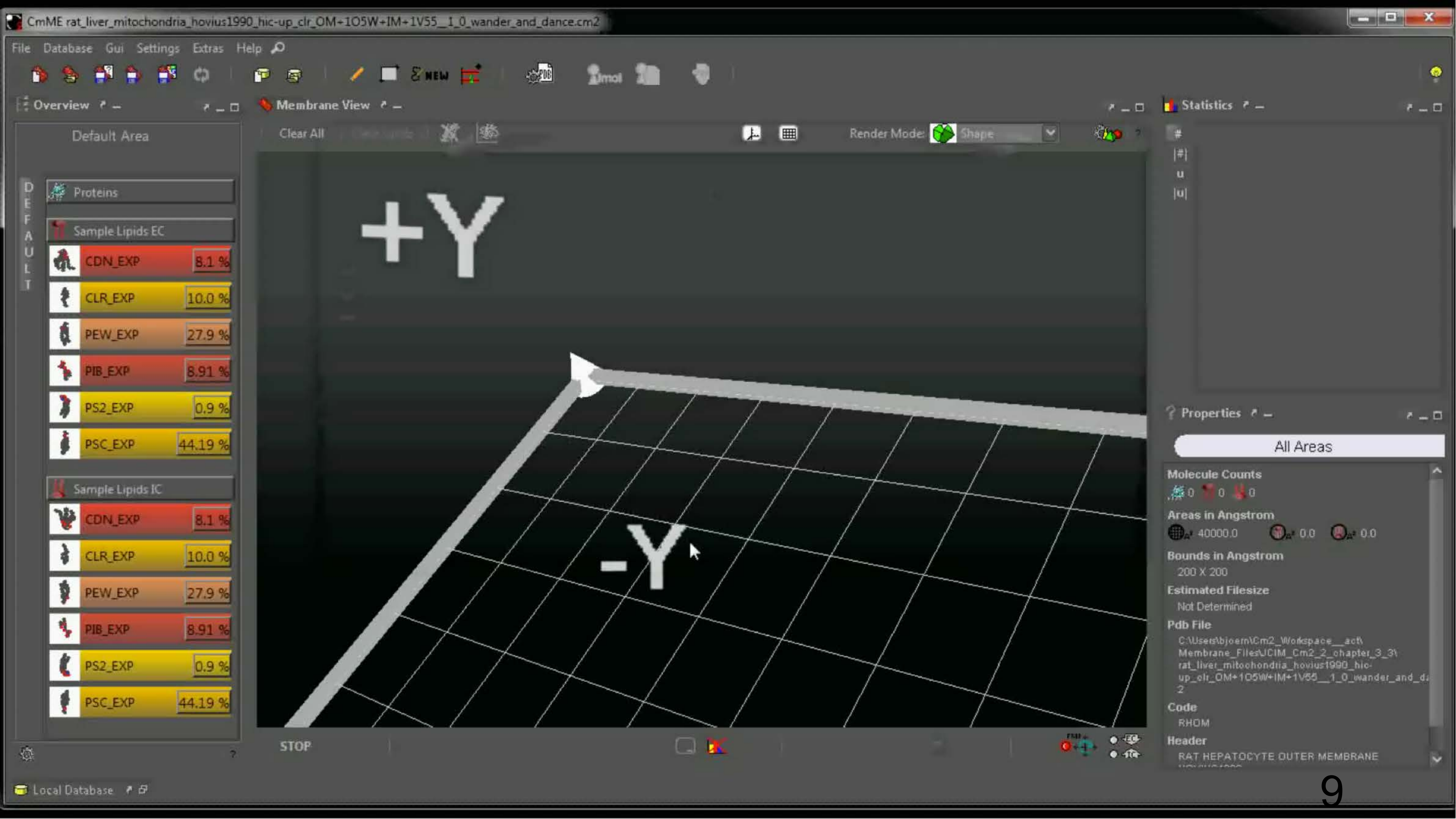
Hovius et al. 1990

Sample Lipids EC

CDN_EXP	9.0 %
PEW_EXP	31.0 %
PIB_EXP	9.9 %
PS2_EXP	1.0 %
PSC_EXP	49.1 %

B. Sommer et al. JCIM 5(51):1165–82, 2011.

Hovius, R. et al. BBA-Biomembranes 1021(2):217–26, 1990.



CmME rat_liver_mitochondria_hovius1990_hic-up_clr_OM+1O5W+IM+1V55__1_0_wander_and_dance.cm2

Proteins

Sample Lipids EC

CDN_EXP	8.1 %
CLR_EXP	10.0 %
PEW_EXP	27.9 %
PIB_EXP	8.91 %
PS2_EXP	0.9 %
PSC_EXP	44.19 %

Sample Lipids IC

CDN_EXP	8.1 %
CLR_EXP	10.0 %
PEW_EXP	27.9 %
PIB_EXP	8.91 %
PS2_EXP	0.9 %
PSC_EXP	44.19 %

All Areas

Molecule Counts
0 0 0

Areas in Angstrom
40000.0 0.0 0.0

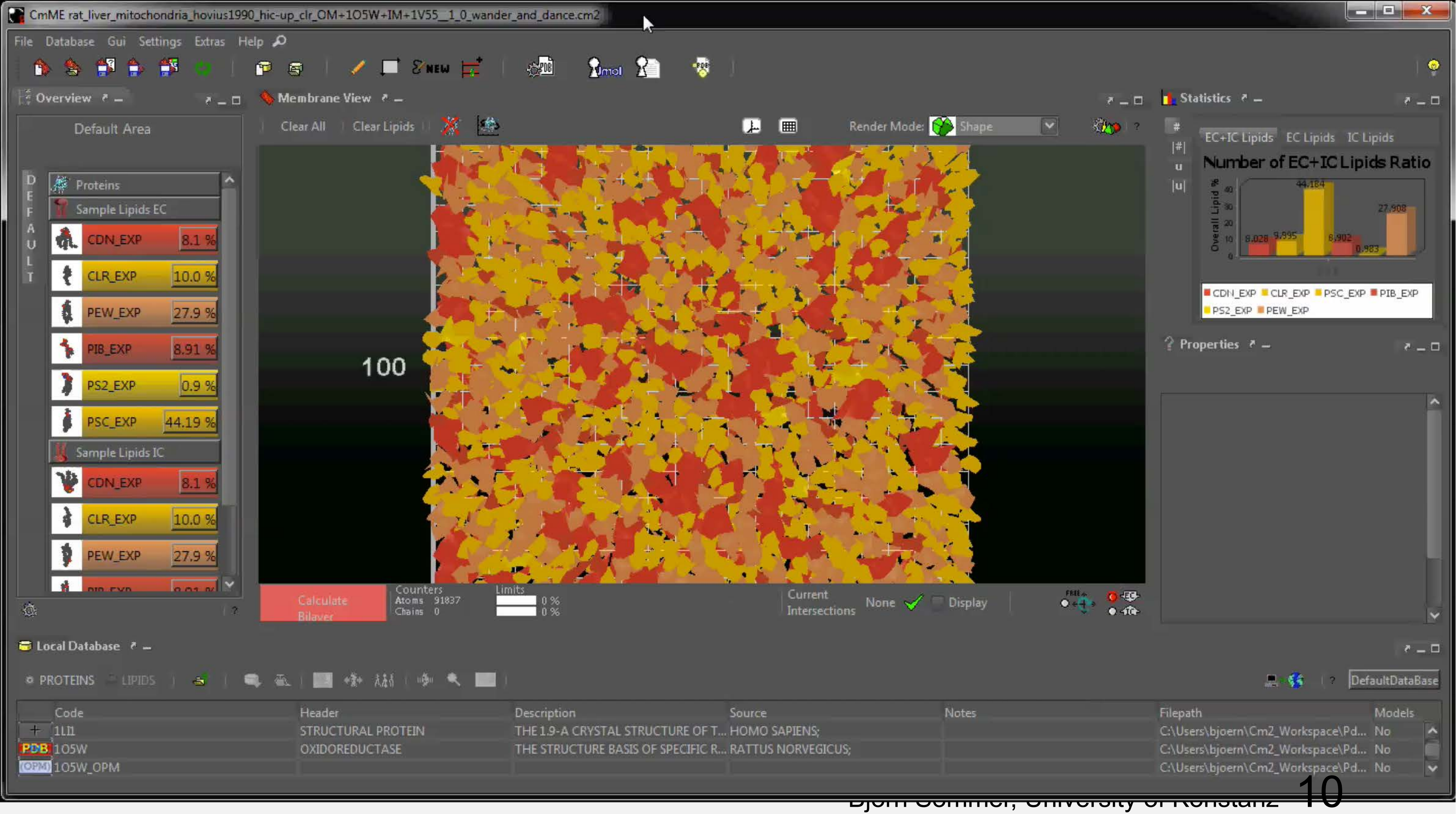
Bounds in Angstrom
200 X 200

Estimated Filesize
Not Determined

Pdb File
C:\Users\bjoern\Cm2_Workspace__act\Membrane_Files\JCIM_Cm2_2_chapter_3_3\rat_liver_mitochondria_hovius1990_hic-up_clr_OM+1O5W+IM+1V55__1_0_wander_and_da
2

Code
RHOM

Header
RAT HEPATOCYTE OUTER MEMBRANE



CmME rat_liver_mitochondria_hovius1990_hic-up_clr_OM+1O5W+IM+1V55__1_0_wander_and_dance.cm2

Sample Lipids EC

CDN_EXP	8.1 %
CLR_EXP	10.0 %
PEW_EXP	27.9 %
PIB_EXP	8.91 %
PS2_EXP	0.9 %
PSC_EXP	44.19 %

Sample Lipids IC

CDN_EXP	8.1 %
CLR_EXP	10.0 %
PEW_EXP	27.9 %
PIB_EXP	8.91 %

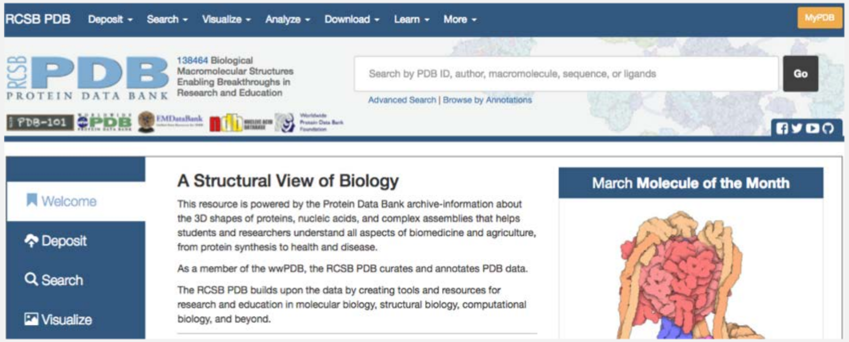
Counters
Atoms 91837
Chains 0

Limits
0 %
0 %

	Code	Header	Description	Source	Notes	Filepath	Models
+	1LI1	STRUCTURAL PROTEIN	THE 1.9-A CRYSTAL STRUCTURE OF T...	HOMO SAPIENS;		C:\Users\bjoern\Cm2_Workspace\Pd...	No
PDB	1O5W	OXIDOREDUCTASE	THE STRUCTURE BASIS OF SPECIFIC R...	RATTUS NORVEGICUS;		C:\Users\bjoern\Cm2_Workspace\Pd...	No
OPM	1O5W_OPM					C:\Users\bjoern\Cm2_Workspace\Pd...	No

Applied Databases by CmME

- PDB for proteins
- OPM for protein placement
- PDB_TM for transmembrane protein placement
- lipid import: various lipid databases



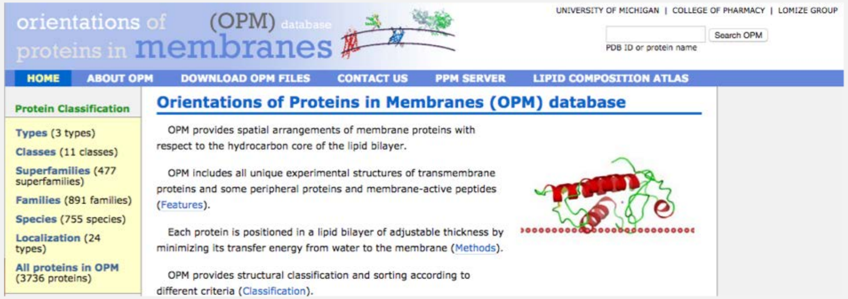
138464 Biological Macromolecular Structures Enabling Breakthroughs in Research and Education

A Structural View of Biology

This resource is powered by the Protein Data Bank archive-information about the 3D shapes of proteins, nucleic acids, and complex assemblies that helps students and researchers understand all aspects of biomedicine and agriculture, from protein synthesis to health and disease.

As a member of the wwPDB, the RCSB PDB curates and annotates PDB data.

The RCSB PDB builds upon the data by creating tools and resources for research and education in molecular biology, structural biology, computational biology, and beyond.



Types (3 types)
Classes (11 classes)
Superfamilies (477 superfamilies)
Families (891 families)
Species (755 species)
Localization (24 types)
All proteins in OPM (3736 proteins)

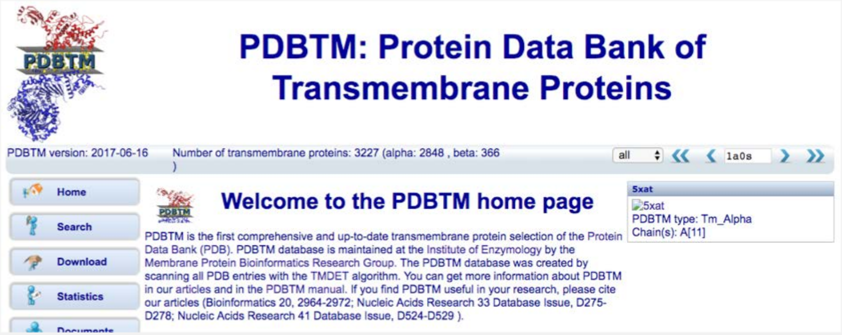
Orientations of Proteins in Membranes (OPM) database

OPM provides spatial arrangements of membrane proteins with respect to the hydrocarbon core of the lipid bilayer.

OPM includes all unique experimental structures of transmembrane proteins and some peripheral proteins and membrane-active peptides (Features).

Each protein is positioned in a lipid bilayer of adjustable thickness by minimizing its transfer energy from water to the membrane (Methods).

OPM provides structural classification and sorting according to different criteria (Classification).



PDBTM: Protein Data Bank of Transmembrane Proteins

PDBTM version: 2017-06-16

Number of transmembrane proteins: 3227 (alpha: 2848 , beta: 366)

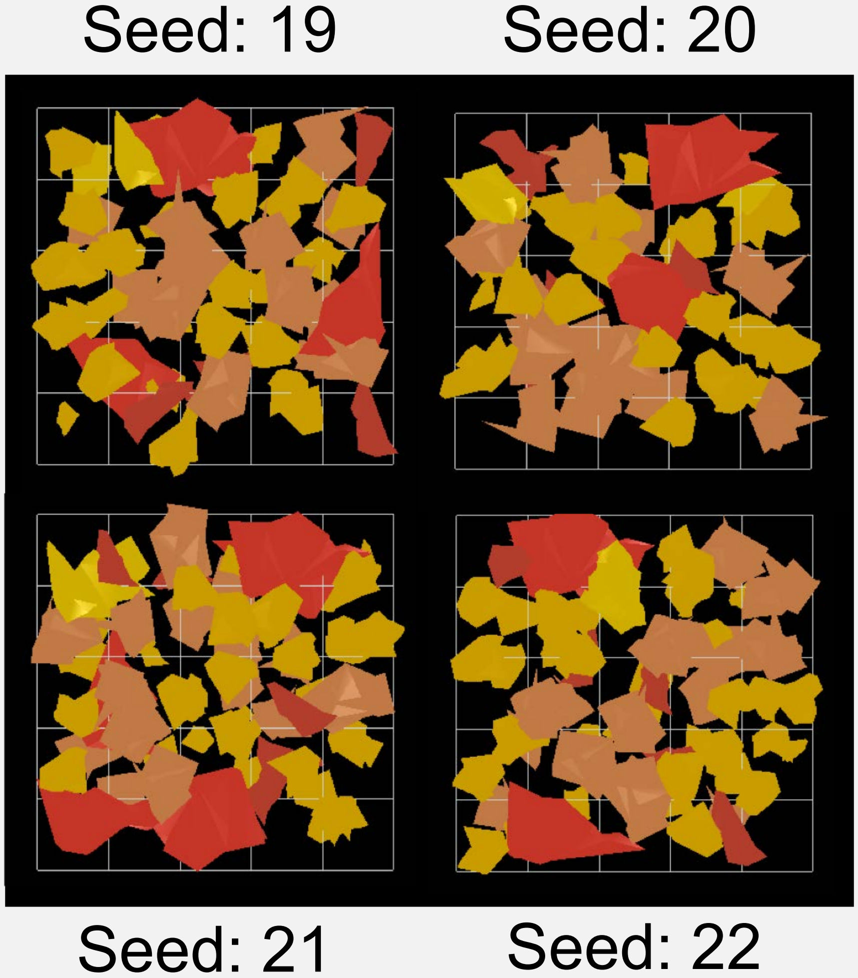
Welcome to the PDBTM home page

PDBTM is the first comprehensive and up-to-date transmembrane protein selection of the Protein Data Bank (PDB). PDBTM database is maintained at the Institute of Enzymology by the Membrane Protein Bioinformatics Research Group. The PDBTM database was created by scanning all PDB entries with the TMDET algorithm. You can get more information about PDBTM in our articles and in the PDBTM manual. If you find PDBTM useful in your research, please cite our articles (Bioinformatics 20, 2964-2972; Nucleic Acids Research 33 Database Issue, D275-D278; Nucleic Acids Research 41 Database Issue, D524-D529).

5xat
5xat
PDBTM type: Tm_Alpha
Chain(s): A[11]

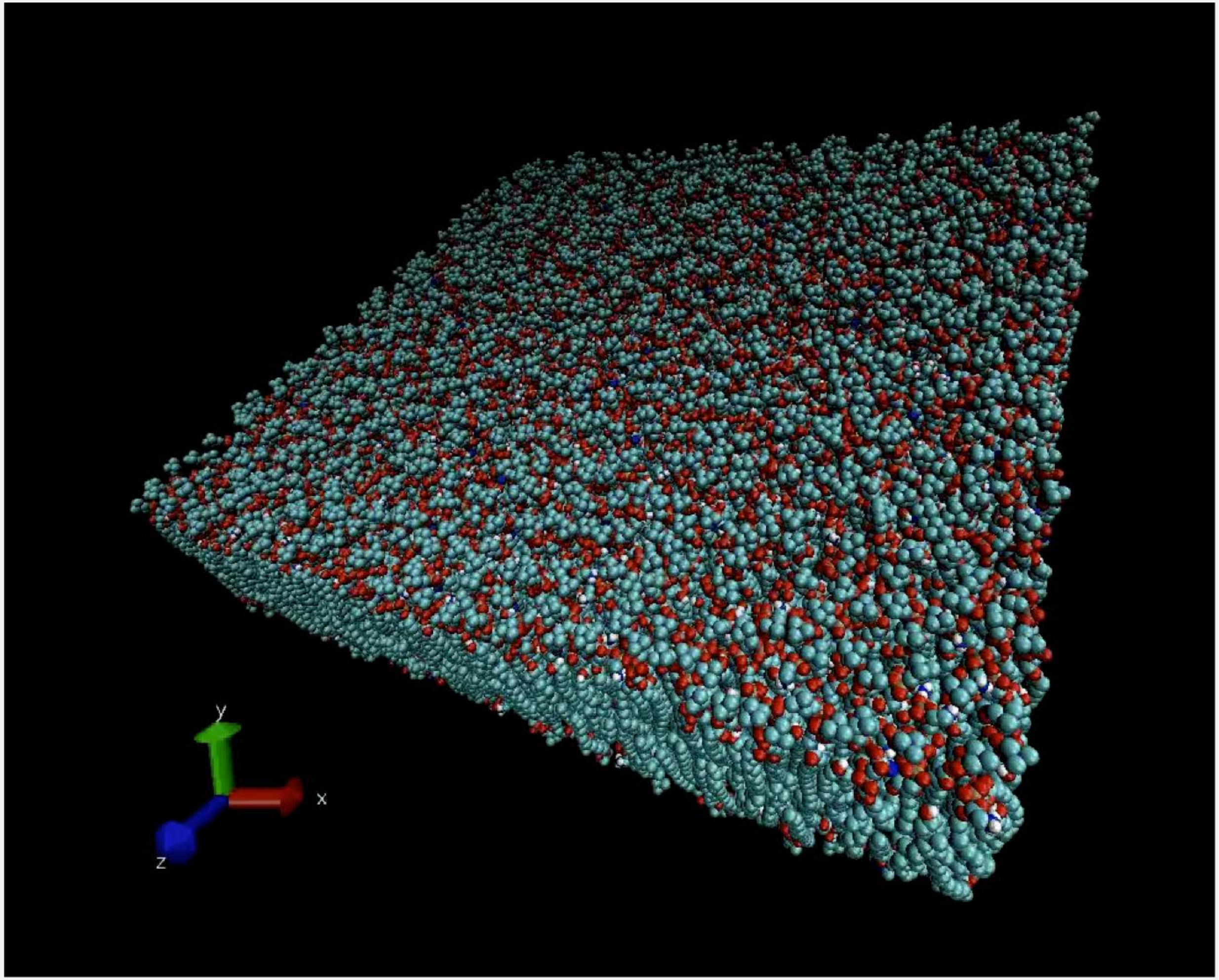
Membranes by Seed

- example from CmME
- using the same random seed with the same algorithm and configuration, the same membrane will be generated.
- using different seeds results in various models of the same configuration



Membrane Simulation

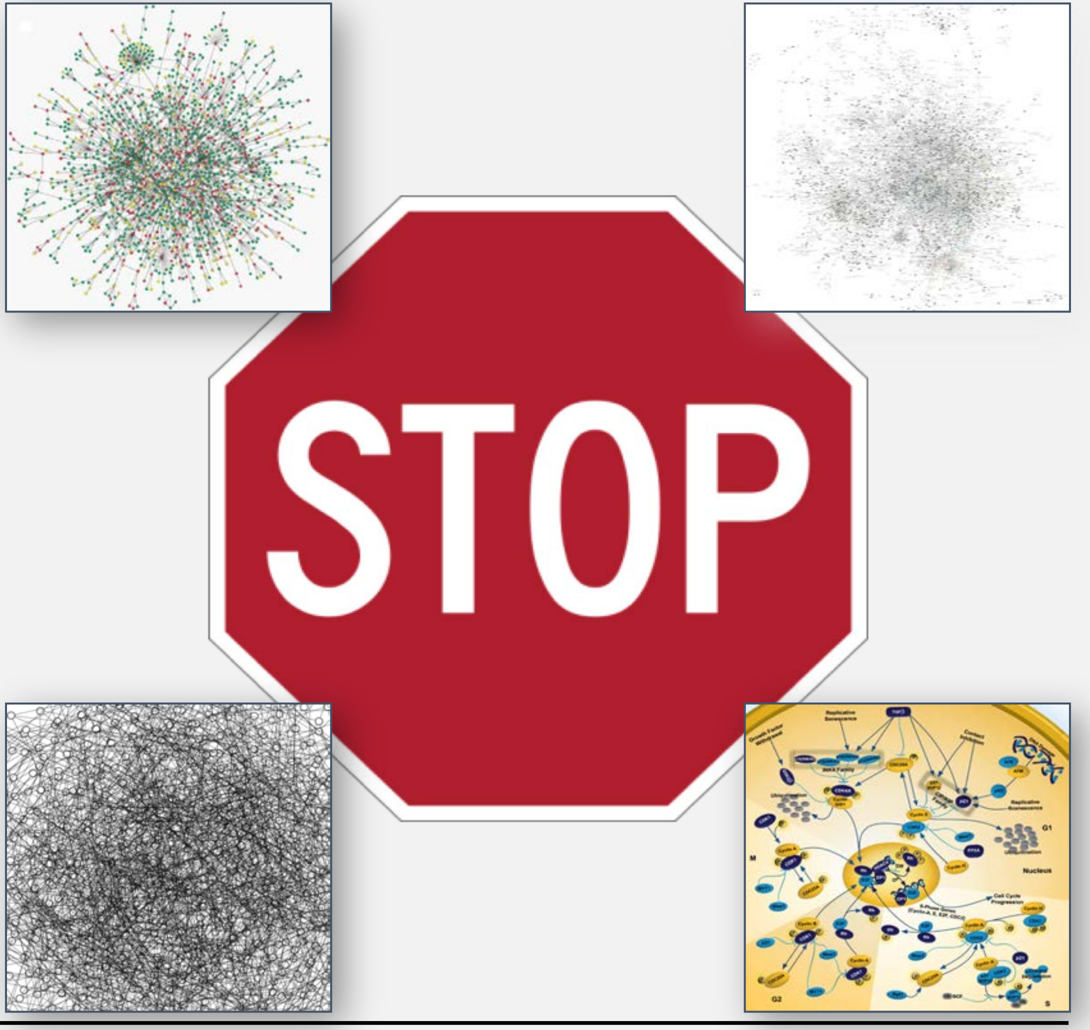
- heterogeneous membrane based on CmME
- simulated with Gromacs



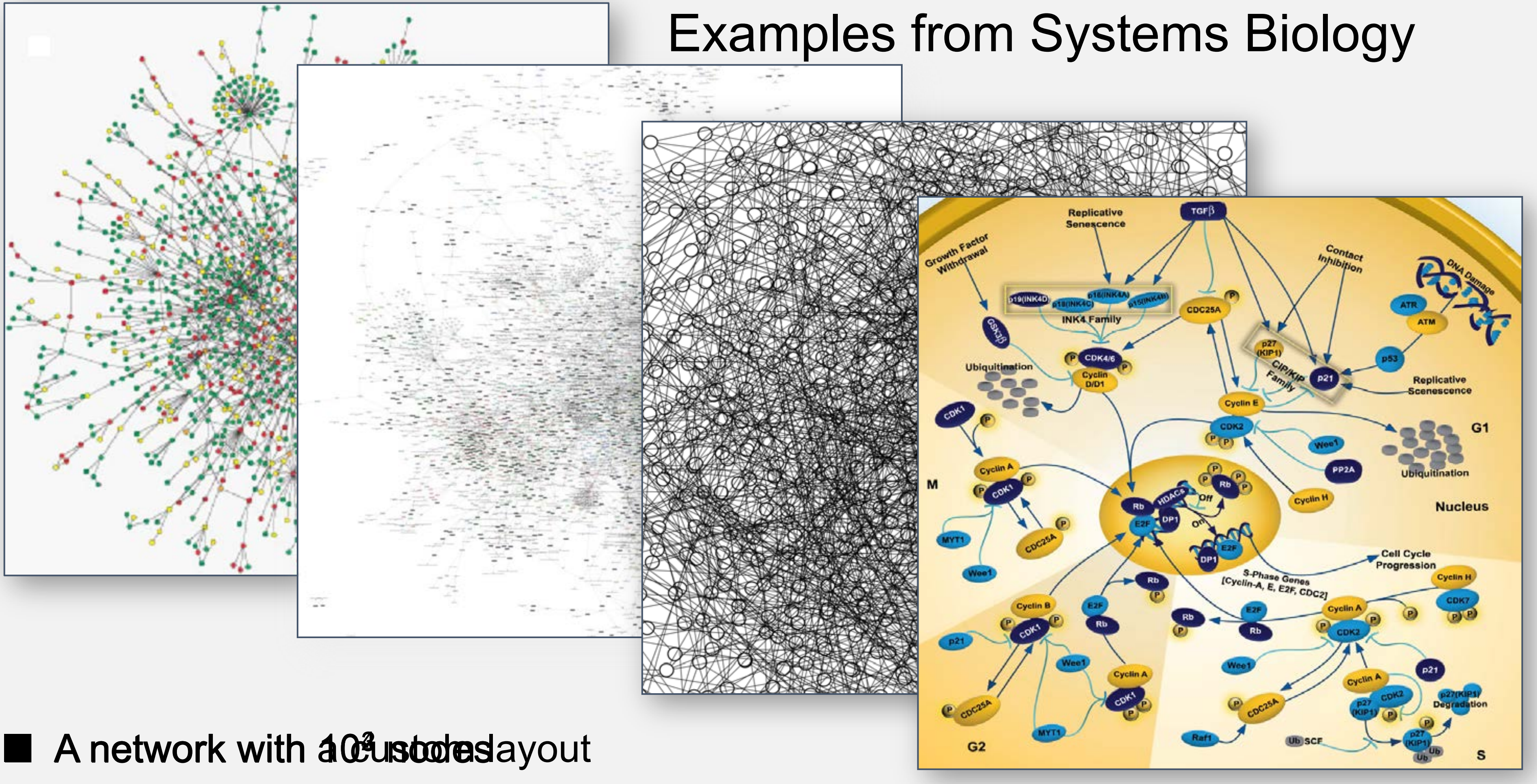
Multiple Membrane Simulations

- vary random seeds creates multiple MD simulations with one membrane configuration
- multiple runs can be used to validate the model
- example:
  - a single 100x100 Ångstrom^2 membrane
  - 1,000 ns simulation
  - 64 cores on the SCC Konstanz cluster
  - requires ca. 2 weeks
  - creates 260 GB of data to be analyzed
- then: analysis and comparison to wet lab evaluations

3. Systems Biology



Examples from Systems Biology



- A network with 10^3 nodes
- A network with a custom layout
- [illegible] Nature, 2001

Most English speaking country

Quebec

Iran

China

Israel

Singapore

Norway

Poland

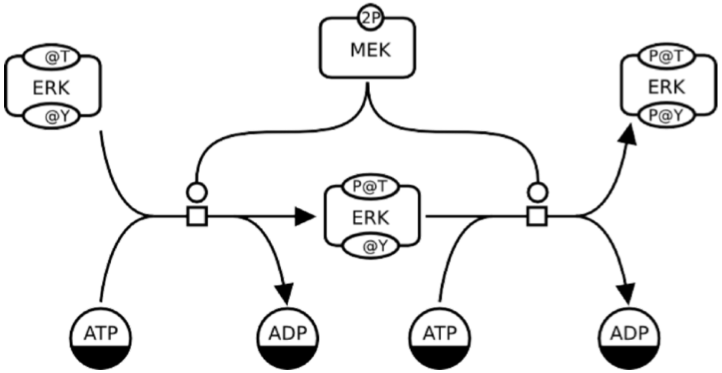
USA and Canada

- Standardized symbols as inspiration

Three languages in one:

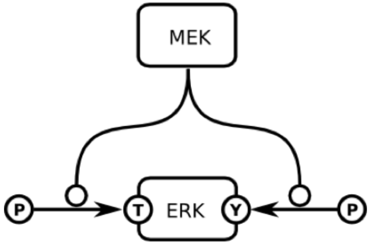
- Systems Biology Graphical Notation
- Le Novère et al. Nature Biotech., 2009

Process Description maps



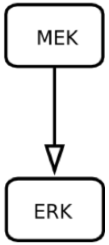
- temporal courses of biochemical interactions

Entity Relationships maps



- relationships in which a given entity participates, regardless of time

Activity Flow maps

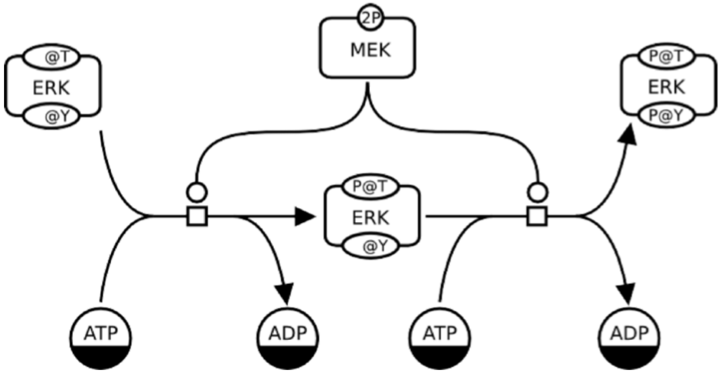


- information flow between biochemical entities in network

Three languages in one:

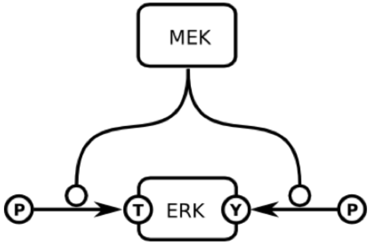
- Systems Biology Graphical Notation
- Le Novère et al. Nature Biotech., 2009

Process Description maps



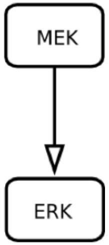
- Unambiguous
- Mechanistic
- Sequential
- Combinatorial explosion

Entity Relationships maps



- Unambiguous
- Mechanistic
- Non-Sequential

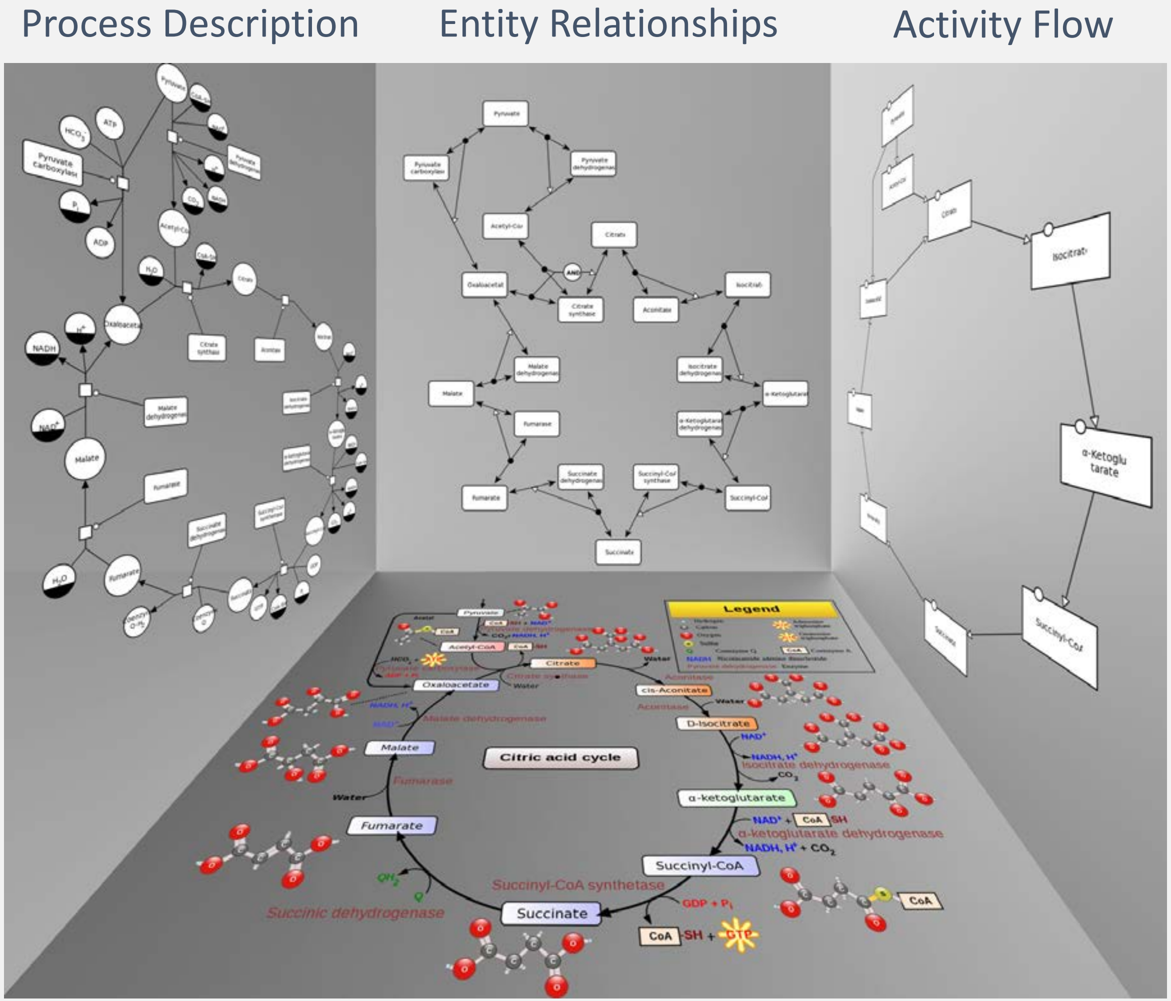
Activity Flow maps

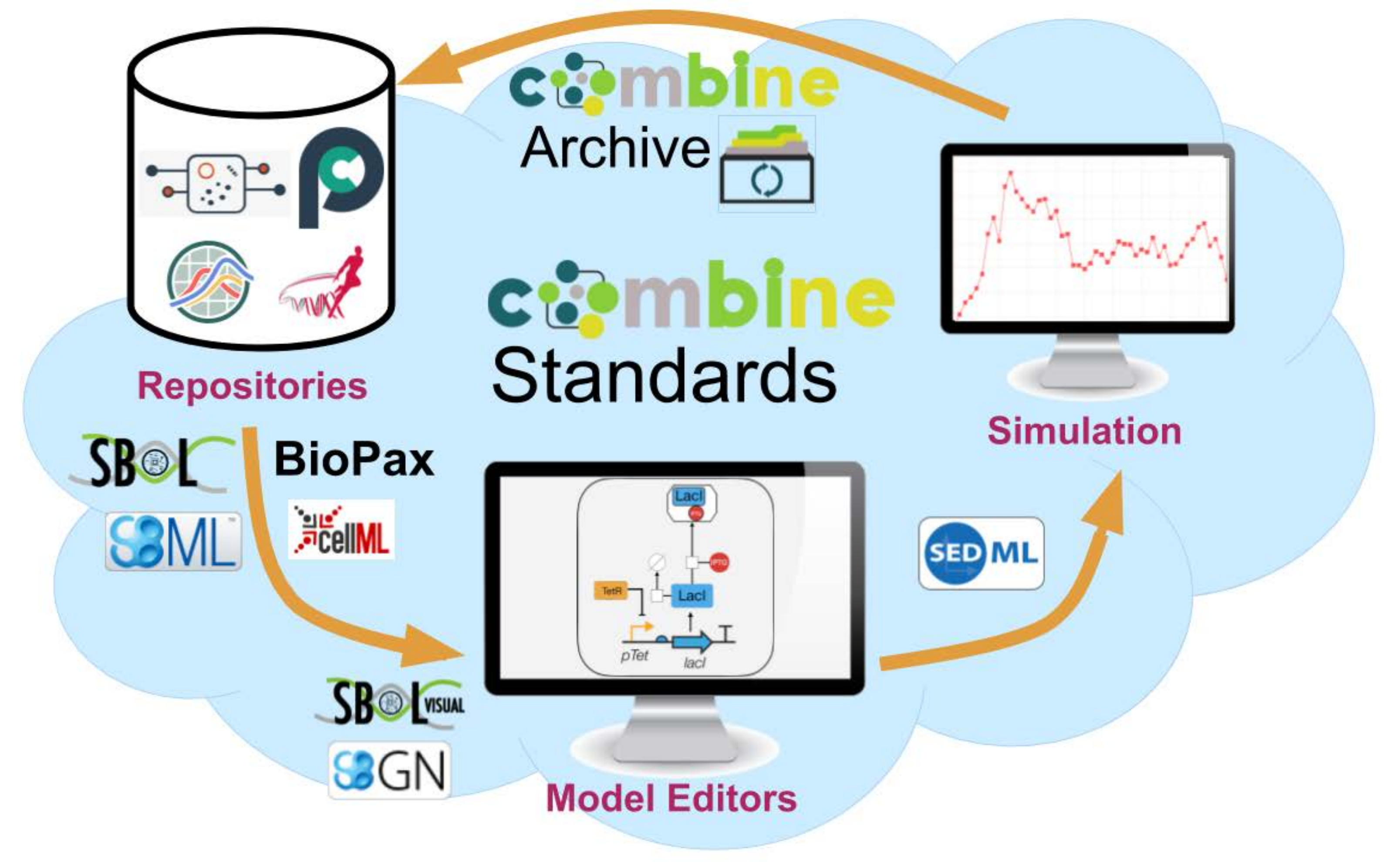


- Ambiguous
- Conceptual
- Sequential

Three languages in one:

- Unambiguously describe biochemical and cellular events in graphs
- Limited amount of symbols (~30) → smooth learning curve
- Can graphically represent quantitative models, biochemical pathways, at different granularity levels
- Developed since 2006 by a growing community, part of COMBINE




combine
Archive
combine
Standards
Repositories
SBOL
BioPax
SBML
cellML
LacI
TetR
IPTG
pTet
lacI
SED ML
Simulation
SBOL VISUAL
SBGN
Model Editors

Repositories affiliated to COMBINE standards

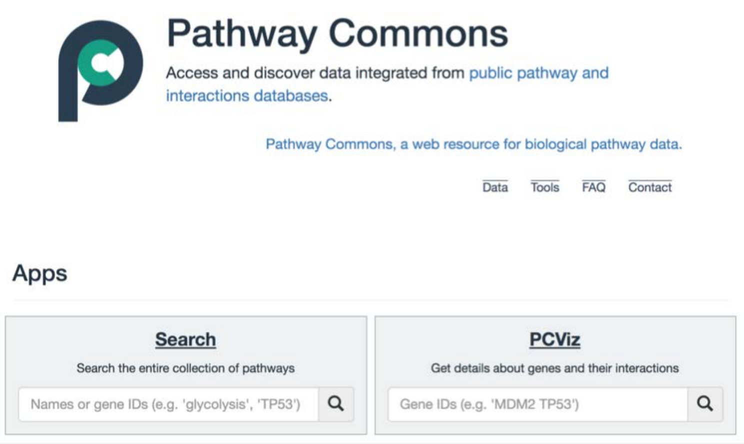
- https://pathwaycommons.org
- https://biomodels.net
- https://synbiohub.org
- https://models.physiomeproject.org

SynBioHub is a *design repository* for people designing biological constructs. It enables DNA and protein designs to be uploaded, then provides a shareable link to allow others to view them. SynBioHub also facilitates searching for information about existing useful parts and designs by combining data from a variety of sources.

Search for useful parts and designs

Upload your design for safekeeping

Share designs for publication or collaboration



Access and discover data integrated from public pathway and interactions databases.

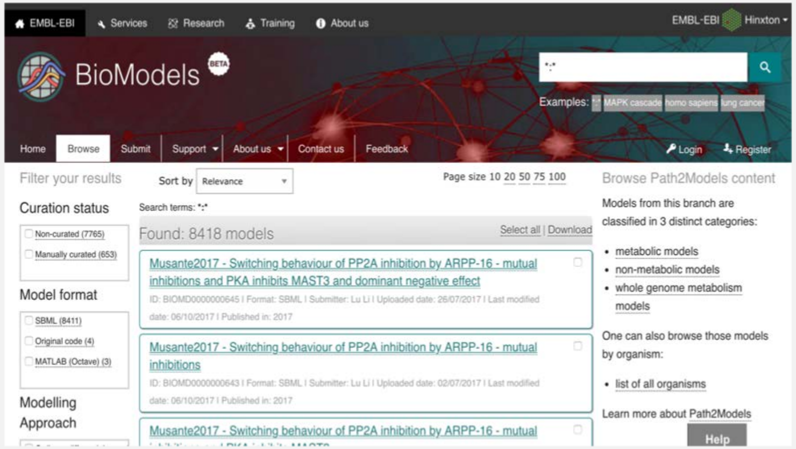
Pathway Commons, a web resource for biological pathway data.

Search

Search the entire collection of pathways

PCViz

Get details about genes and their interactions



Filter your results

Sort by: Relevance

Page size 10 20 50 75 100

Curation status

- ☐ Non-curated (7765)
- ☐ Manually curated (653)

Model format

- ☐ SBML (8411)
- ☐ Original code (4)
- ☐ MATLAB (Octave) (3)

Modelling Approach

Search terms: *:*

Found: 8418 models

Select all | Download

Musante2017 - Switching behaviour of PP2A inhibition by ARPP-16 - mutual inhibitions and PKA inhibits MAST3 and dominant negative effect

ID: BIOMD0000000645 | Format: SBML | Submitter: Lu Li | Uploaded date: 26/07/2017 | Last modified date: 06/10/2017 | Published in: 2017

Musante2017 - Switching behaviour of PP2A inhibition by ARPP-16 - mutual inhibitions

ID: BIOMD0000000643 | Format: SBML | Submitter: Lu Li | Uploaded date: 02/07/2017 | Last modified date: 06/10/2017 | Published in: 2017

Musante2017 - Switching behaviour of PP2A inhibition by ARPP-16 - mutual

Browse Path2Models content

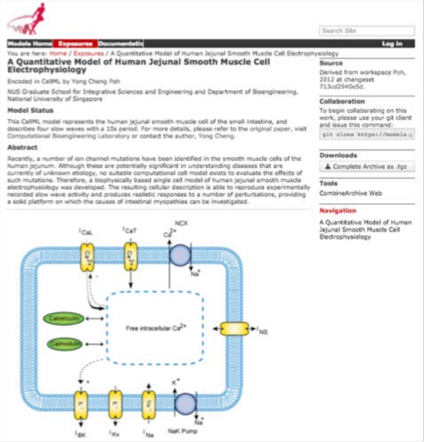
Models from this branch are classified in 3 distinct categories:

- metabolic models
- non-metabolic models
- whole genome metabolism models

One can also browse those models by organism:

- list of all organisms

Learn more about Path2Models



You are here: Home / Exposures / A Quantitative Model of Human Jejunal Smooth Muscle Cell Electrophysiology

A Quantitative Model of Human Jejunal Smooth Muscle Cell Electrophysiology

Encoded in CellML by Yong Cheng Poh

NUS Graduate School for Integrative Sciences and Engineering and Department of Bioengineering, National University of Singapore

Model Status

This CellML model represents the human jejunal smooth muscle cell of the small intestine, and describes four slow waves with a 10s period. For more details, please refer to the original paper, visit Computational Bioengineering Laboratory or contact the author, Yong Cheng.

Abstract

Recently, a number of ion channel mutations have been identified in the smooth muscle cells of the human jejunum. Although these are potentially significant in understanding diseases that are currently of unknown etiology, no suitable computational cell model exists to evaluate the effects of such mutations. Therefore, a biophysically based single cell model of human jejunal smooth muscle electrophysiology was developed. The resulting cellular description is able to reproduce experimentally recorded slow wave activity and produces realistic responses to a number of perturbations, providing a solid platform on which the causes of intestinal myopathies can be investigated.

Source

Derived from workspace Poh, 2012 at changeset 713cd2940e5d.

Collaboration

To begin collaborating on this work, please use your git client and issue this command:

`git clone https://models.`

Downloads

Tools

CombineArchive Web

Navigation

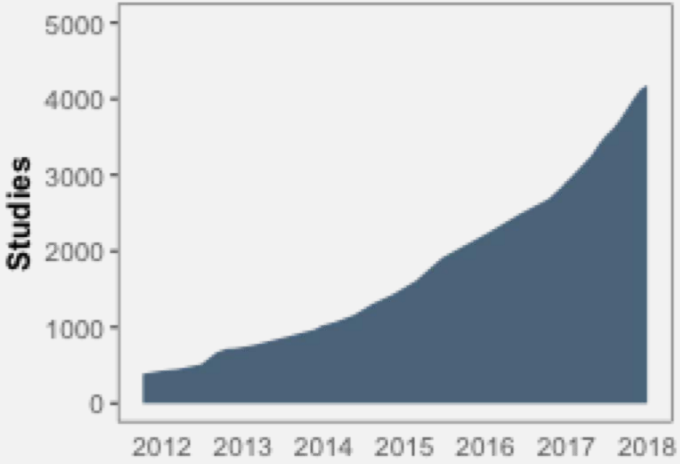
A Quantitative Model of Human Jejunal Smooth Muscle Cell Electrophysiology

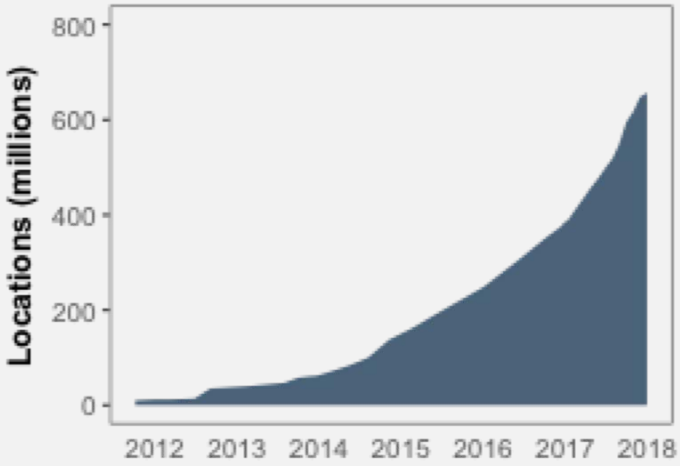
5. Collective Behaviour

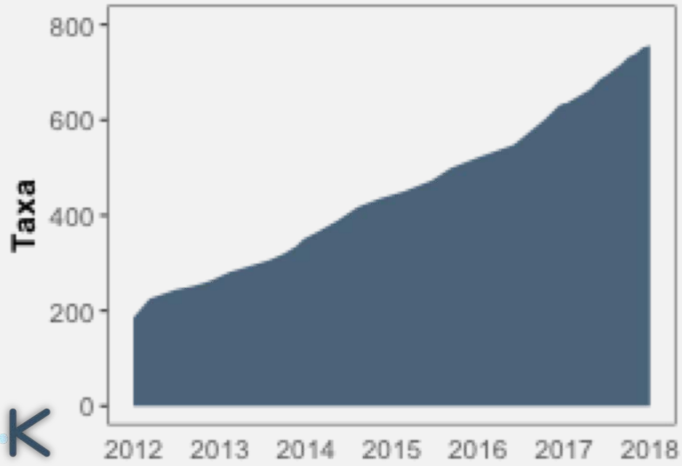
MoveBank in 2017

- over 600 million animal locations
- over 1.1 billion measurements from non-location sensors

- 4,200 studies
- 756 taxa
- 5,000 contributors
- 800 active data feeds



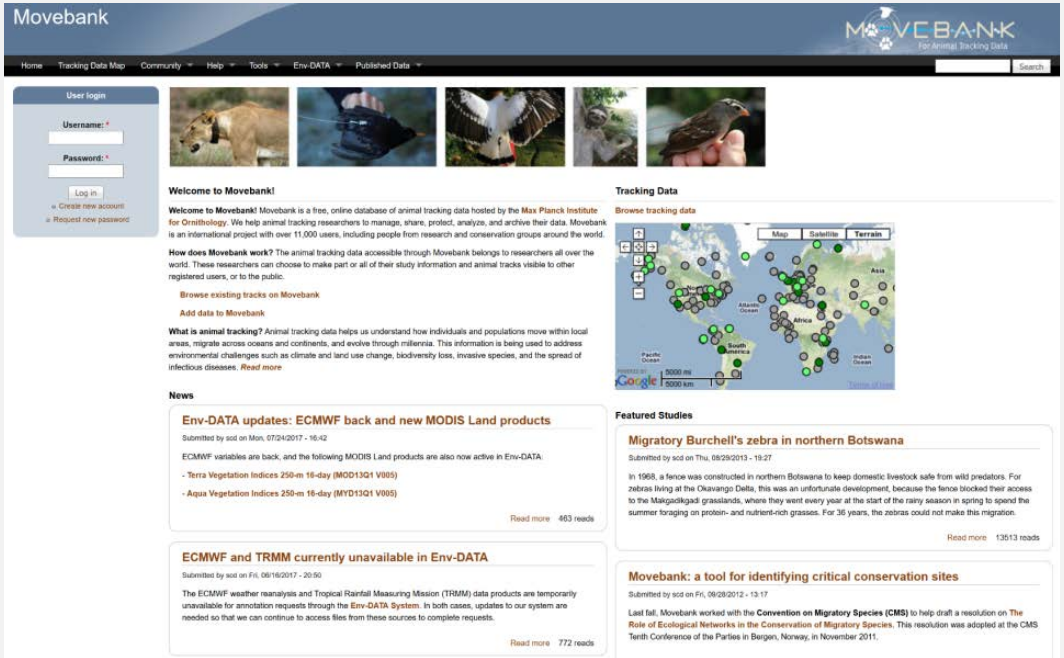




Research Questions Bird Behaviour

- How do birds move with the wind?
- Who leads a flock?
- Where and how do birds die?
- Can we predict disasters based on animal behavior?

- Next: Accumulated Bird Data over several years projected to one year



Welcome to Movebank!

Welcome to Movebank! Movebank is a free, online database of animal tracking data hosted by the **Max Planck Institute for Ornithology**. We help animal tracking researchers to manage, share, protect, analyze, and archive their data. Movebank is an international project with over 11,000 users, including people from research and conservation groups around the world.

How does Movebank work? The animal tracking data accessible through Movebank belongs to researchers all over the world. These researchers can choose to make part or all of their study information and animal tracks visible to other registered users, or to the public.

Browse existing tracks on Movebank

Add data to Movebank

What is animal tracking? Animal tracking data helps us understand how individuals and populations move within local areas, migrate across oceans and continents, and evolve through millennia. This information is being used to address environmental challenges such as climate and land use change, biodiversity loss, invasive species, and the spread of infectious diseases. *Read more*

News

Env-DATA updates: ECMWF back and new MODIS Land products

Submitted by sod on Mon, 07/24/2017 - 16:42

ECMWF variables are back, and the following MODIS Land products are also now active in Env-DATA:

- Terra Vegetation Indices 250-m 16-day (MOD13Q1 V005)
- Aqua Vegetation Indices 250-m 16-day (MYD13Q1 V005)

Read more 463 reads

ECMWF and TRMM currently unavailable in Env-DATA

Submitted by sod on Fri, 06/16/2017 - 20:50

The ECMWF weather reanalysis and Tropical Rainfall Measuring Mission (TRMM) data products are temporarily unavailable for annotation requests through the **Env-DATA System**. In both cases, updates to our system are needed so that we can continue to access files from these sources to complete requests.

Read more 772 reads

Featured Studies

Migratory Burchell's zebra in northern Botswana

Submitted by scd on Thu, 08/29/2013 - 19:27

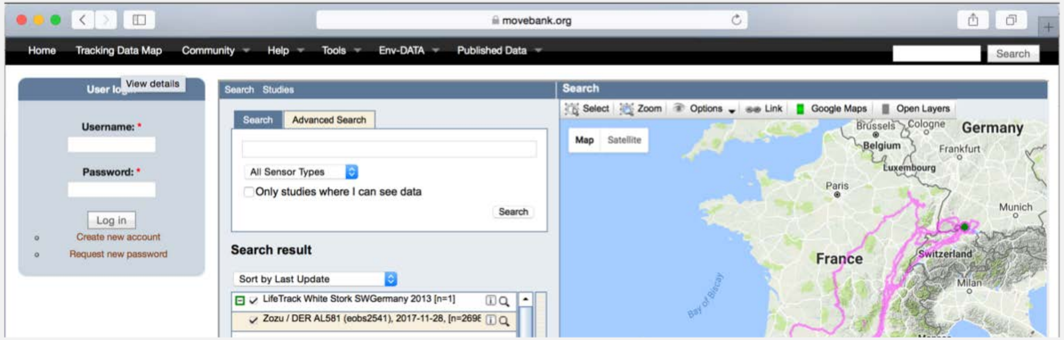
In 1968, a fence was constructed in northern Botswana to keep domestic livestock safe from wild predators. For zebras living at the Okavango Delta, this was an unfortunate development, because the fence blocked their access to the Makgadikgadi grasslands, where they went every year at the start of the rainy season in spring to spend the summer foraging on protein- and nutrient-rich grasses. For 36 years, the zebras could not make this migration.

Read more 13513 reads

Movebank: a tool for identifying critical conservation sites

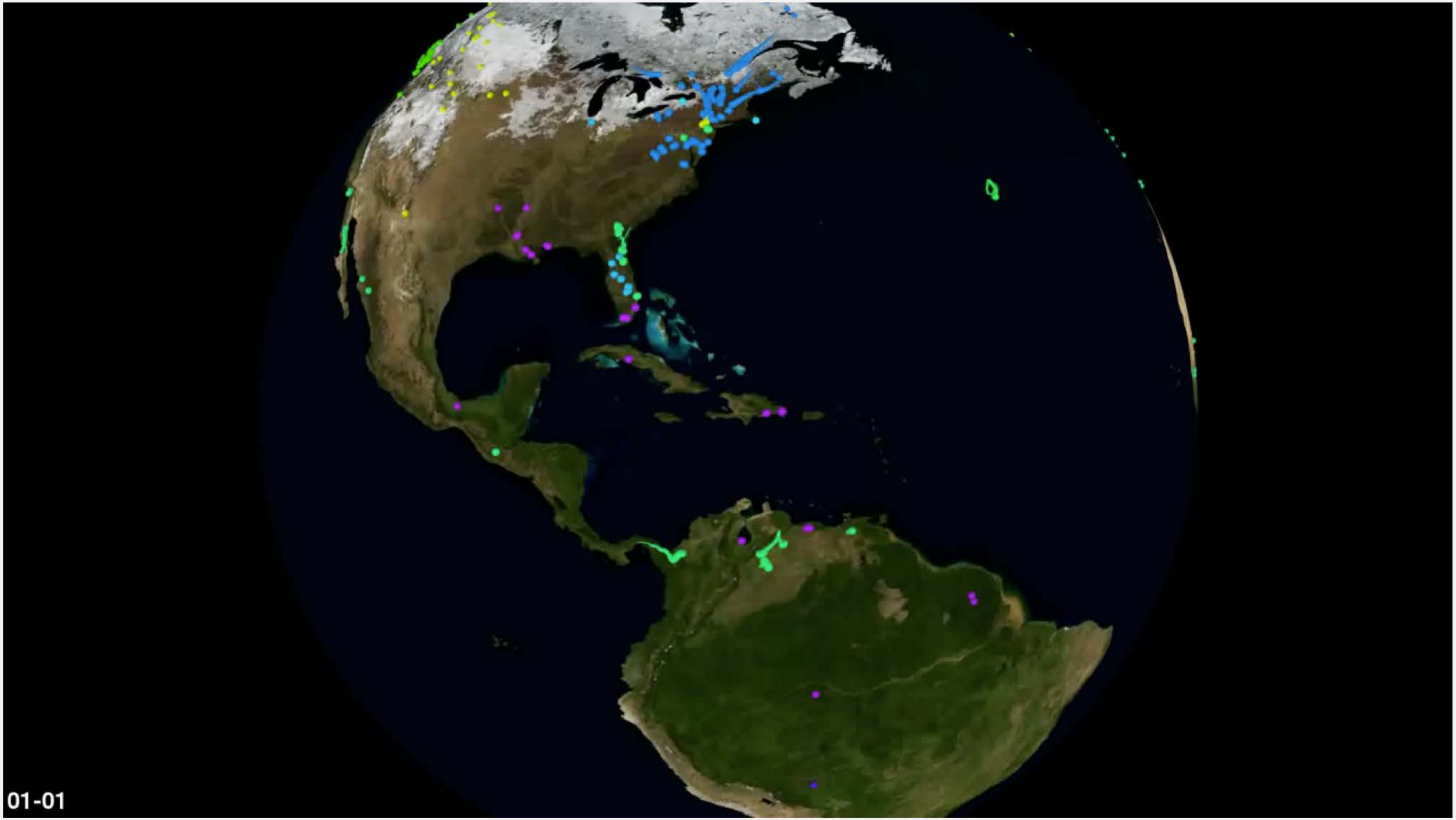
Submitted by scd on Fri, 09/28/2012 - 13:17

Last fall, Movebank worked with the **Convention on Migratory Species (CMS)** to help draft a resolution on **The Role of Ecological Networks in the Conservation of Migratory Species**. This resolution was adopted at the CMS Tenth Conference of the Parties in Bergen, Norway, in November 2011.

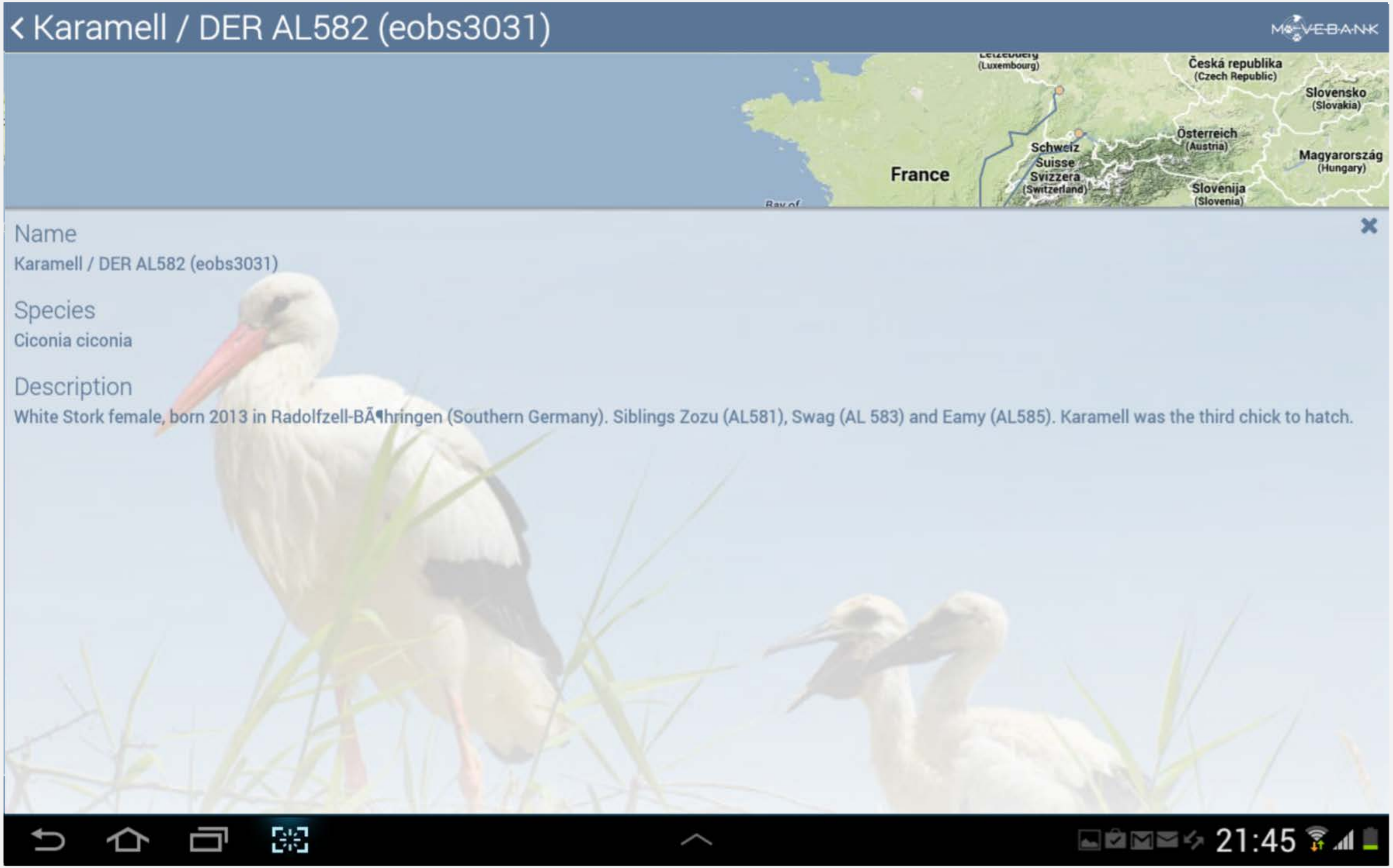


LifeTrack White Stork SWGermany 2013 [n=1]

Zozu / DER AL581 (eobs2541), 2017-11-28, [n=2696


01-01

Life of a Storch



< Karamell / DER AL582 (eobs3031)

Name

Karamell / DER AL582 (eobs3031)

Species

Ciconia ciconia

Description

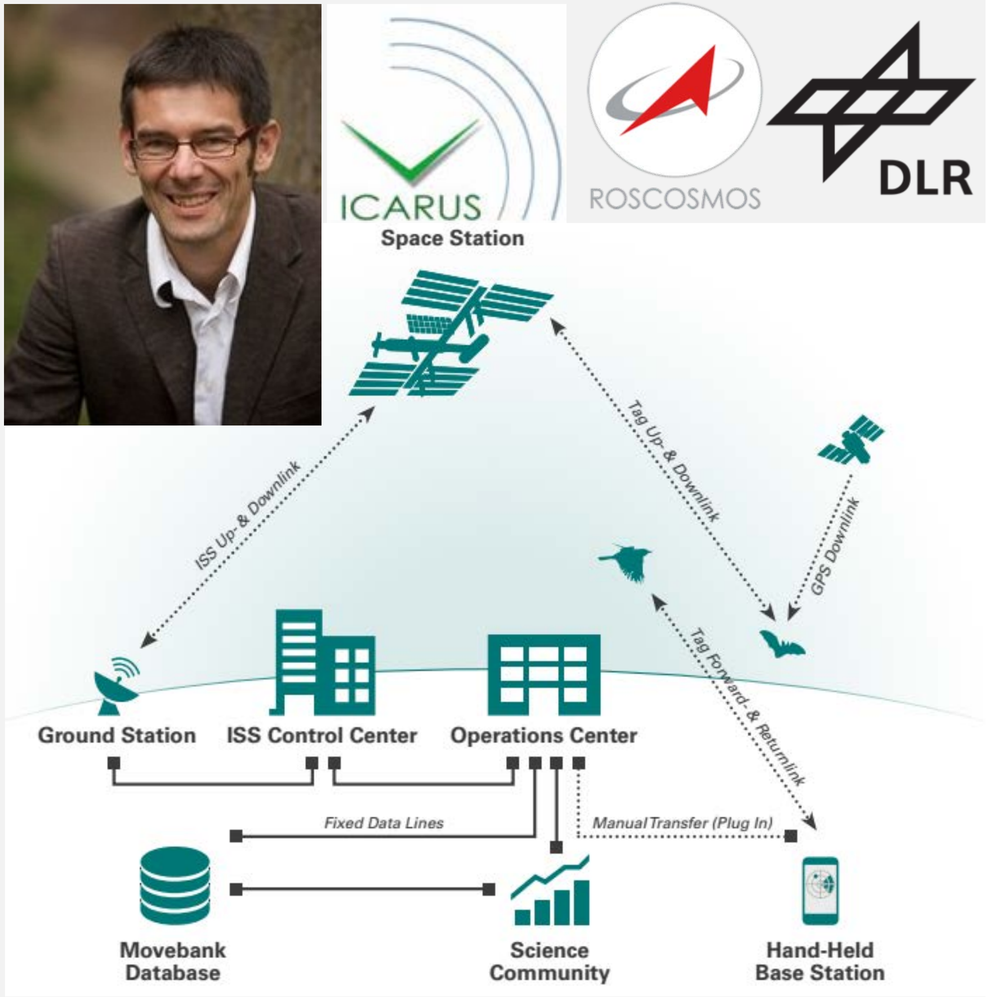
White Stork female, born 2013 in Radolfzell-BÃ¶hringen (Southern Germany). Siblings Zozu (AL581), Swag (AL 583) and Eamy (AL585). Karamell was the third chick to hatch.

Data is collected via GPS sensors

- Differ in timely resolution
  - 2 times 24h, or
  - every second
- Usually only Longitude/Latitude are recorded – Altitude is missing
- here: storchs

ICARUS

- International Cooperation for Animal Research Using Space
- Driven by Martin Wikelski et al.
- ICARUS <5-g tags including GPS, acceleration and other sensors, ground-based base stations and software supporting two-way communication
- Data will be distributed to owners via automated feeds in Movebank

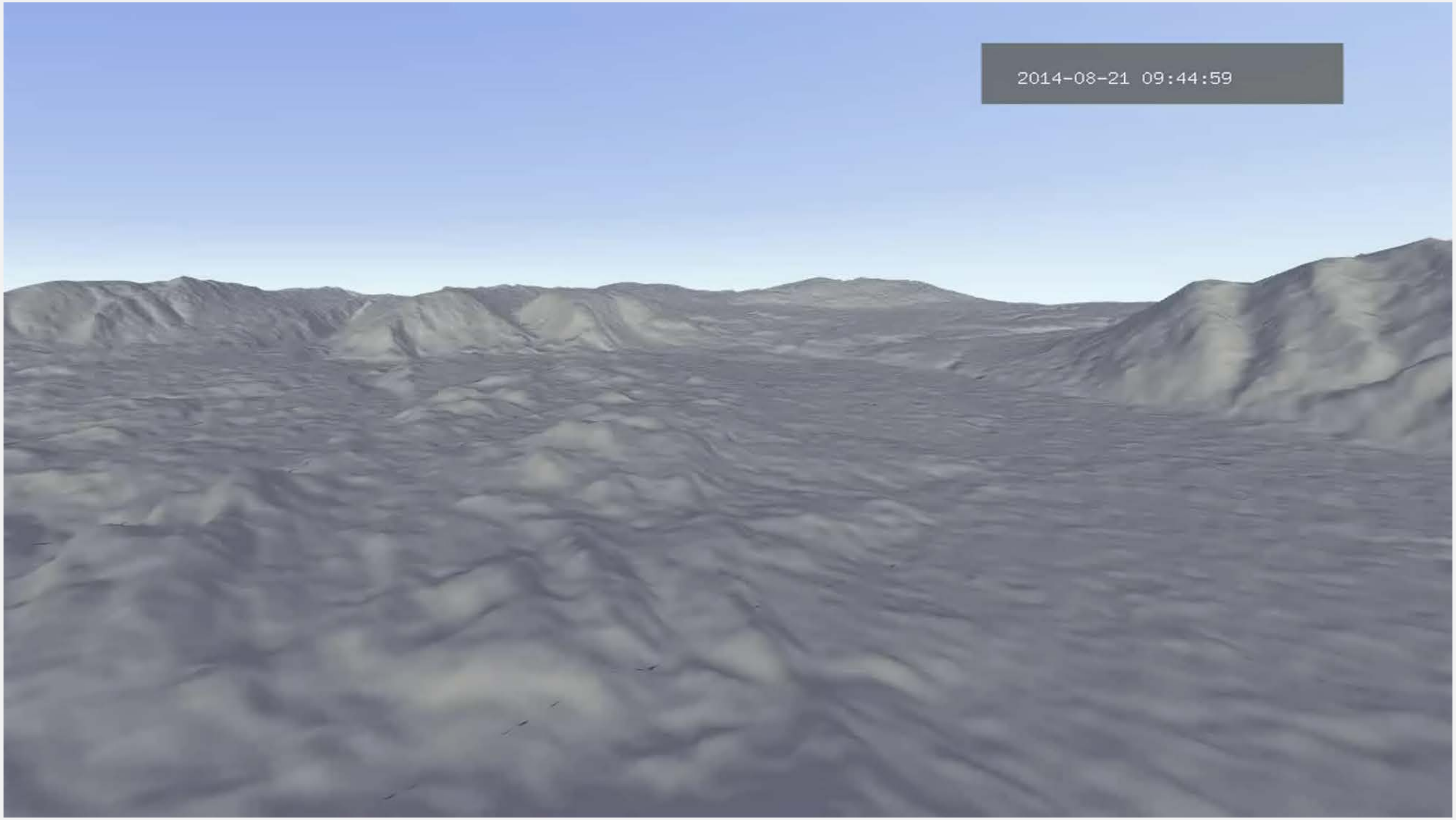


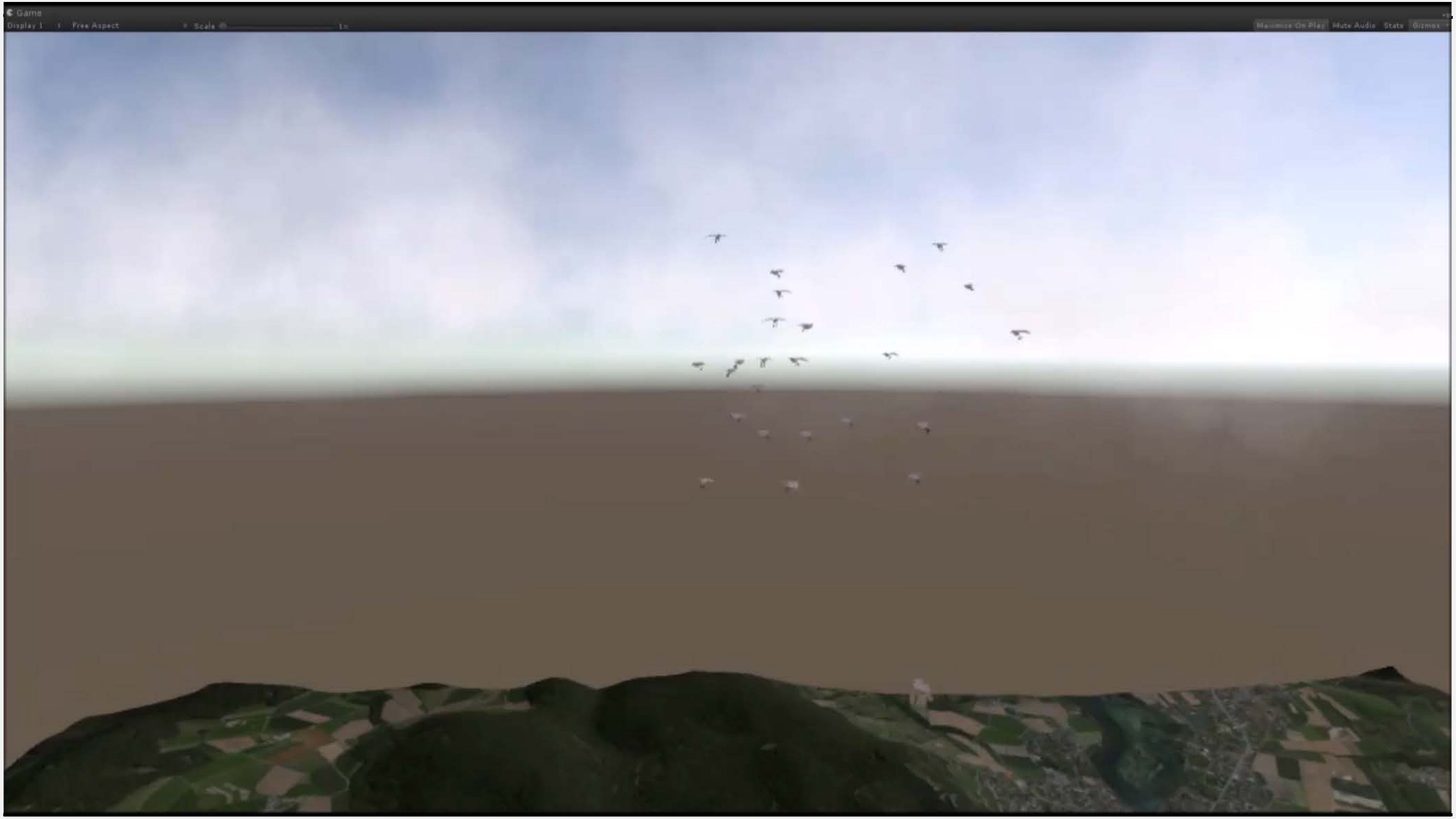
Data is collected via GPS sensors

- here: storchs

- next:
  - storch & thermals

- After next:
  - Fly with the storchs in VR!


2014-08-21 09:44:59


Game
Display 1
Free Aspect
Scale
1x
Maximize On Play
Mute Audio
Stats
Gizmos

Immersive Analytics of Collective Bird Behaviour

- Stand at BLS 6 2018 in Konstanz

BLS 6

INTERNATIONAL
BIO-LOGGING
SOCIETY

6. Acknowledgements

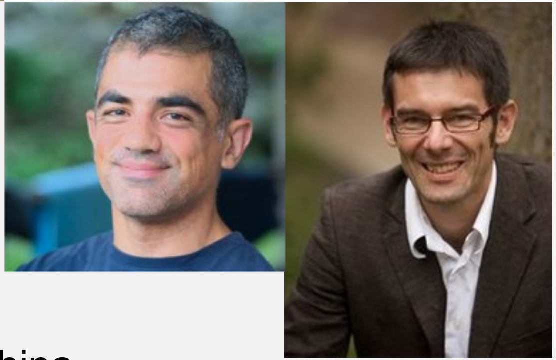
- MPI for Ornithology
  - *Martin Wikelski*
  - *Kamran Safi*
  - Mate Nagy
  - Michael Quetting
  - Andrea Flack
  - Wolfgang Fiedler
- Computational Life Sciences Konstanz
  - *Falk Schreiber*
  - *Karsten Klein*
  - Hieu Nim (Monash Uni.)
  - Our bird expert students:
    - Alexej Gluschkow
    - Artur Baltabayev
    - *Dimitar Garkov*
    - *Kim Rehberg*
    - Stefan Feyer
    - Thorsten Breitkreuz
    - *Ying Wang*

Universität Konstanz

MAX-PLANCK-GESELLSCHAFT

Max-Planck-Institut für Ornithologie

Thanks for your attention!!!



CELLmicrocosmos project

- Funding
  - DFG GK635
  - BMBF, DAAD, NNSF China
- And our many talented students over the years!
  - http://team.CELLmicrocosmos.org

6. Acknowledgements

BDVA'18

Universität
Konstanz

4th International Symposium on Big Data Visual & Immersive Analytics 2018

OCTOBER 17-19, KONSTANZ, GERMANY

General Chair
Falk Schreiber

Program Chairs
Eduard Gröller, Kim Marriott

Organising Chairs
Alexandra Diehl, Björn Sommer

Follow @BDVAsymposium

University of Konstanz