# How mathematical AI is transforming biosciences

Guo-Wei Wei

Mathematics

Michigan State University

http://www.math.msu.edu/~wei

Mathematical and Computational Biology

Jun 12 - 16, 2023

Research partnerships:

Bristol-Myers Squibb Pfizer NVIDIA FDA

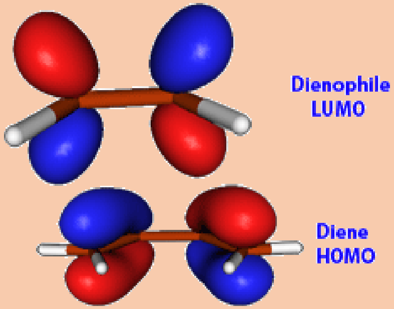
Grant support: NIH, NSF, NASA, Pfizer, BMS, MEDC, Georgetown U, COVID-19 HPC Consortium, MSU Foundation, and MSU iCER

# Four paradigms of scientific research

1^St^ Paradigm:
Empirical sciences

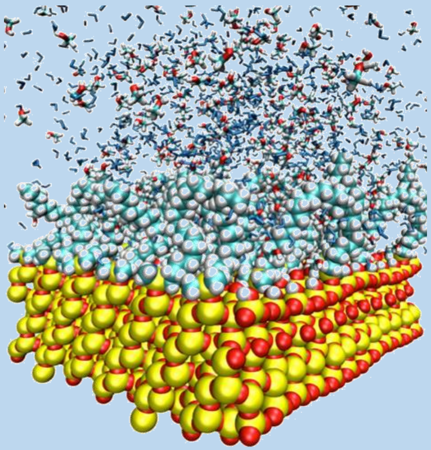
**Experiments**

2^nd^ Paradigm:
Model-based theoretical sciences



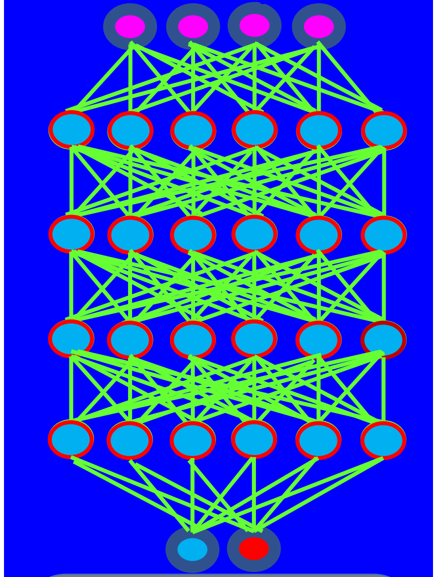
**Math/Phys models**

3^rd^ Paradigm:
Computational sciences



**Computing, simulation, algorithms**

4^th^ Paradigm:
Data-driven scientific discovery



**AI, machine learning, data science**

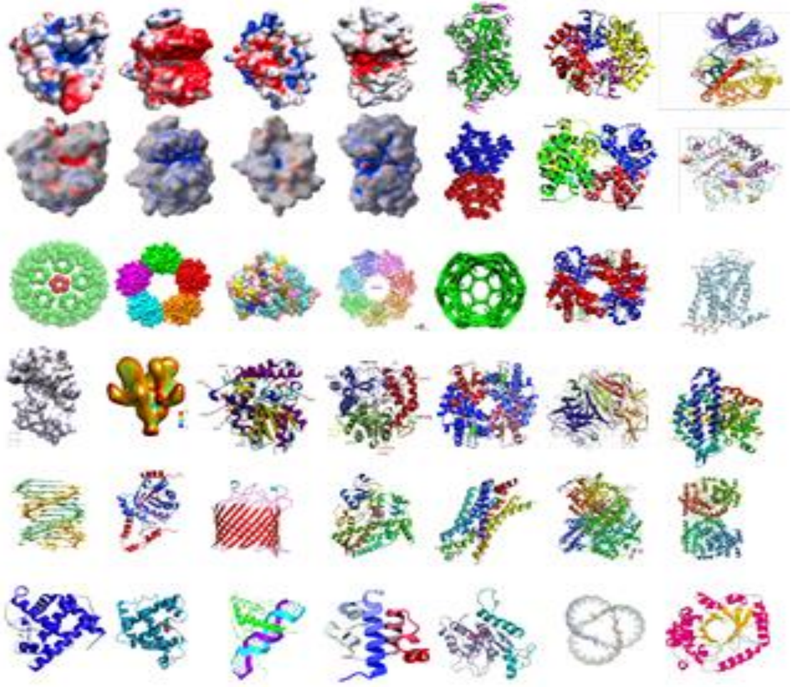
1600

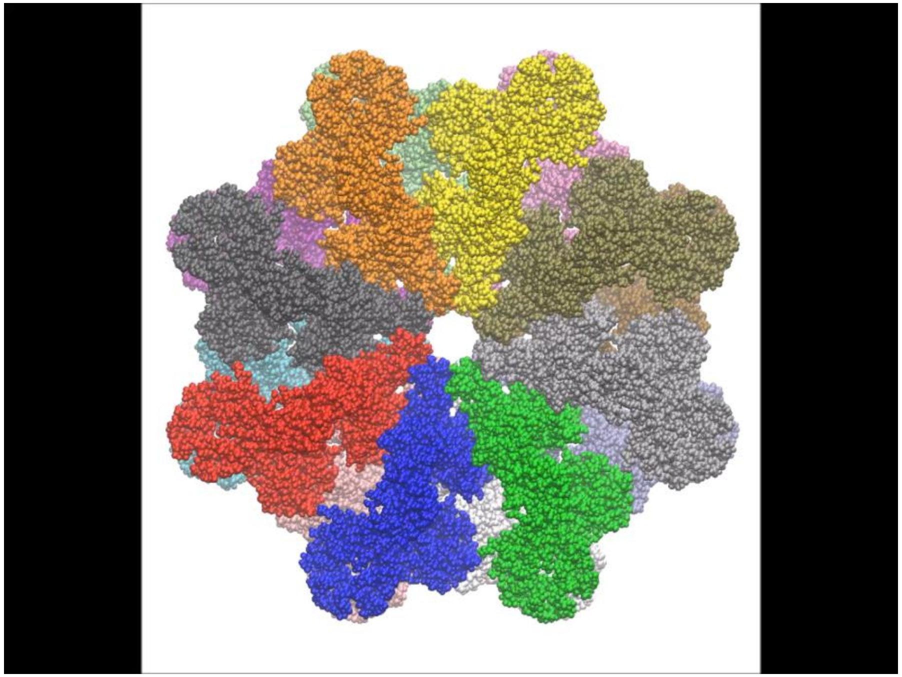
1950

2000

# Challenges of AI in biomolecular systems

- ***Geometric dimensionality***: $\mathbb{R}^{3N}$, **where *N* ~5000 for a protein.**
- ***Machine learning dimensionality***: **> $1024^3$*m*, where *m* is the number of atom types in a protein.**
- ***Non-scalability***: **different sizes.**
- ***Complexity***: **intermolecular & intramolecular interactions.**





# Two schools of thinking

Given a protein with *N* atoms and an average of *n* electrons in each atom

Fundamentalism; Mechanistic

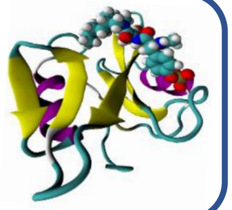
**Quantum Mechanics** $\mathbb{R}^{3Nn+3N}$

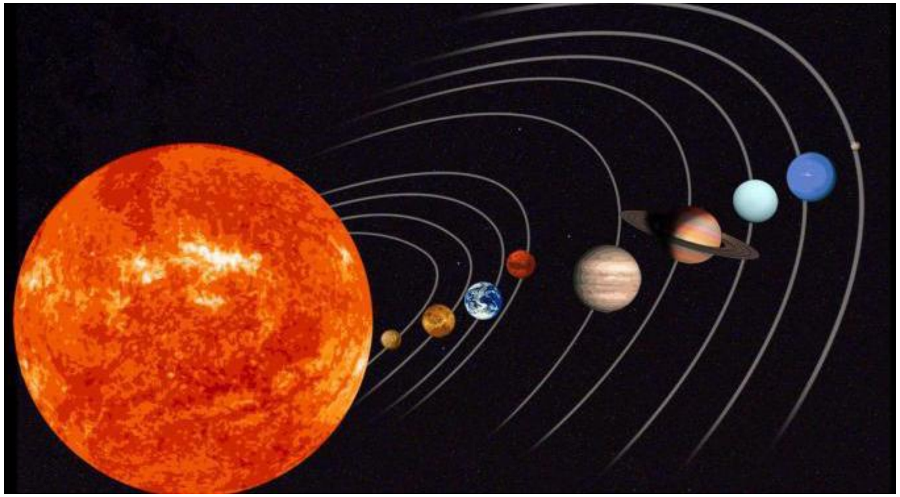
**QM/MM** $\mathbb{R}^{K}$ $3N< K <3N(n+1)$

**Molecular Mechanics** $\mathbb{R}^{3N}$

**Multiscale Coarse-grain** $\mathbb{R}^{M}$ $(3<M<3N)$

**Poisson-Boltzmann, PNP, etc.** $\mathbb{R}^{3}$





**Differentiable Manifold** $\mathbb{R}^{2}$

**Algebraic Topology** $\mathbb{R}^{1}$

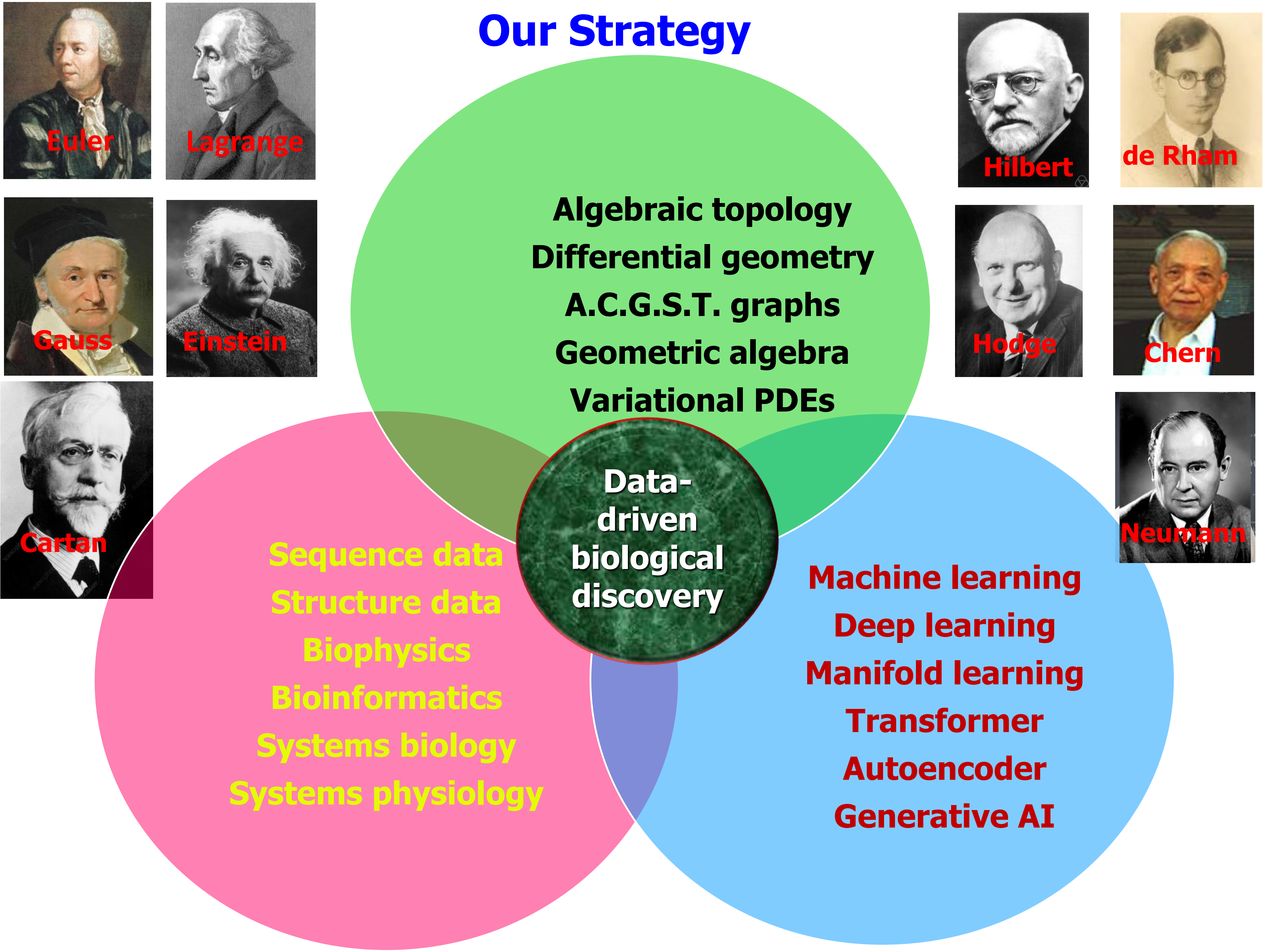
**Graph Theory** $\mathbb{R}^{0}$

**Index Theory** $\mathbb{R}^{0}$

Reductionism; Data-driven

Basic hypothesis: Intrinsic physics lies on low-dimensional manifolds in a high dimensional space

# Our Strategy



Algebraic topology
Differential geometry
A.C.G.S.T. graphs
Geometric algebra
Variational PDEs

Data-driven biological discovery

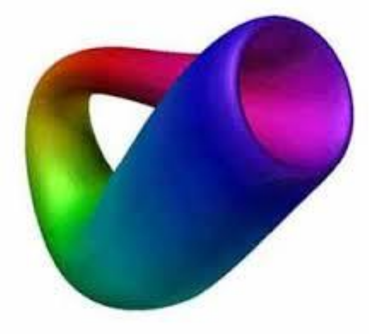
Sequence data
Structure data
Biophysics
Bioinformatics
Systems biology
Systems physiology

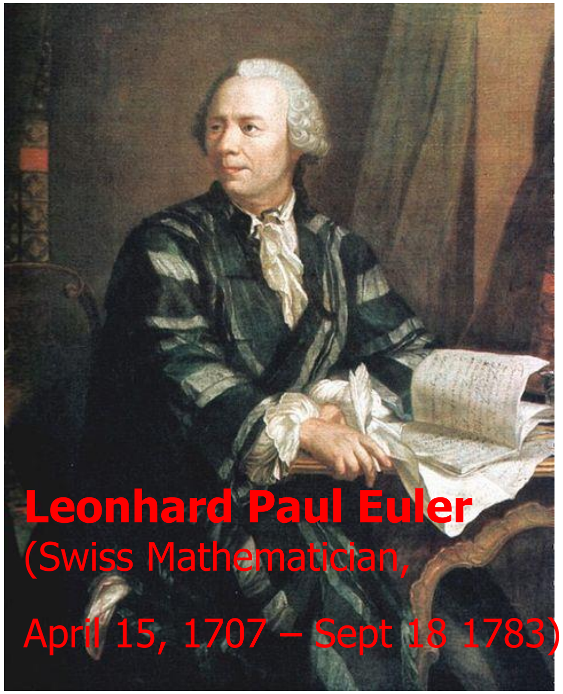
Machine learning
Deep learning
Manifold learning
Transformer
Autoencoder
Generative AI

# Topology

Möbius Strips (1858)

Klein Bottle (1882)

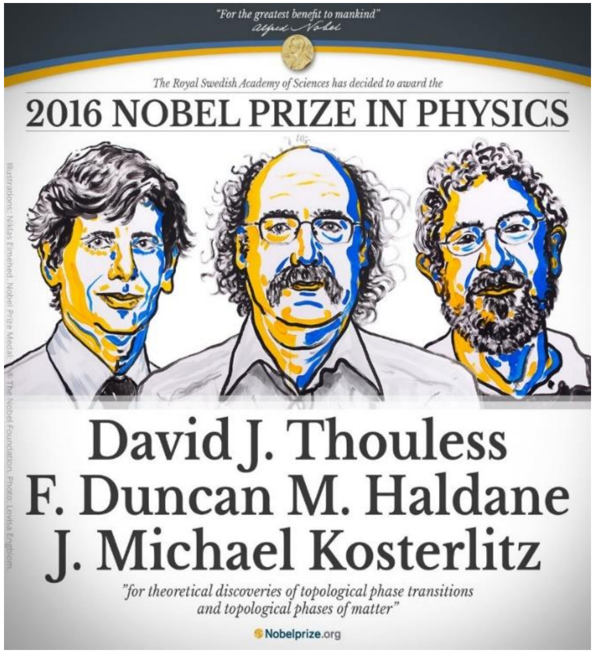




**Leonhard Paul Euler**
(Swiss Mathematician,
April 15, 1707 – Sept 18 1783)

Torus

Double Torus



*"For the greatest benefit to mankind"*
*Alfred Nobel*

*The Royal Swedish Academy of Sciences has decided to award the*

2016 NOBEL PRIZE IN PHYSICS

David J. Thouless
F. Duncan M. Haldane
J. Michael Kosterlitz

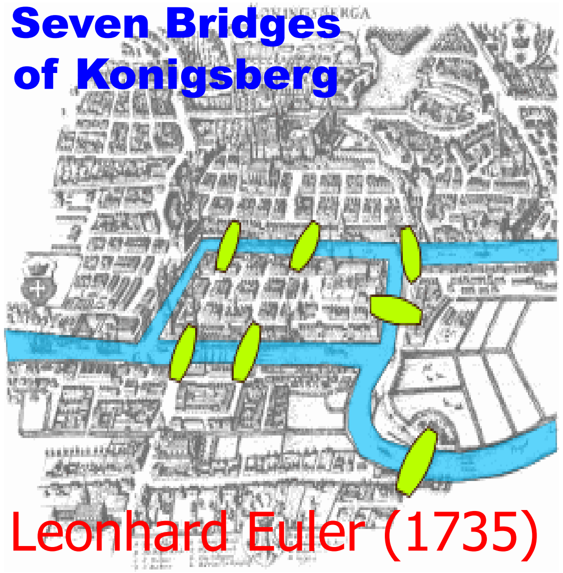
*"for theoretical discoveries of topological phase transitions and topological phases of matter"*

Nobelprize.org

Illustrations: Niklas Elmehed. Nobel Prize Medal®. The Nobel Foundation. Photo: Lovisa Engblom.

Augustin-Louis Cauchy,
Ludwig Schläfli,
Johann Benedict Listing,
Bernhard Riemann, and
Enrico Betti

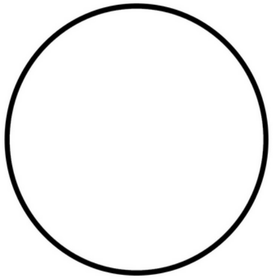
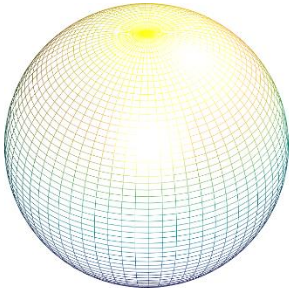
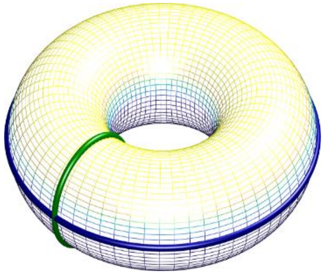
**Seven Bridges of Konigsberg**



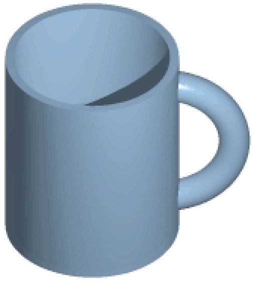
Leonhard Euler (1735)

# Topological invariants: Betti numbers

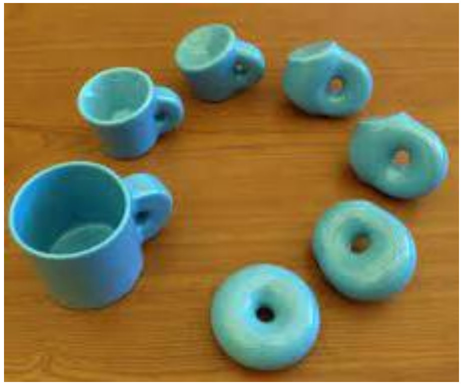
$\beta_0$ **is the number of connected components.**
$\beta_1$ **is the number of tunnels or circles.**
$\beta_2$ **is the number of cavities or voids.**

| **Point** | **Circle** | **Sphere** | **Torus** |
|---|---|---|---|
| |  |  |  |
| $\beta_0 = 1$ | $\beta_0 = 1$ | $\beta_0 = 1$ | $\beta_0 = 1$ |
| $\beta_1 = 0$ | $\beta_1 = 1$ | $\beta_1 = 0$ | $\beta_1 = 2$ |
| $\beta_2 = 0$ | $\beta_2 = 0$ | $\beta_2 = 1$ | $\beta_2 = 1$ |

**Limitation**



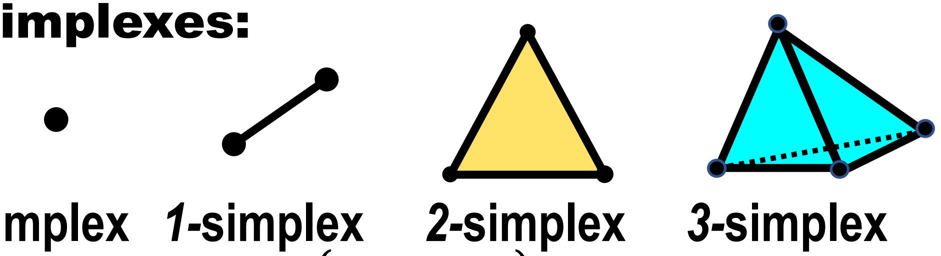
L. Vieira



Crane and Segerman

# Persistent homology induced by filtration

**Simplexes:**



Frosini and Nandi (1999), Robins (1999), Edelsbrunner, Letscher and Zomorodian (2002), Zomorodian and Carlsson (2005), Edelsbrunner and Harer, (2007) Kaczynski, Mischaikow and Mrozek (2004), Ghrist (2008), …

**k-chain:** $K = \left\{ \sum_j c_j \sigma_j^q \right\}$

**Chain group:** $C_q(K, \mathbb{Z}_2)$

**Boundary operator:**

$$\partial_q \sigma^q = \sum_{j=0}^{q} (-1)^j \{v_0, v_1, \dots, \widehat{v_j}, \dots, v_k\}$$

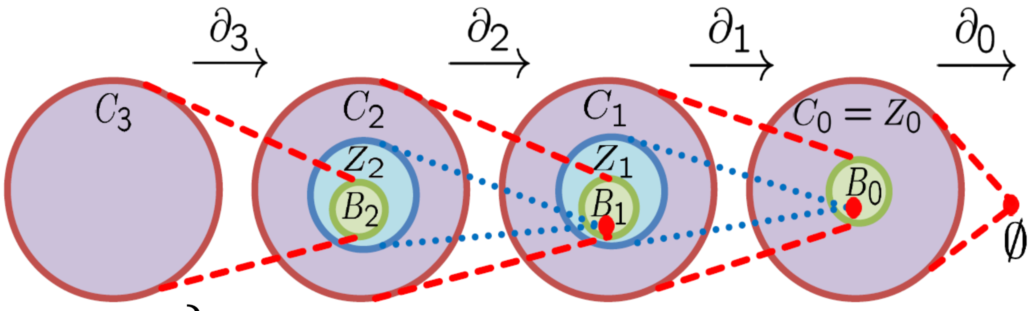
**Cycle group:** $Z_q = \operatorname{Ker} \partial_q$

**Boundary group:** $B_q = \operatorname{Im} \partial_{q+1}$

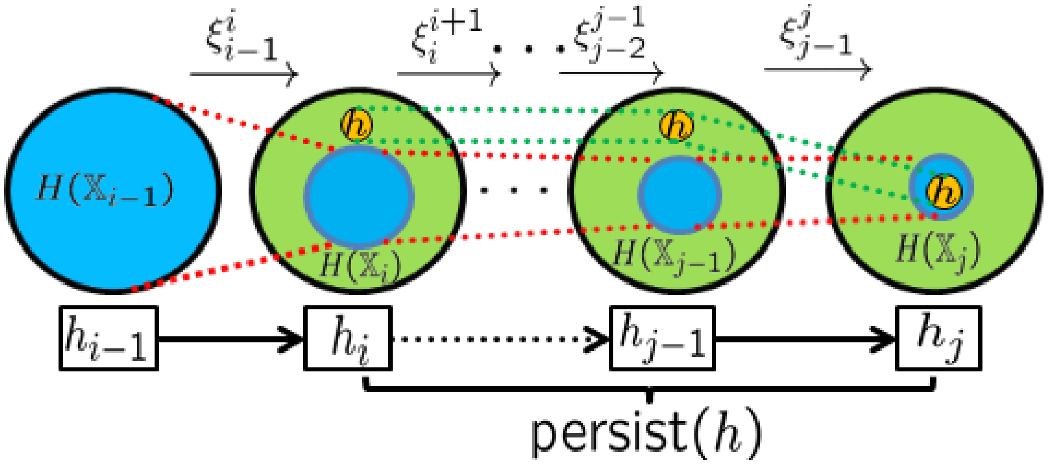
**Homology group:** $H_q = Z_q / B_q$

**Betti number:** $\beta_q = \operatorname{Rank}(H_q)$

Xia, Wei, IJNMBE, 2014;
Xia, Feng, Tong, Wei, JCC, 2015


$\xrightarrow{\partial_3}$
$\xrightarrow{\partial_2}$
$\xrightarrow{\partial_1}$
$\xrightarrow{\partial_0}$
$C_3$
$C_2$
$Z_2$
$B_2$
$C_1$
$Z_1$
$B_1$
$C_0 = Z_0$
$B_0$
$\emptyset$

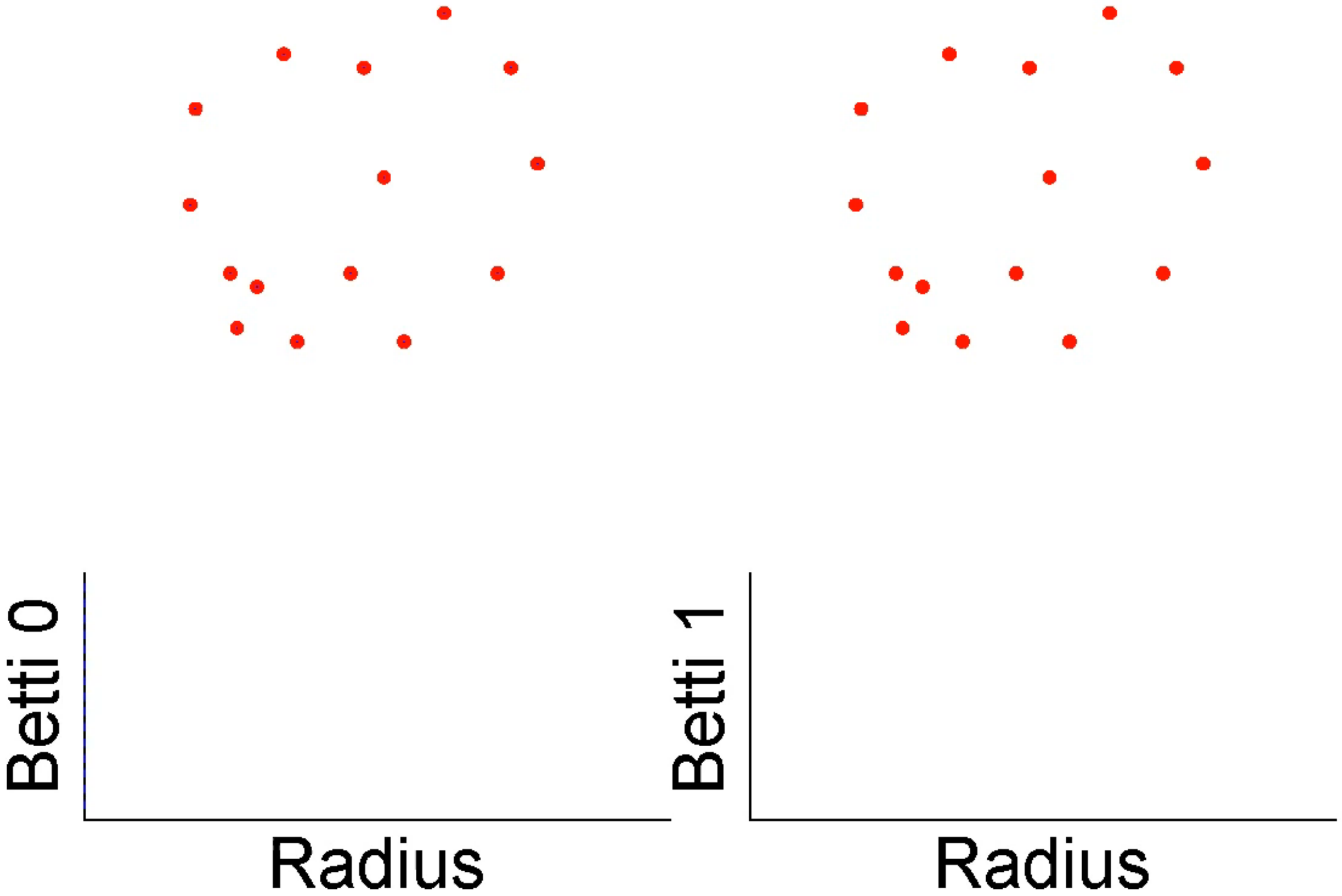
Filtration:



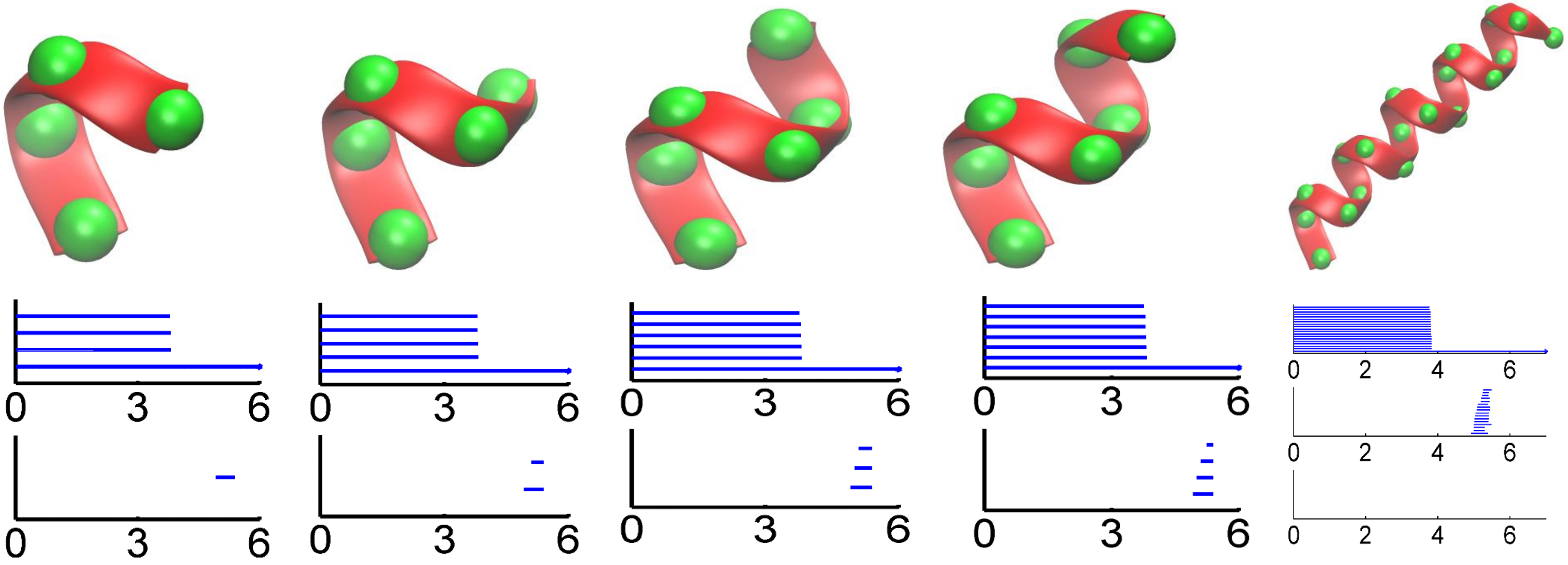
# Topological data analysis

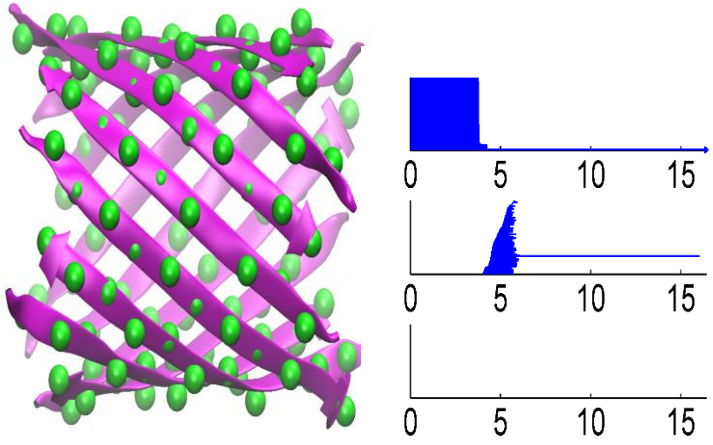
Vietoris-Rips complexes, persistent homology and topological fingerprint

(Xia, Wei, 2014)

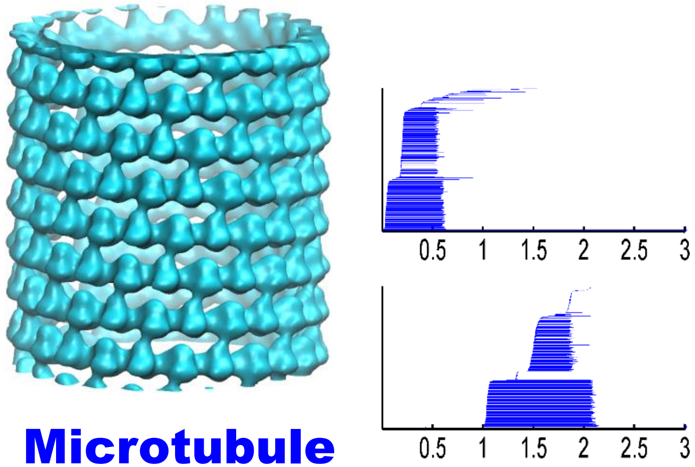


# Topological fingerprints of an alpha helix, beta barrel, etc.





**Beta barrel**

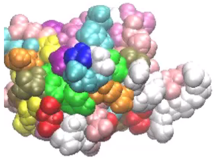


**Microtubule**

**(Xia & Wei, IJNMBE, 2014, 2015)**

# Topological data analysis

## 2D persistent homology of protein unfolding (1UBQ)



$\mathbb{R}^{3N+1} \to \mathbb{R}^2$

Kelin Xia

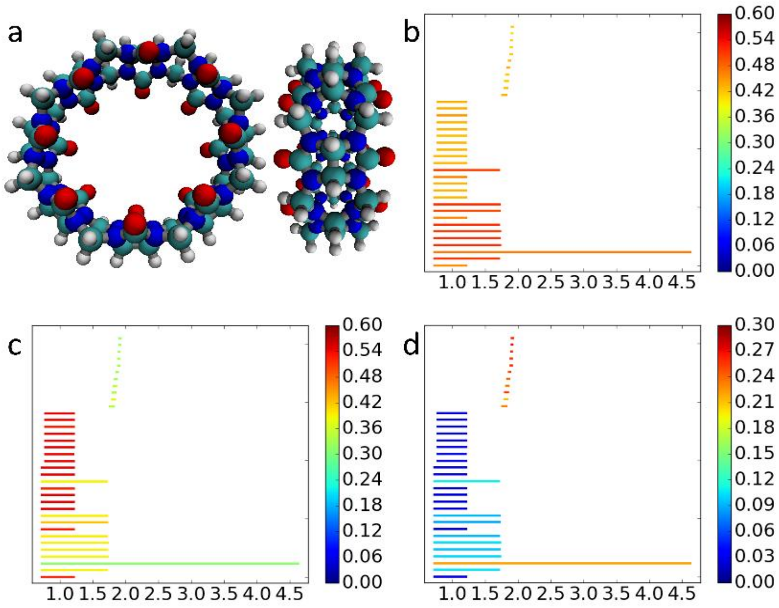
(Xia & Wei, JCC, 2015)

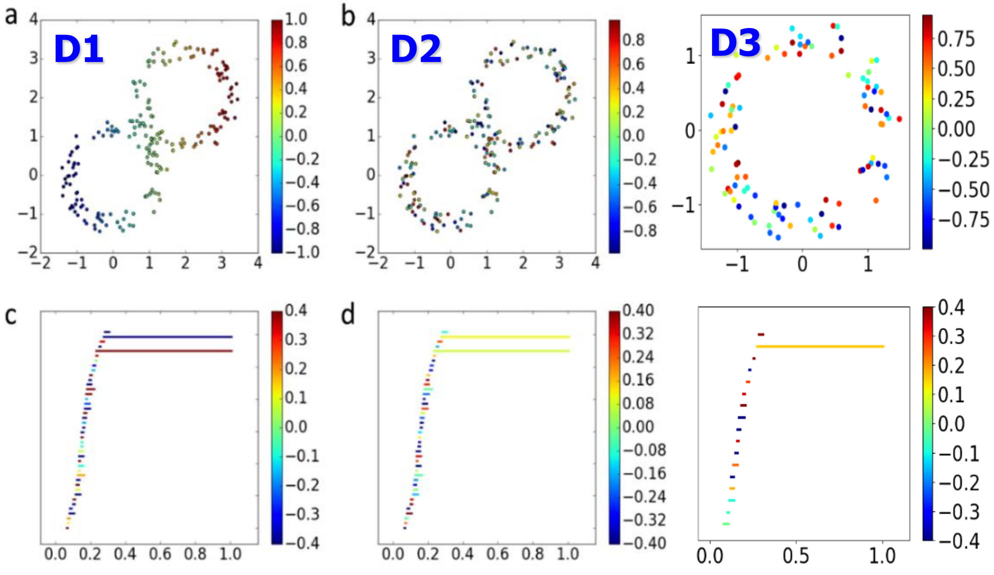
# Limitations of persistent homology that prevent it from working well for many data

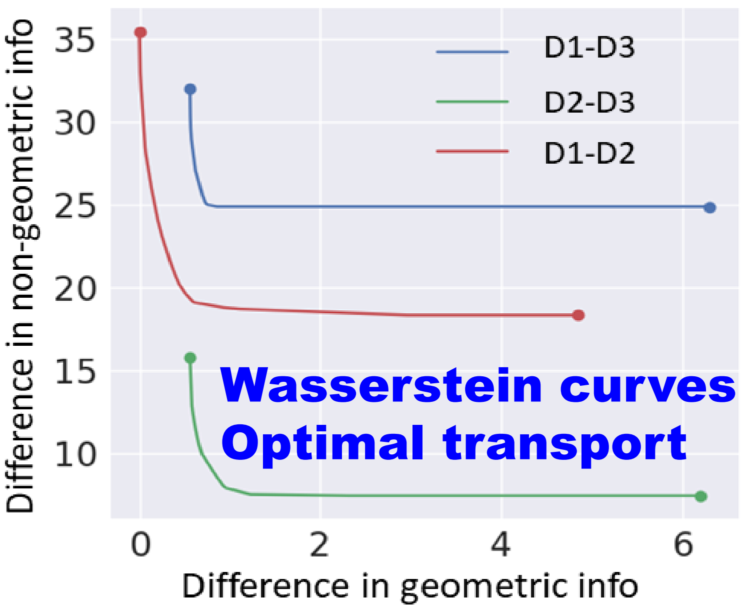
- It cannot handle heterogeneous information (i.e., different type of objects in the data)
- It is qualitative rather than quantitative (e.g., a 5-member ring is counted the same as a 6-member ring)
- It cannot describe non-topological changes (i.e., homotopic shape evolution over filtration)
- It is incapable of dealing with directed networks and digraphs (polarization, regulation, control issues)
- It is unable to characterize structured data (e.g., hypergraphs, directed networks)

We address these limitations with new topological methods

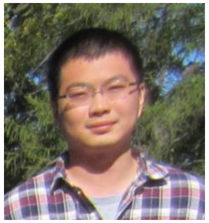
# Persistent cohomology for heterogeneous data



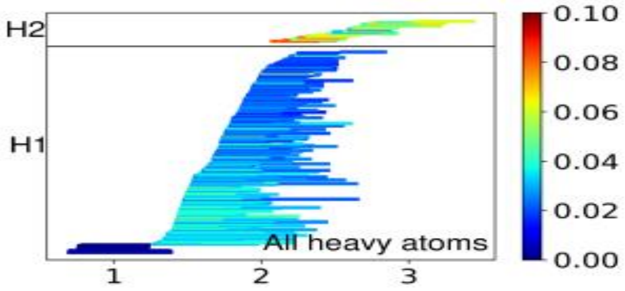


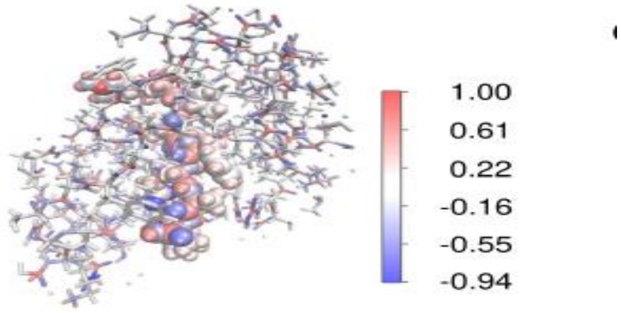


**Wasserstein curves**
**Optimal transport**



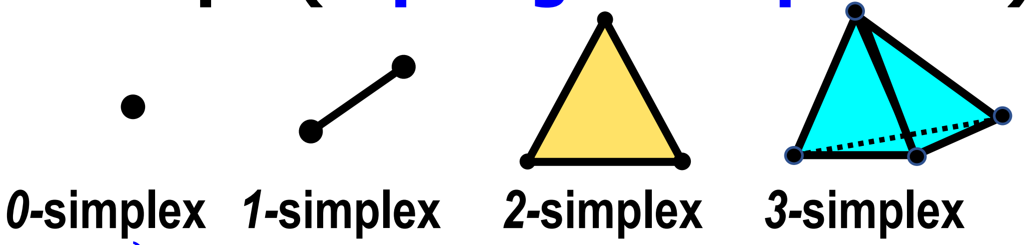
Zixuan Cang
And Wei, SIAM
JMDS 2020





# Combinatorial Graph (topological Laplacian)

- Simplexes ($\sigma^q$):



- $K$-chain: $K = \left\{ \sum_j w_j \sigma_j^q \right\}$

(Eckmann 1944; Goldberg 2002; Horak, Jost, AIM, 2013; Serrano, Gomze, 2019,…)

- Chain group: $C_q(K, \mathbb{Z}_2)$
- Boundary operator: $\partial_q : C_q(K) \to C_{q-1}(K)$

$$\partial_q \sigma^q = \sum_{j=0}^{q} (-1)^j \{v_0, v_1, \dots, \widehat{v_j}, \dots, v_q\}$$

- Adjoint boundary operator: $\partial_q^* : C_{q-1}(K) \to C_q(K)$
- $q$-combinatorial Laplacian operator: $\Delta_q = \partial_{q+1}\partial_{q+1}^* + \partial_q^* \partial_q$
- $q$-combinatorial Laplacian matrix: $\mathcal{L}_q = \mathcal{B}_{q+1}\mathcal{B}_{q+1}^T + \mathcal{B}_q^T \mathcal{B}_q$
- Betti numbers:

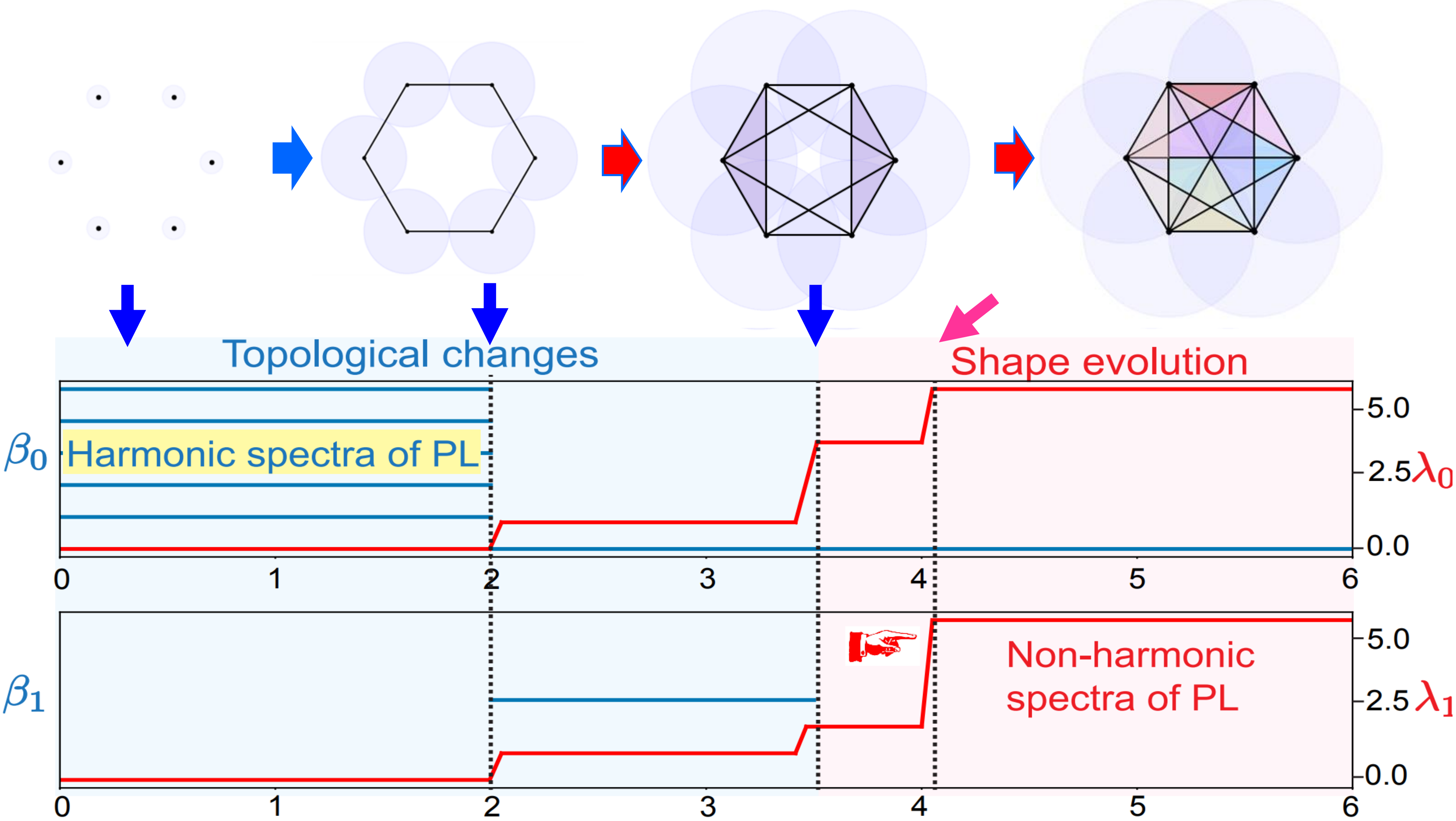
$\beta_q = \dim(\mathcal{L}_q(K)) - \operatorname{rank}\left(\mathcal{L}_q(K)\right)$=# of zero eigenvalues of $\mathcal{L}_q(K)$

# Persistent (Combinatorial) Laplacians

$$\mathcal{L}_q^{t+p} = \mathcal{B}_{q+1}^{t+p}\left(\mathcal{B}_{q+1}^{t+p}\right)^T + \left(\mathcal{B}_q^{t+p}\right)^T \mathcal{B}_q^{t+p}$$

Rui Wang

(Wang, Nguyen, Wei, 2019; Meng et al. 2021; Memoli et al. 2022; Liu and Wu 2023)



Alternative: Persistent Dirac by Maroulas and coworkers, Xia and coworkers

# More in our toolbox for TDA

## Evolutionary Homology

Zixuan Cang, Munch, Wei, J. Appl. Comput. Topology, 2020

## Persistent sheaf Laplacians

Xiaoqi Wei, Wei, under review, 2021

## Persistent Path Laplacians

Rui Wang,  Wei, Foundation of Data Science, 2023

## Persistent hypergraph Laplacians

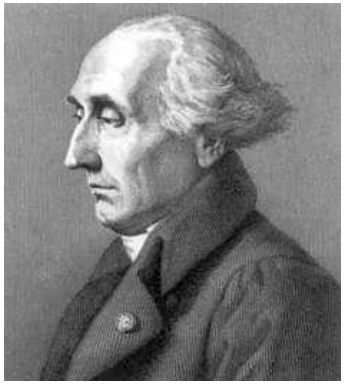
Dong Chen, Liu, Wu, Wei, 2023

## Persistent hyperdigraph Laplacians

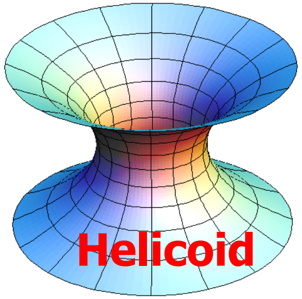
Dong Chen, Liu, Wu, Wei, 2023

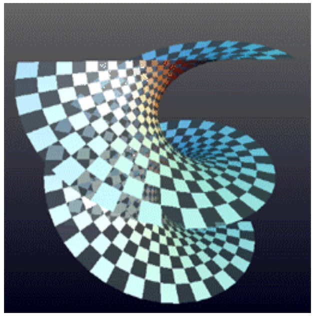
# Differential geometry

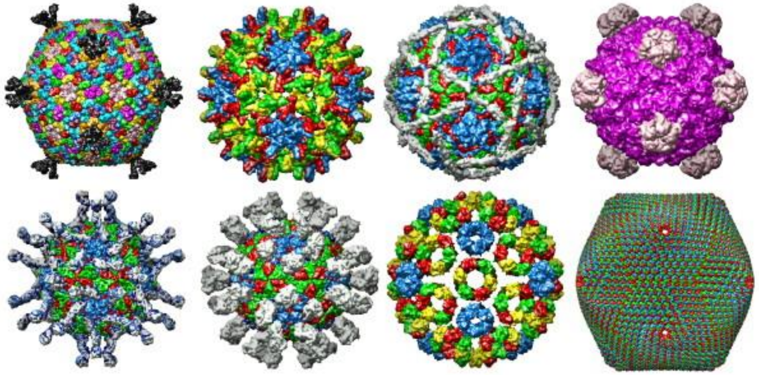
**Leonhard P. Euler**
(Swiss Mathematician,
April 15, 1707 – Sept
18 1783



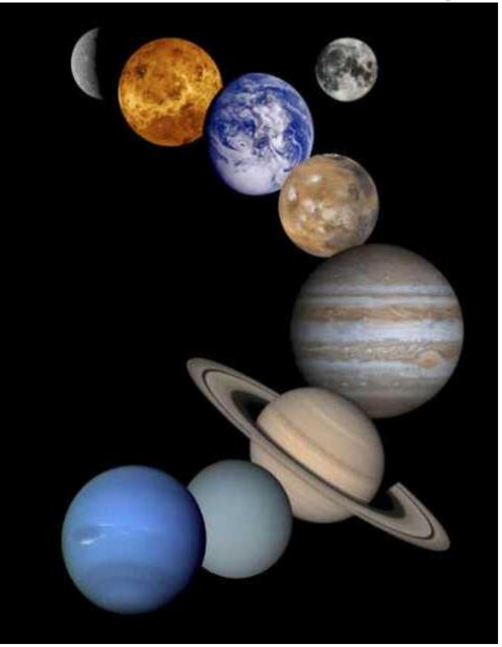
Joseph L. Lagrange
(Italian
Mathematician,
January 25 1736 –
April 10, 1813)





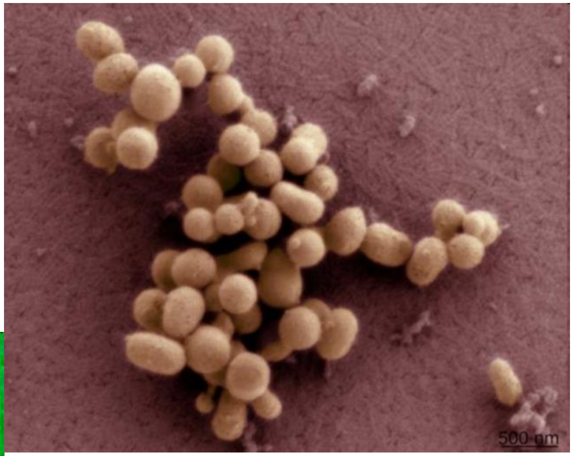


Viral morphology



**Minimal Surfaces**

**A way to minimize energy and maximize stability**

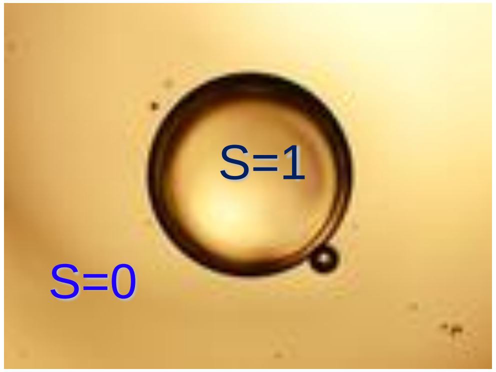


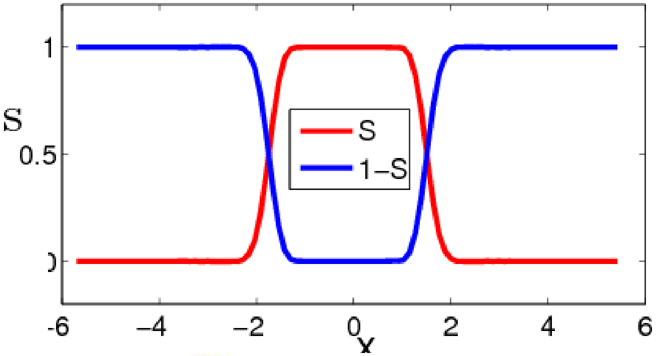
Man-made life,
Mycoplasma
mycoides

# Differential geometry based multiscale model

$$G = \int \gamma[\text{area}]\, d\boldsymbol{r} \qquad \text{area} = |\nabla S|$$

where G is the surface energy, gamma $(\gamma)$ is the surface tension, and S is a surface characteristic function:

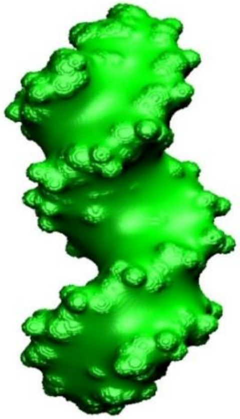


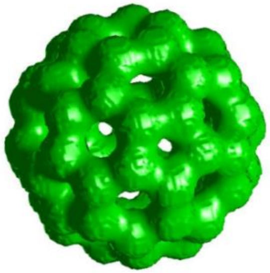


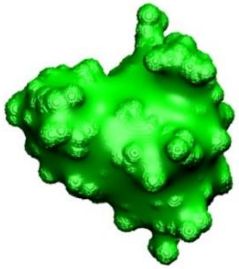
Generalized Laplace-Beltrami flow:

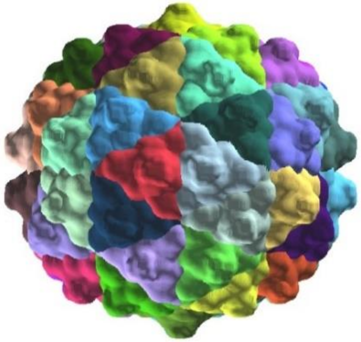
$$\frac{\partial S}{\partial t} = |\nabla S| \left[ \nabla \cdot \frac{\gamma \nabla S}{|\nabla S|} \right]$$

Mean curvature









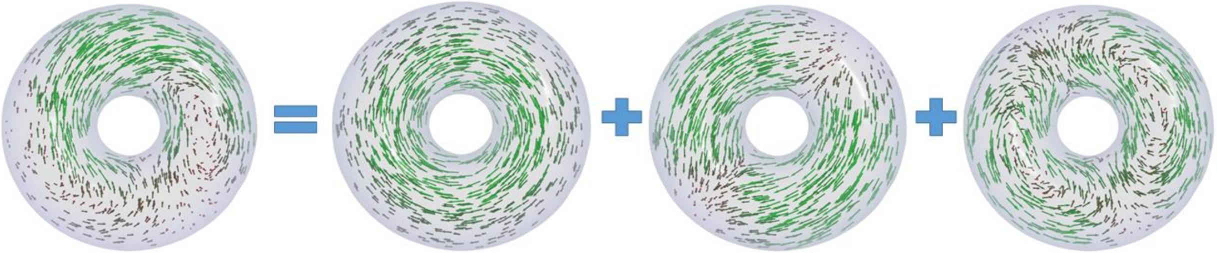
Shan Zhao

(Bates, Wei, Zhao, 2006; JCC,2008; Zhao, Cang, Tong & Wei, Bioinformatics 2018 )

# De Rham-Hodge theory and discrete exterior calculus

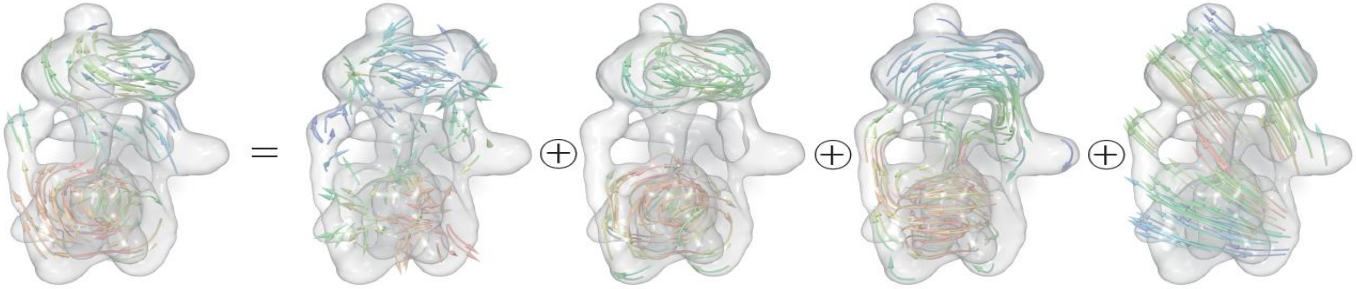
**Hodge decomposition:**

(Zhao, Wang, Chen, Tong & Wei, BMB, 2020)



**A vector field = Harmonic + curl-free + divergent-free**

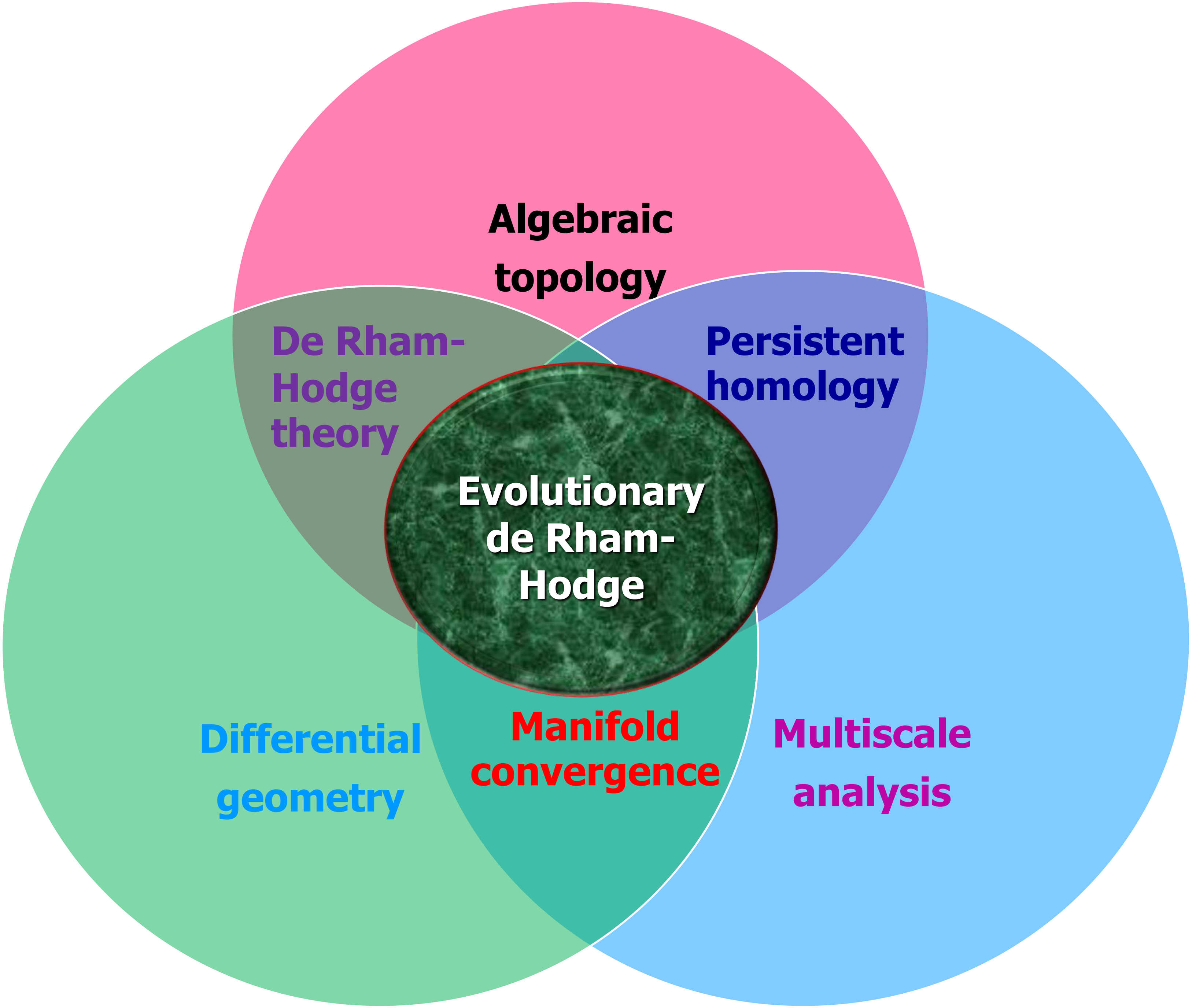
**Cryo-EM data:**



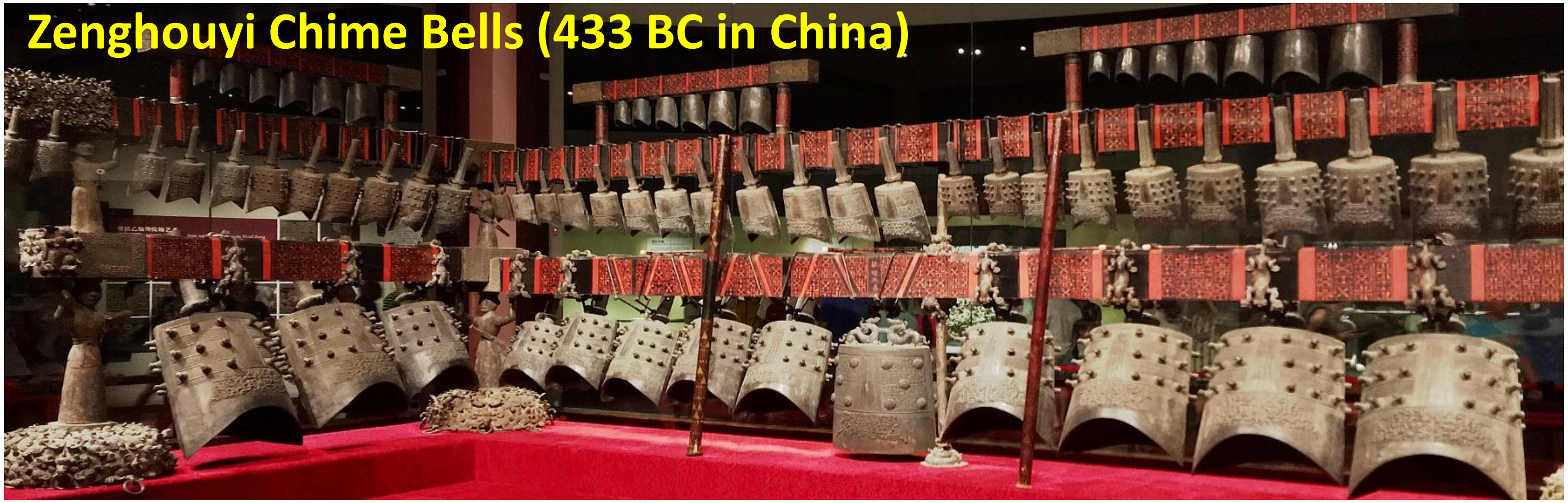
$$\begin{tikzcd}
\text{0-Form} \arrow[r, "d_0"] \arrow[d, shift right, "\star_0"'] \arrow[dr, phantom, "\circlearrowleft"] & \text{1-Form} \arrow[r, "d_1"] \arrow[d, shift right, "\star_1"'] \arrow[dr, phantom, "\circlearrowleft"] & \text{2-Form} \arrow[r, "d_2"] \arrow[d, shift right, "\star_2"'] \arrow[dr, phantom, "\circlearrowleft"] & \text{3-Form} \arrow[d, shift right, "\star_3"'] \\
\text{Dual 3-Form} \arrow[u, shift right, "\star_0^{-1}"'] & \text{Dual 2-Form} \arrow[l, "d_0^T"'] \arrow[u, shift right, "\star_1^{-1}"'] & \text{Dual 1-Form} \arrow[l, "d_1^T"'] \arrow[u, shift right, "\star_2^{-1}"'] & \text{Dual 0-Form} \arrow[l, "d_2^T"'] \arrow[u, shift right, "\star_3^{-1}"']
\end{tikzcd}$$

$$\begin{tikzcd}
\textit{point-based} \text{ scalar field} \arrow[r, "\nabla"] \arrow[d, leftrightarrow, "\star"] & \textit{edge-based} \text{ vector field} \arrow[r, "\nabla\times"] \arrow[d, leftrightarrow, "\star"] & \textit{face-based} \text{ vector field} \arrow[r, "\nabla\cdot"] \arrow[d, leftrightarrow, "\star"] & \textit{cell-based} \text{ scalar field} \arrow[d, leftrightarrow, "\star"] \\
\textit{cell-based} \text{ scalar field} & \textit{face-based} \text{ vector field} \arrow[l, "\nabla\cdot"'] & \textit{edge-based} \text{ vector field} \arrow[l, "\nabla\times"'] & \textit{point-based} \text{ scalar field} \arrow[l, "\nabla"']
\end{tikzcd}$$

(Douglas Arnold, M Desbrun, AN Hirani, …)


Algebraic topology
De Rham-Hodge theory
Persistent homology
Evolutionary de Rham-Hodge
Differential geometry
Manifold convergence
Multiscale analysis

# Evolutionary de Rham-Hodge



Zenghouyi Chime Bells (433 BC in China)

## Manifold filtration

$$\begin{tikzcd}
M_0 \arrow[r, "\mathfrak{I}_{0,1}"] & M_1 \arrow[r, "\mathfrak{I}_{1,2}"] & M_2 \arrow[r, "\mathfrak{I}_{2,3}"] & \cdots \arrow[r, "\mathfrak{I}_{n-1,n}"] & M_n \arrow[r, "\mathfrak{I}_{n,n+1}"] & M
\end{tikzcd}$$

## Filtration-induced de Rham complexes:

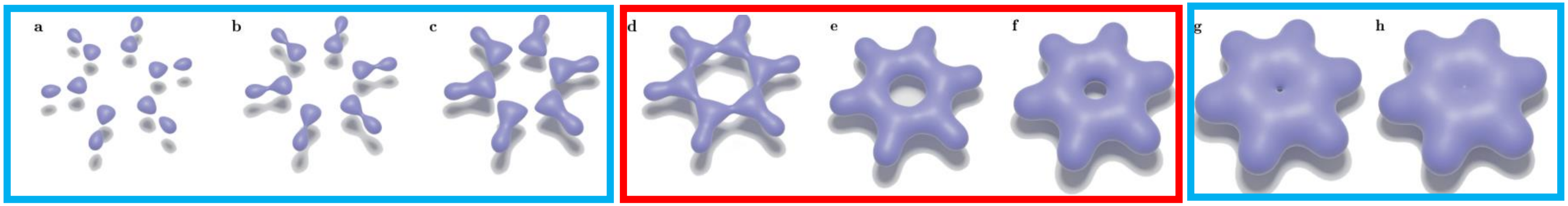
$$\begin{tikzcd}
\Omega^0_n(M_0) \arrow[r, "d^0"] \arrow[d, "\mathfrak{E}_{0,1}"] & \Omega^1_n(M_0) \arrow[r, "d^1"] \arrow[d, "\mathfrak{E}_{0,1}"] & \Omega^2_n(M_0) \arrow[r, "d^2"] \arrow[d, "\mathfrak{E}_{0,1}"] & \Omega^3_n(M_0) \arrow[d, "\mathfrak{E}_{0,1}"] \\
\Omega^0_n(M_1) \arrow[r, "d^0"] \arrow[d, "\mathfrak{E}_{1,1}"] & \Omega^1_n(M_1) \arrow[r, "d^1"] \arrow[d, "\mathfrak{E}_{1,1}"] & \Omega^2_n(M_1) \arrow[r, "d^2"] \arrow[d, "\mathfrak{E}_{1,1}"] & \Omega^3_n(M_1) \arrow[d, "\mathfrak{E}_{1,1}"] \\
\Omega^0_n(M_2) \arrow[r, "d^0"] \arrow[d, "\mathfrak{E}_{2,1}"] & \Omega^1_n(M_2) \arrow[r, "d^1"] \arrow[d, "\mathfrak{E}_{2,1}"] & \Omega^2_n(M_2) \arrow[r, "d^2"] \arrow[d, "\mathfrak{E}_{2,1}"] & \Omega^3_n(M_2) \arrow[d, "\mathfrak{E}_{2,1}"] \\
\cdots & \cdots & \cdots & \cdots
\end{tikzcd}$$

(Chen, Zhao, Tong & Wei, DCDS-B, 2020)

# Evolutionary de Rham-Hodge Laplacians

$$\Delta_k^{l,p} = \partial_{k+1}^{l} d_k^{l} + d_{k-1}^{l+p} \partial_k^{l+p}$$

**Manifold filtration:**

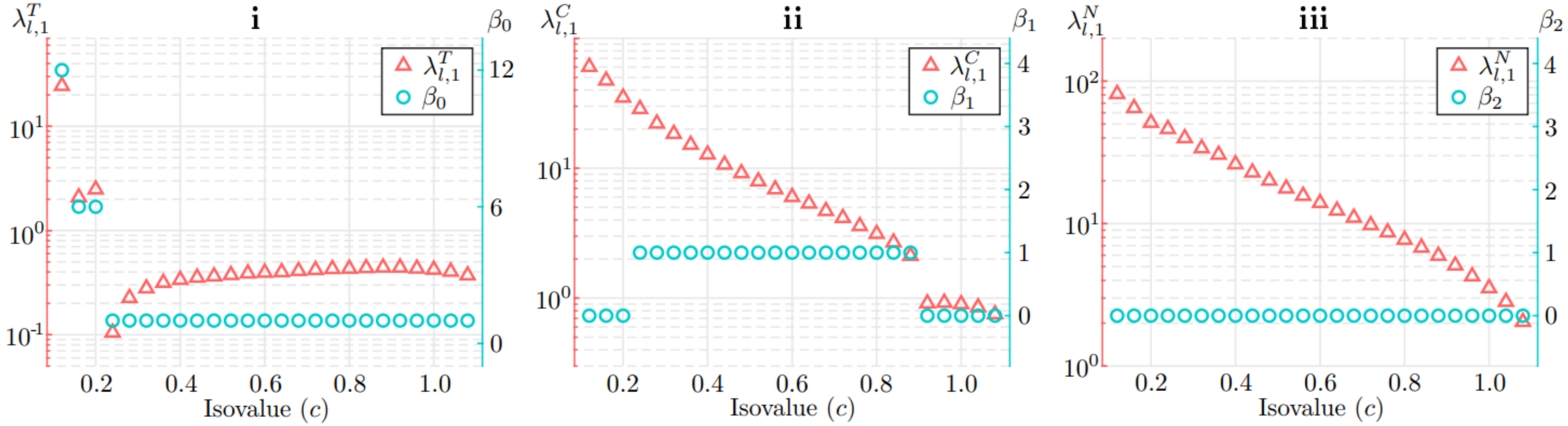


**Topological persistence**

**Homotopic shape evolution**

**Topological persistence**

## Discontinuous harmonic (topological) and continuous non-harmonic spectra



(Chen, Zhao, Tong & Wei, DCDS-B, 2020)

# Mathematical learning algorithms

Logistic regression

Support vector machine

Random forest

Ensemble methods

Transfer learning

Active learning

Deep neural network

Convolutional neural network

Nature language processing

Recurrent neural network

Long-short term memory

Graph neural network

Generative AI

ChatGPT

Autoencoder

Transformer

Manifold learning

Graph learning

Geometric learning

PCA

UMAP

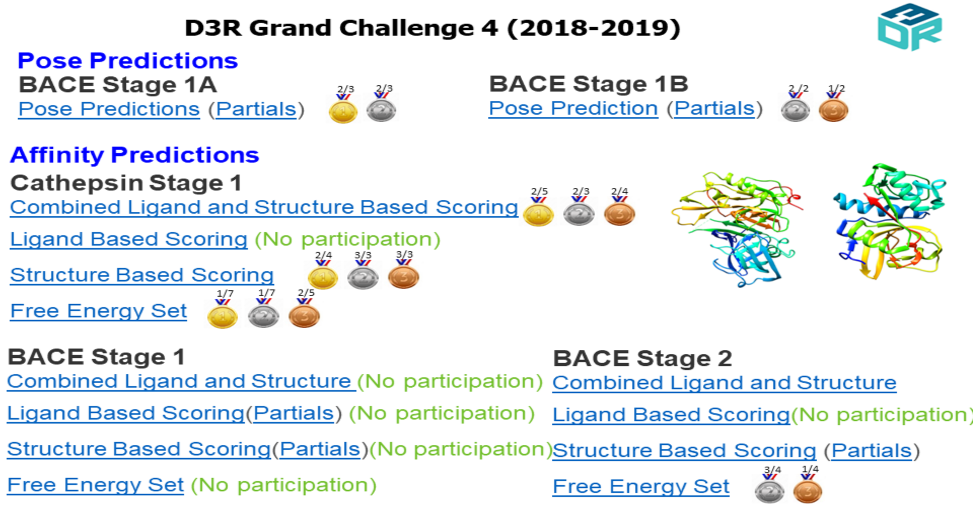
t-SNE

Correlated clustering and projection

Topological deep learning

Multiscale Laplacian learning

# Our performance in D3R Grand Challenges, worldwide competitions in computer-aided drug design organized by NIH, 2016-2019.



## D3R Grand Challenge 4 (2018-2019)

**Pose Predictions**

**BACE Stage 1A**
Pose Predictions (Partials) 2/3 🥇 2/3 🥈

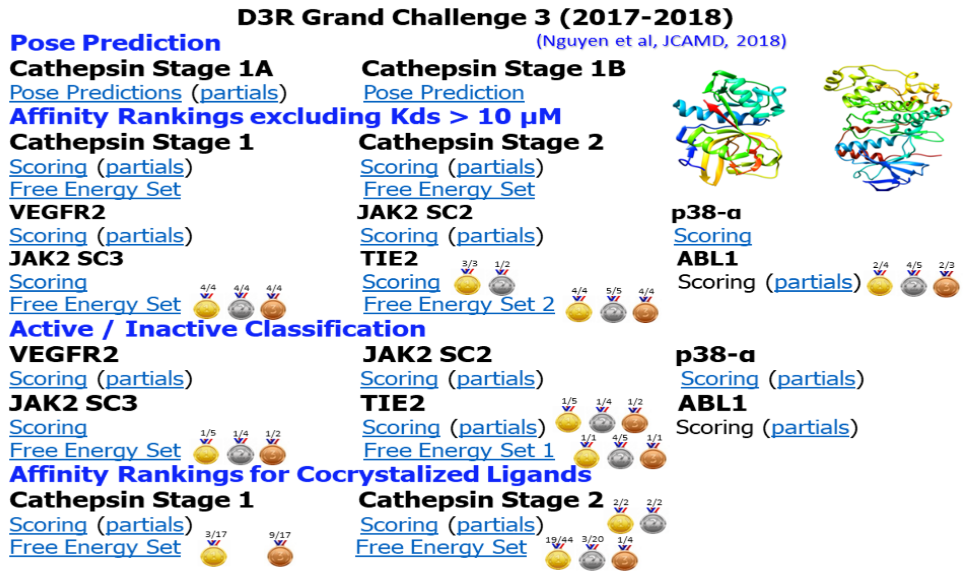
**BACE Stage 1B**
Pose Prediction (Partials) 2/2 🥈 1/2 🥉

**Affinity Predictions**

**Cathepsin Stage 1**
Combined Ligand and Structure Based Scoring 2/5 🥇 2/3 🥈 2/4 🥉
Ligand Based Scoring (No participation)
Structure Based Scoring 2/4 🥇 3/3 🥈 3/3 🥉
Free Energy Set 1/7 🥇 1/7 🥈 2/5 🥉

**BACE Stage 1**
Combined Ligand and Structure (No participation)
Ligand Based Scoring(Partials) (No participation)
Structure Based Scoring(Partials)(No participation)
Free Energy Set (No participation)

**BACE Stage 2**
Combined Ligand and Structure
Ligand Based Scoring(No participation)
Structure Based Scoring (Partials)
Free Energy Set 3/4 🥈 1/4 🥉



## D3R Grand Challenge 3 (2017-2018)

(Nguyen et al, JCAMD, 2018)

**Pose Prediction**

**Cathepsin Stage 1A**
Pose Predictions (partials)

**Cathepsin Stage 1B**
Pose Prediction

**Affinity Rankings excluding Kds > 10 μM**

**Cathepsin Stage 1**
Scoring (partials)
Free Energy Set

**Cathepsin Stage 2**
Scoring (partials)
Free Energy Set

**VEGFR2**
Scoring (partials)

**JAK2 SC3**
Scoring
Free Energy Set 4/4 🥇 4/4 🥈 4/4 🥉

**JAK2 SC2**
Scoring (partials)

**TIE2**
Scoring 3/3 🥇 1/2 🥈
Free Energy Set 2 4/4 🥇 5/5 🥈 4/4 🥉

**p38-α**
Scoring

**ABL1**
Scoring (partials) 2/4 🥇 4/5 🥈 2/3 🥉

**Active / Inactive Classification**

**VEGFR2**
Scoring (partials)

**JAK2 SC3**
Scoring
Free Energy Set 1/5 🥇 1/4 🥈 1/2 🥉

**JAK2 SC2**
Scoring (partials)

**TIE2**
Scoring (partials) 1/5 🥇 1/4 🥈 1/2 🥉
Free Energy Set 1 1/1 🥇 4/5 🥈 1/1 🥉

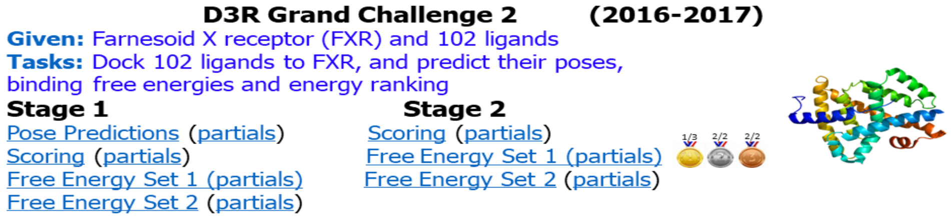
**p38-α**
Scoring (partials)

**ABL1**
Scoring (partials)

**Affinity Rankings for Cocrystalized Ligands**

**Cathepsin Stage 1**
Scoring (partials)
Free Energy Set 3/17 🥇 9/17 🥉

**Cathepsin Stage 2** 2/2 🥇 2/2 🥈
Scoring (partials)
Free Energy Set 19/44 🥇 3/20 🥈 1/4 🥉



## D3R Grand Challenge 2 (2016-2017)

**Given:** Farnesoid X receptor (FXR) and 102 ligands
**Tasks:** Dock 102 ligands to FXR, and predict their poses, binding free energies and energy ranking

**Stage 1**
Pose Predictions (partials)
Scoring (partials)
Free Energy Set 1 (partials)
Free Energy Set 2 (partials)

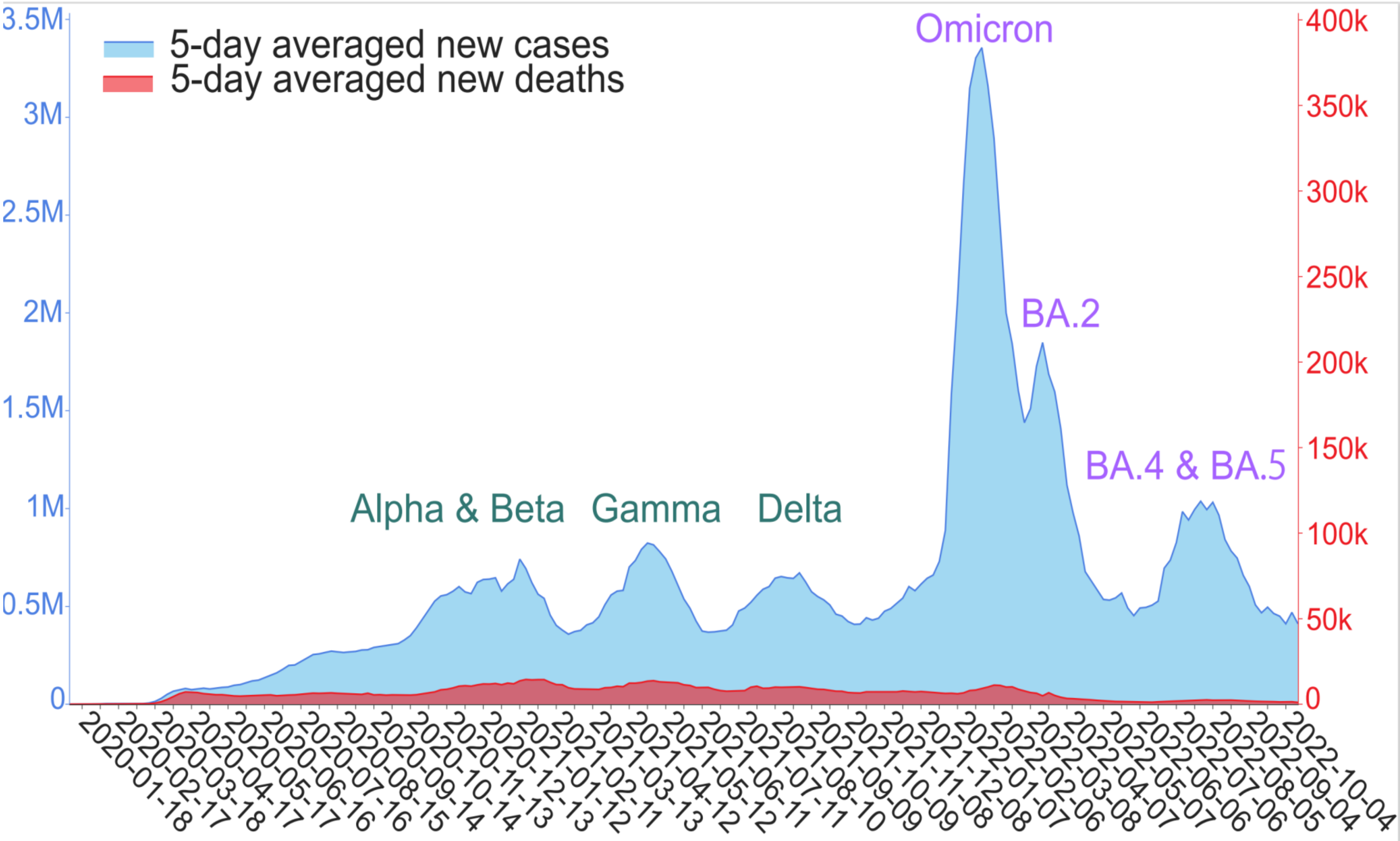
**Stage 2**
Scoring (partials)
Free Energy Set 1 (partials) 1/3 🥇 2/2 🥈 2/2 🥉
Free Energy Set 2 (partials)

D Nguyen

Zixuan Cang

Kaifu Gao

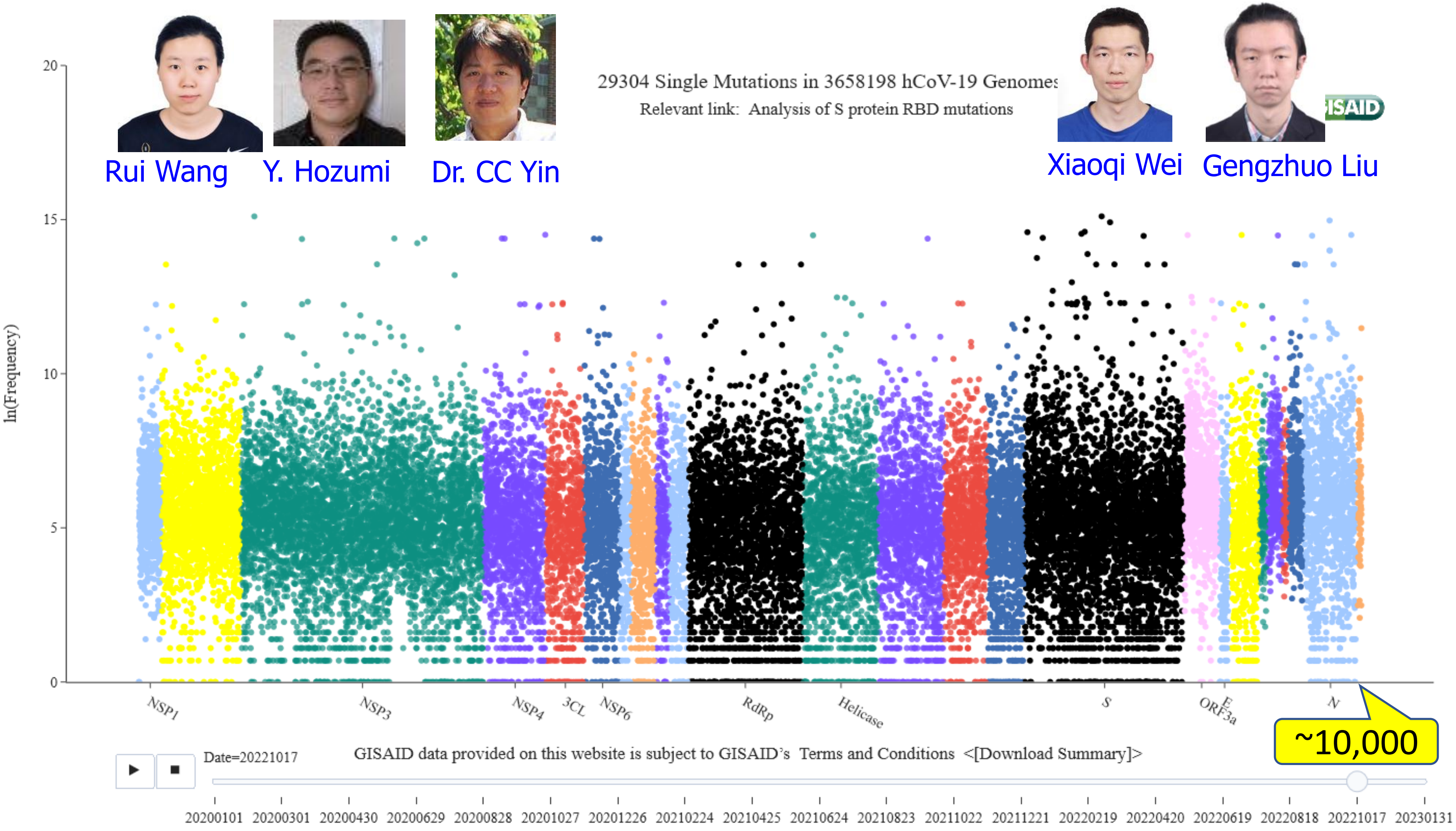
# Evolution of SARS-CoV-2 variants



What are the evolutionary mechanisms?

# Mutation Tracker

https://users.math.msu.edu/users/weig/SARS-CoV-2_Mutation_Tracker.html



Rui Wang Y. Hozumi Dr. CC Yin Xiaoqi Wei Gengzhuo Liu

29304 Single Mutations in 3658198 hCoV-19 Genomes

Relevant link: Analysis of S protein RBD mutations

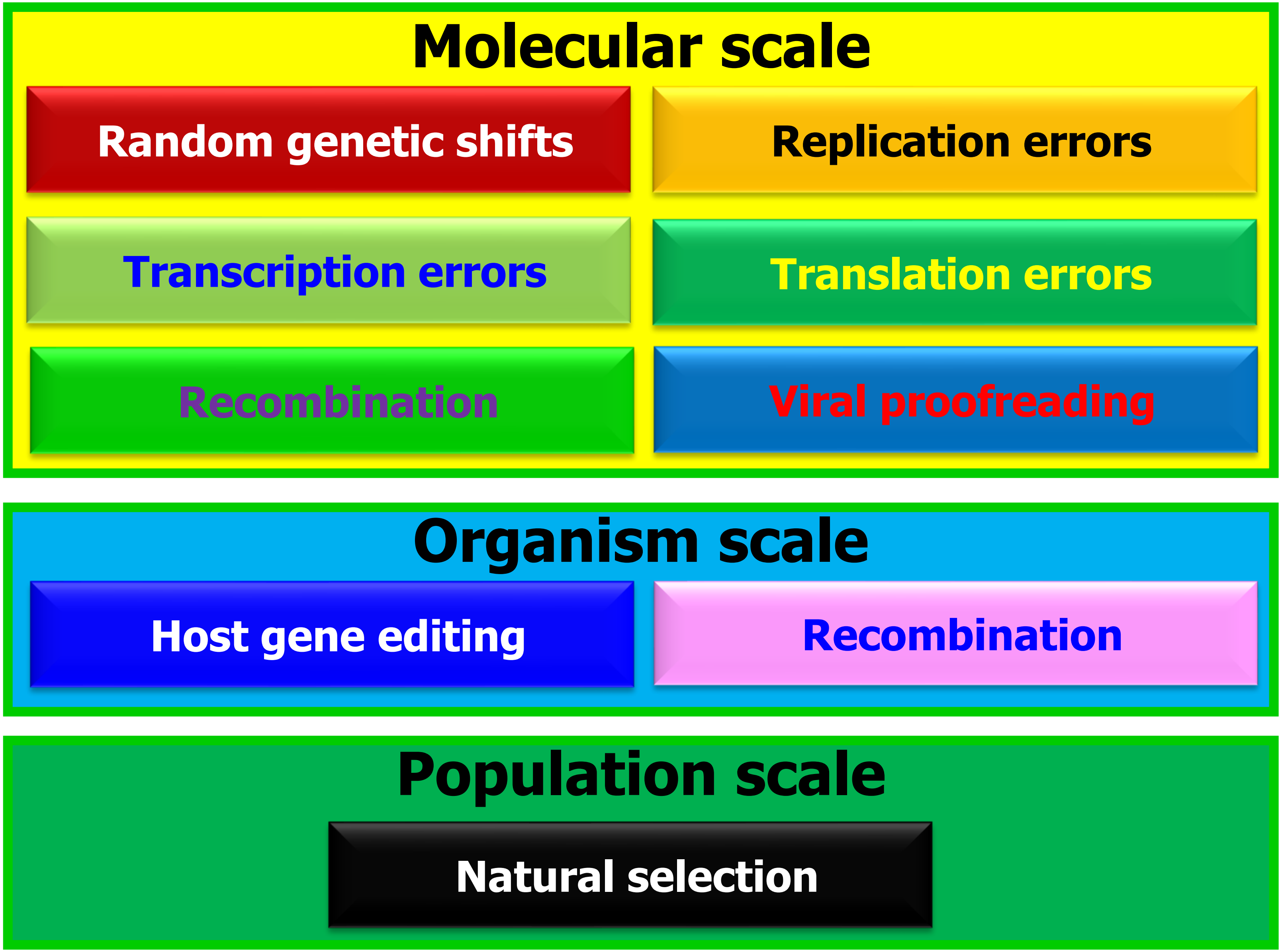
Date=20221017

GISAID data provided on this website is subject to GISAID's Terms and Conditions <[Download Summary]>

~10,000

**What governs SARS-CoV-2 transmission and evolution?**

# Competing mechanisms of SARS-CoV mutations



## Molecular scale

Random genetic shifts

Replication errors

Transcription errors

Translation errors

Recombination

Viral proofreading

## Organism scale

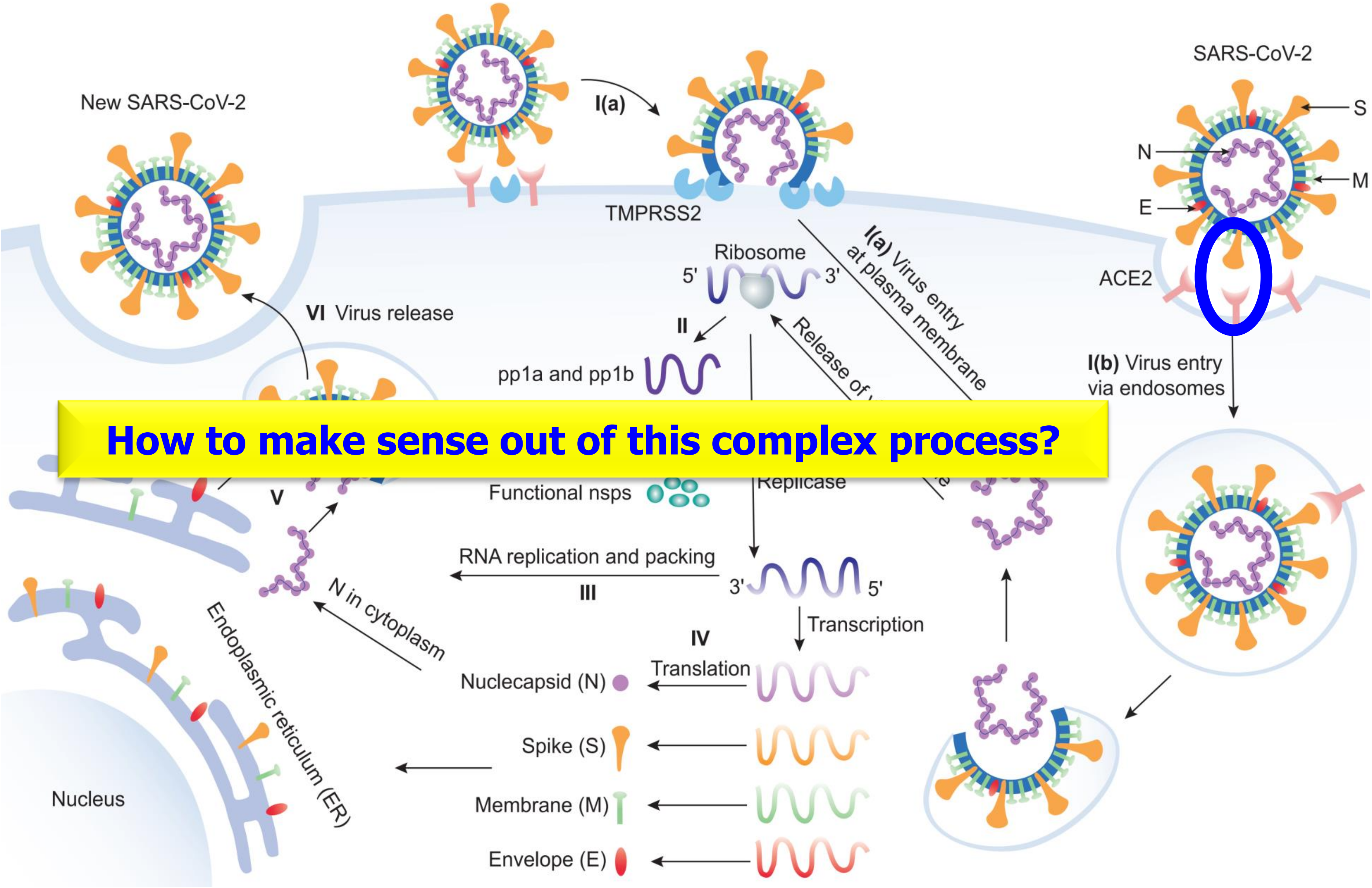
Host gene editing

Recombination

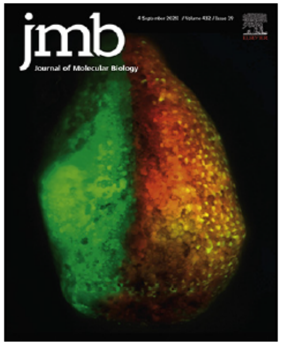
## Population scale

Natural selection

# Life cycle of SARS-CoV-2 in a host cell



**How to make sense out of this complex process?**

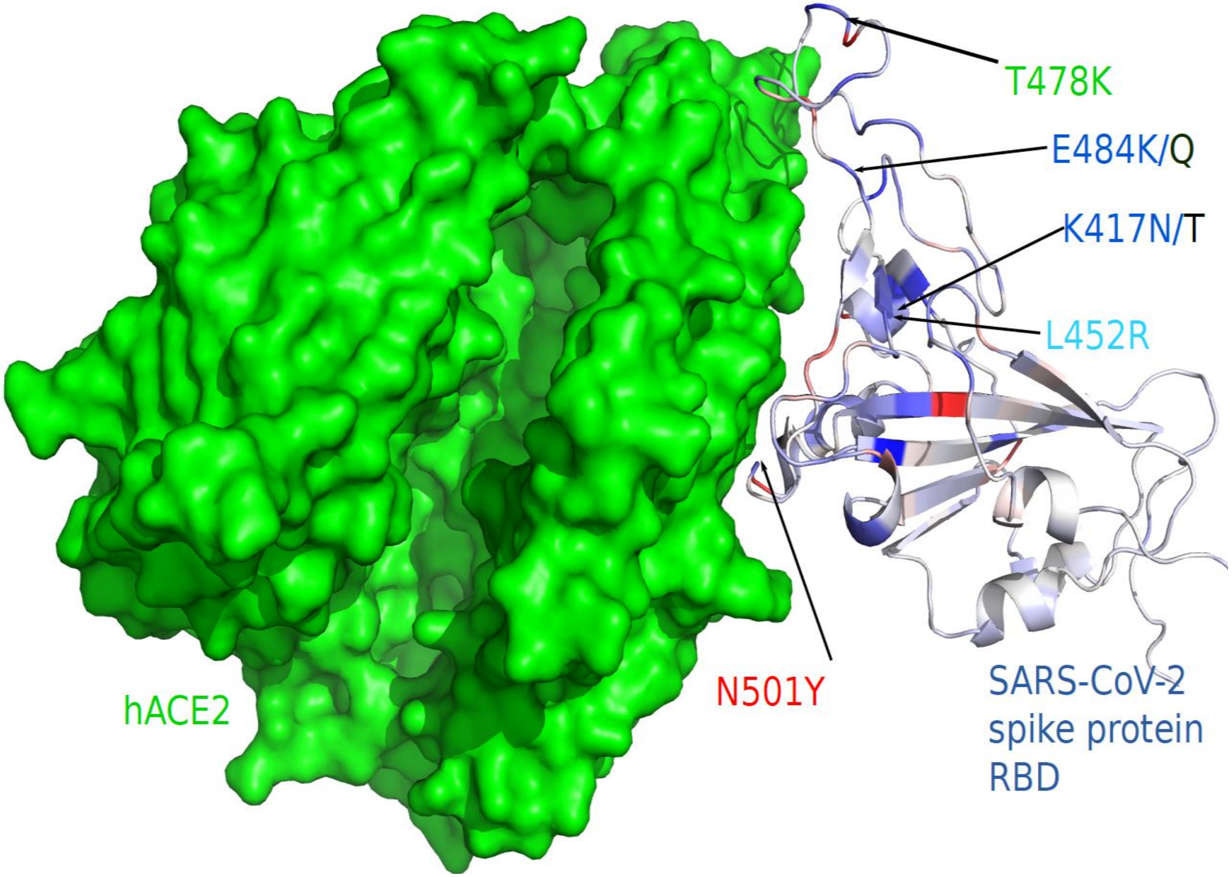


# Mutations Strengthened SARS-CoV-2 Infectivity

We predicted prevailing SARS-CoV-2 variants to occur at residues 452 and 501

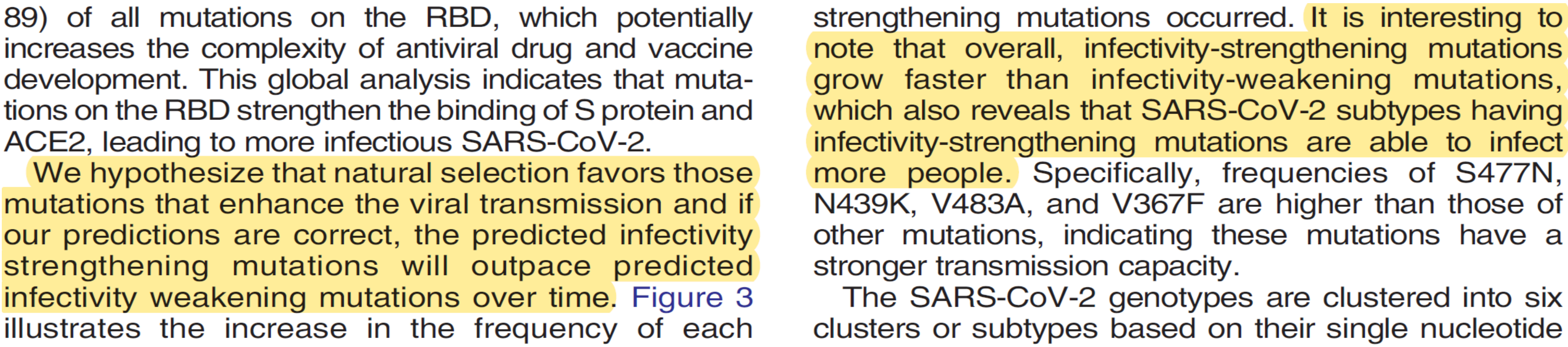
Dr Jiahui Chen

Jiahui Chen[1], Rui Wang[1], Menglun Wang[1] and Guo-Wei Wei[1,2,3]



**Alpha**: N501Y
**Beta**: K417N, E484K, N501Y
**Gamma**: K417T, E484K, N501Y
**Delta**: L452R, T478K
Epsilon: L452R
Theta: E484K, N501Y
Kappa: L452R, E484Q
Lambda: L452Q, F490S
Mu: R346K,E484K,N501Y
**Omicron**: G339D, S371L, S373P, S375F, K417N, N440K, G446S, S477N, T478K, E484A, Q493R, G496S, Q498R, N501Y, Y505H;
**BA.2**.12.1: Omicron + L452Q;
**BA.4/BA.5**: Omicron + L452R

# We discovered the mechanism of viral transmission and evolution



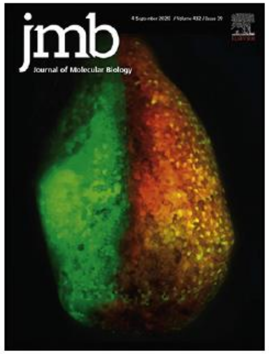
89) of all mutations on the RBD, which potentially increases the complexity of antiviral drug and vaccine development. This global analysis indicates that mutations on the RBD strengthen the binding of S protein and ACE2, leading to more infectious SARS-CoV-2.

We hypothesize that natural selection favors those mutations that enhance the viral transmission and if our predictions are correct, the predicted infectivity strengthening mutations will outpace predicted infectivity weakening mutations over time. Figure 3 illustrates the increase in the frequency of each strengthening mutations occurred. It is interesting to note that overall, infectivity-strengthening mutations grow faster than infectivity-weakening mutations, which also reveals that SARS-CoV-2 subtypes having infectivity-strengthening mutations are able to infect more people. Specifically, frequencies of S477N, N439K, V483A, and V367F are higher than those of other mutations, indicating these mutations have a stronger transmission capacity.

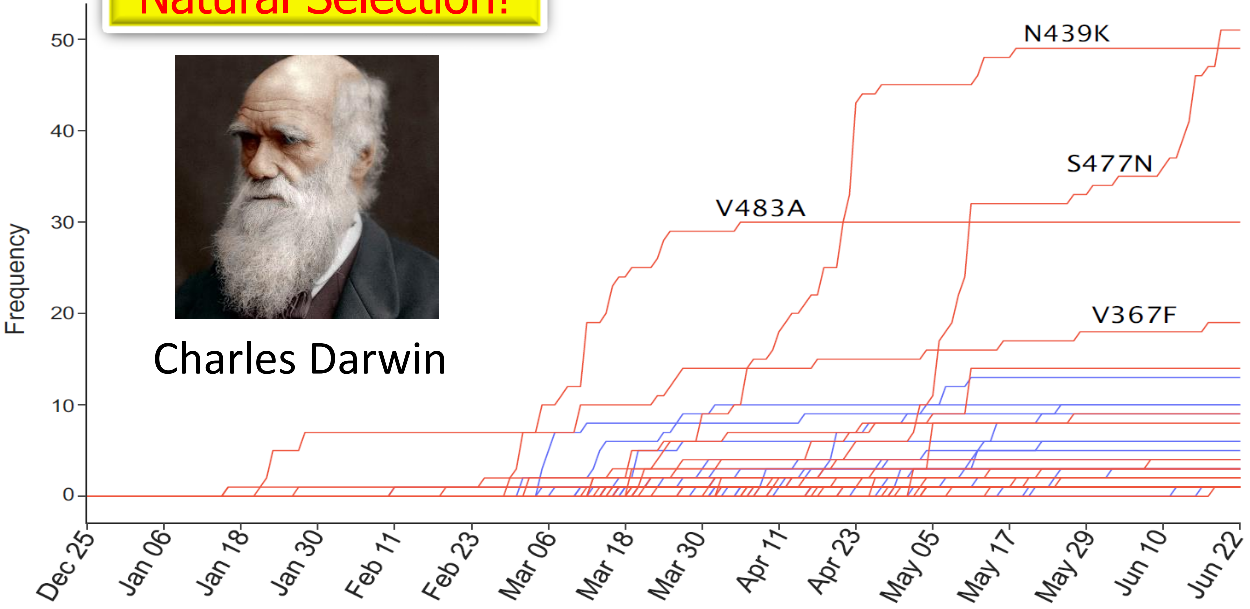
The SARS-CoV-2 genotypes are clustered into six clusters or subtypes based on their single nucleotide

Based on 15,140 viral genomes

Natural Selection!



Chen, Wang, Wang, Wei, JMB, 432, 5212, July 2020



Charles Darwin

Dr Jiahui Chen

Rui Wang

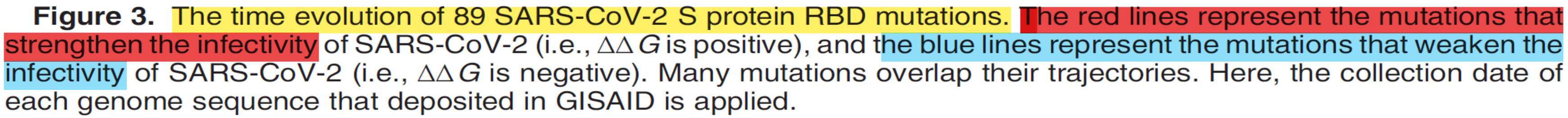


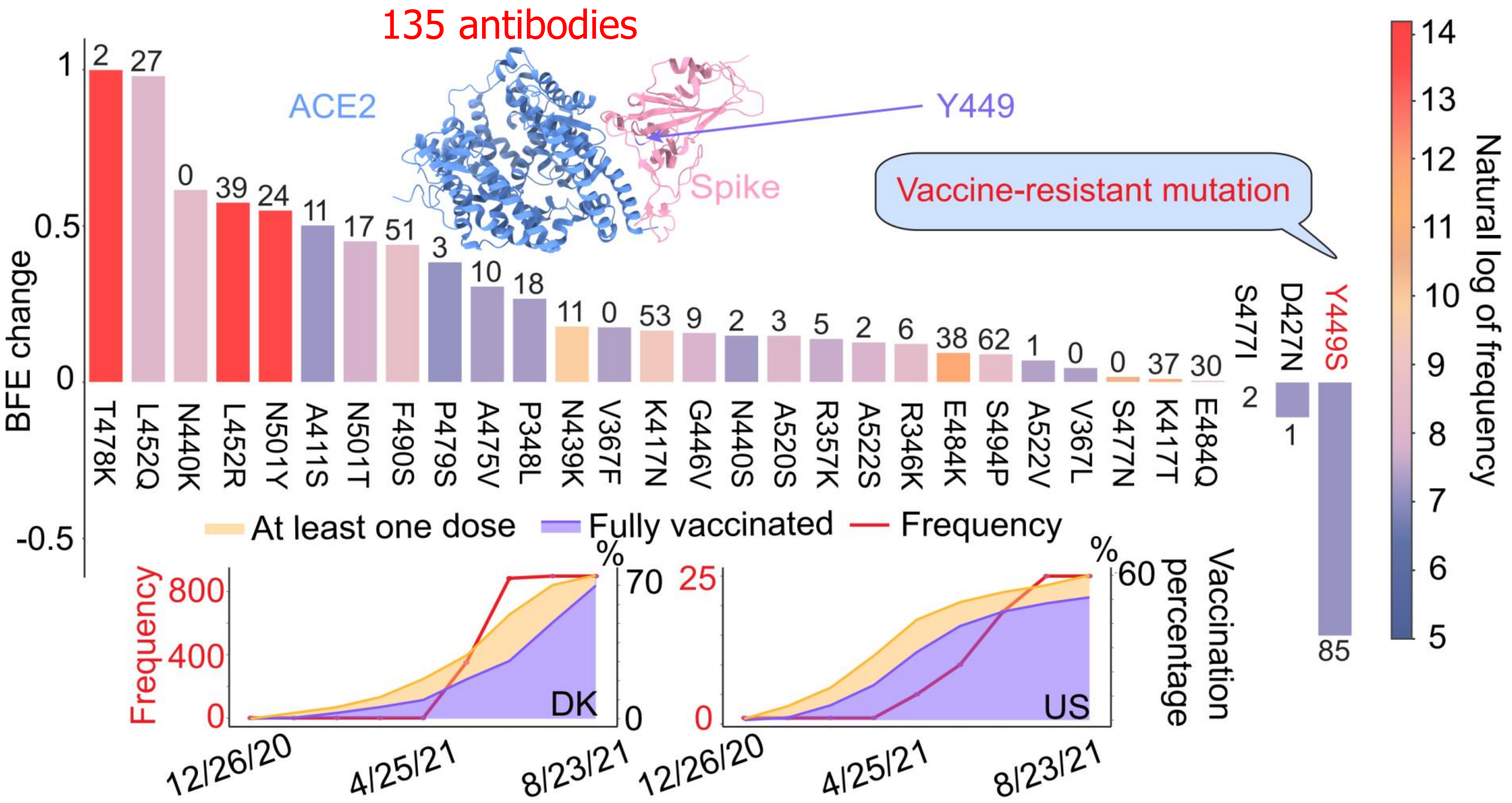
**Figure 3.** The time evolution of 89 SARS-CoV-2 S protein RBD mutations. The red lines represent the mutations that strengthen the infectivity of SARS-CoV-2 (i.e., $\Delta\Delta G$ is positive), and the blue lines represent the mutations that weaken the infectivity of SARS-CoV-2 (i.e., $\Delta\Delta G$ is negative). Many mutations overlap their trajectories. Here, the collection date of each genome sequence that deposited in GISAID is applied.

# Vaccine-breakthrough mutations

## By genotyping 2,298,349 viral genomes isolated from patients

Wang, Chen, and Wei, J. Phys. Chem. Letter, 12. 11850-11857 2021

Rui Wang



Evolution mechanisms --- Natural selection via two complementary transmission pathways: Infectivity strengthening and vaccine breakthrough

# Omicron BA.2 (B.1.1.529.2): high potential to becoming the next dominating variant

Jiahui Chen[1] and Guo-Wei Wei[1,3,4*]
[1] Department of Mathematics,
Michigan State University, MI 48824, USA.

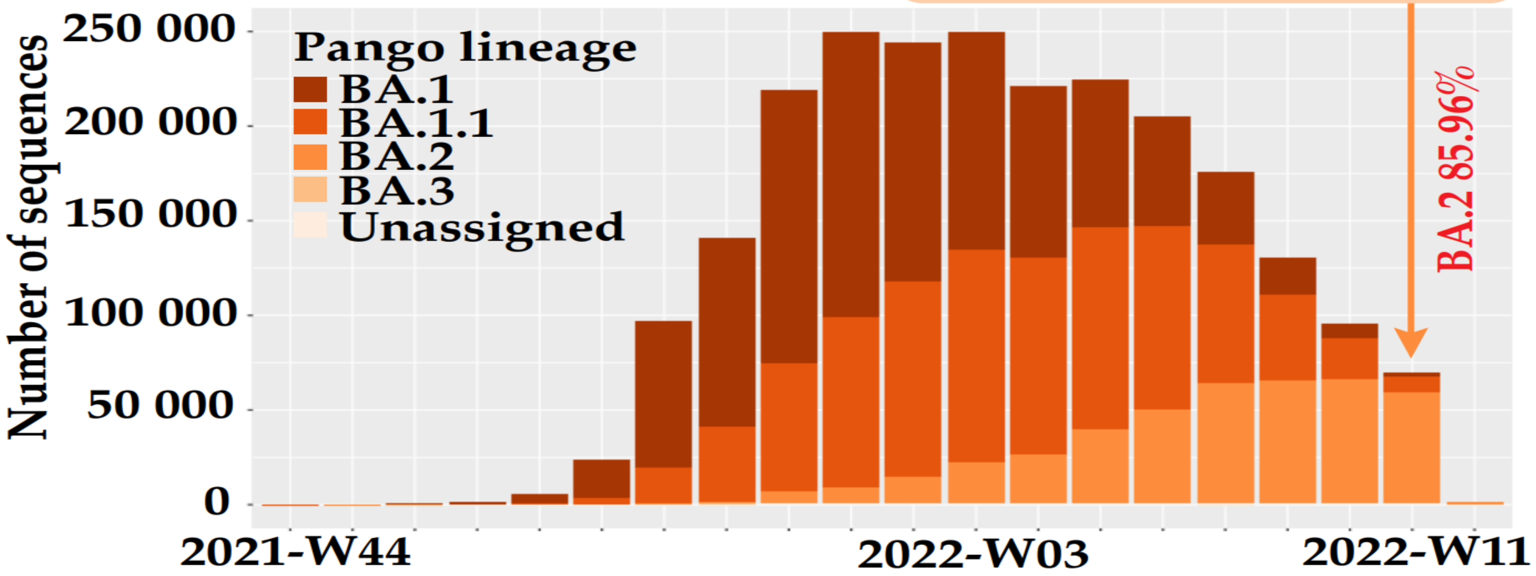
arXiv:2202.05031v1 [q-bio.PE] 10 Feb 2022

Dr Jiahui Chen

World Health Organization

COVID-19 Weekly Epidemiological Update
Edition 84, published 22 March 2022

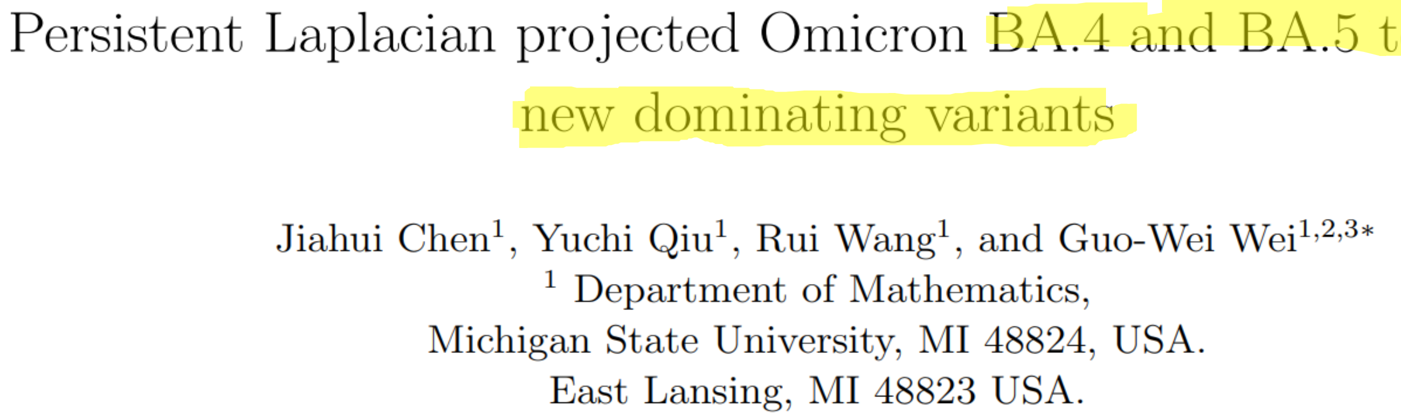
On 2/10/2022, we predicted that BA.2 will become the dominant variant. This became the reality in later March according to WHO



**This was confirmed by WHO on March 22, 2022!**

**All other predictions were confirmed within 50 days**

Chen, Wei, J. Phys. Chem. Lett., 13, 2840-3849, 2022.

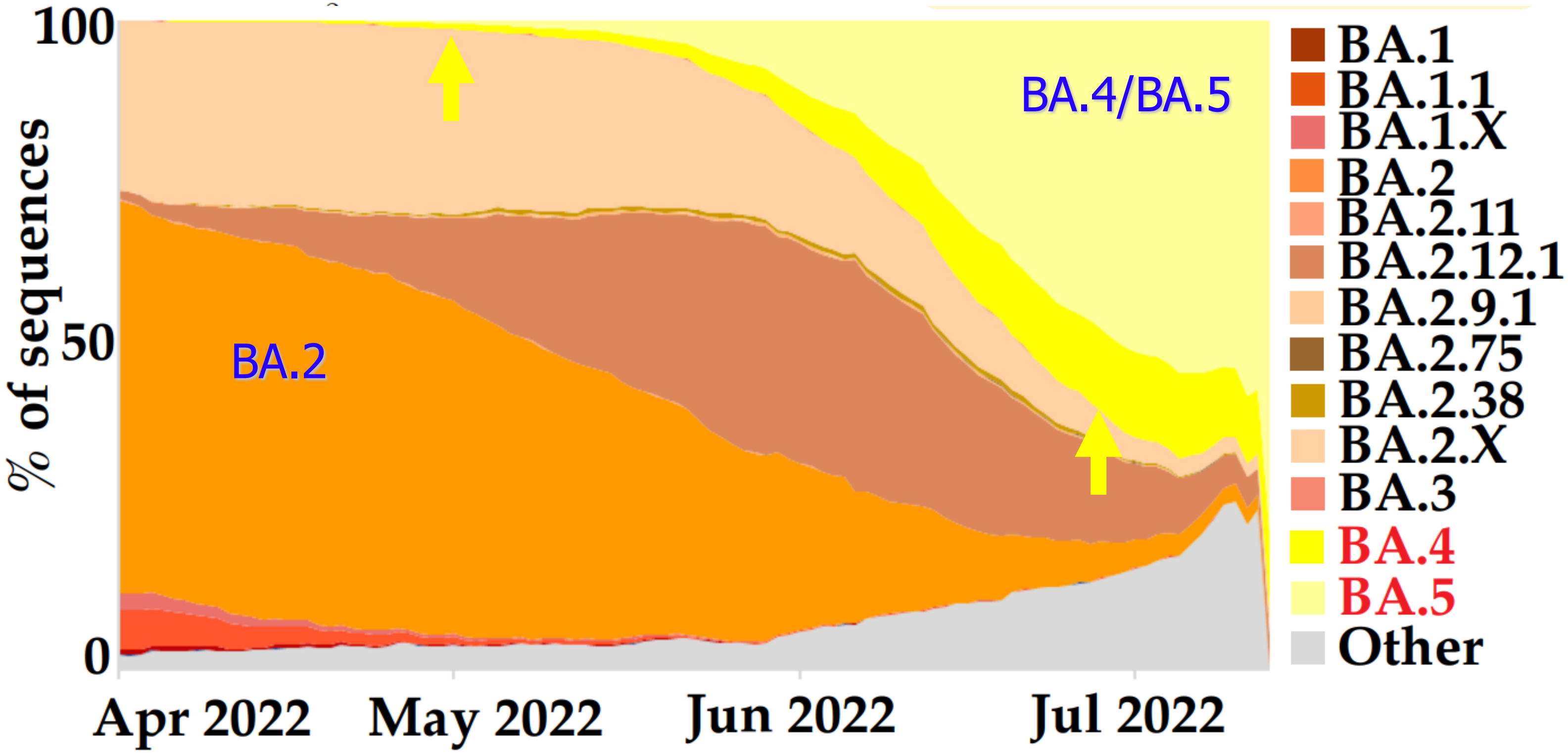


# Persistent Laplacian projected Omicron BA.4 and BA.5 to become new dominating variants

Jiahui Chen[1], Yuchi Qiu[1], Rui Wang[1], and Guo-Wei Wei[1,2,3]*
[1] Department of Mathematics,
Michigan State University, MI 48824, USA.
East Lansing, MI 48823 USA.

:submit/4287023 [q-bio.BM] 1 May 2022

Dr Jiahui Chen

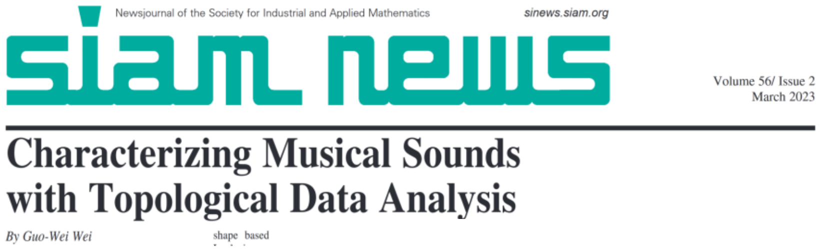


World Health Organization

**This was confirmed by WHO in early July (WHO weekly update release number 101)**

Jiahui Chen, Qiu, Wang, and Wei, Computers in Biology and Medicine, 151, 106262 (2023)

**Over one hundred of news and media coverages**



Newsjournal of the Society for Industrial and Applied Mathematics

sinews.siam.org

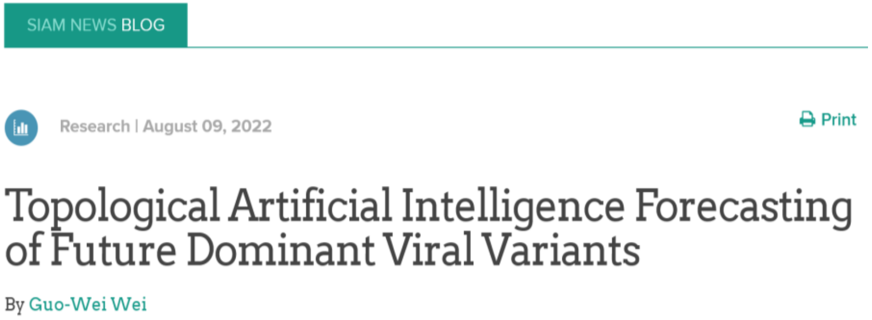
siam news

Volume 56/ Issue 2
March 2023

## Characterizing Musical Sounds with Topological Data Analysis

By Guo-Wei Wei

shape based



SIAM NEWS BLOG

Research | August 09, 2022

Print

## Topological Artificial Intelligence Forecasting of Future Dominant Viral Variants

By Guo-Wei Wei

SIAM NEWS MAY 2020

Research | May 01, 2020

Print

## Math and AI-based Repositioning of Existing Drugs for COVID-19

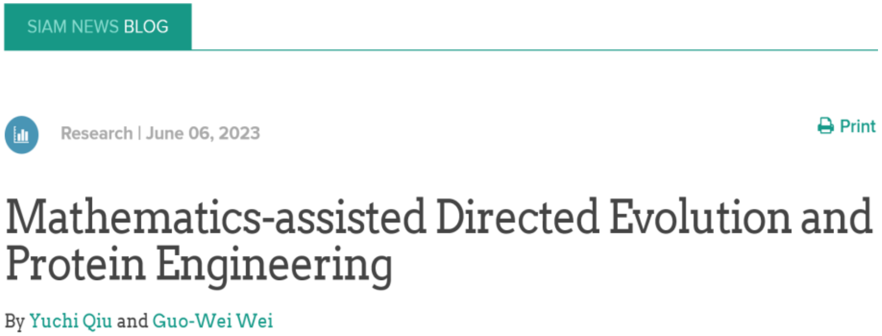
By Duc D. Nguyen and Guo-Wei Wei

SIAM NEWS DECEMBER 2017

Research | December 01, 2017

## Persistent Homology Analysis of Biomolecular Data

By Guo-Wei Wei



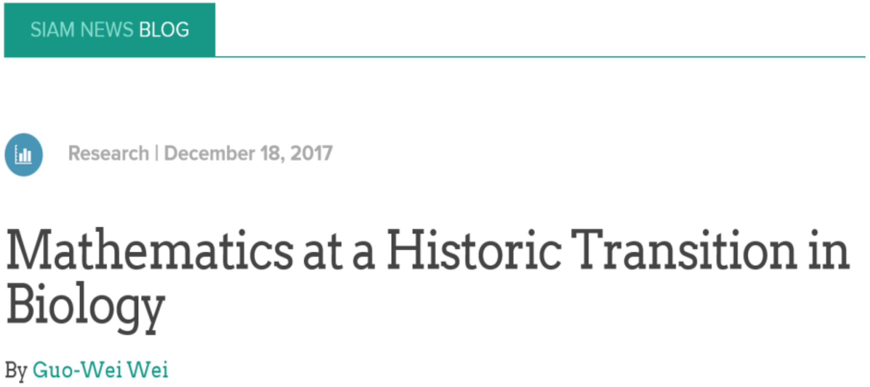
SIAM NEWS BLOG

Research | June 06, 2023

Print

## Mathematics-assisted Directed Evolution and Protein Engineering

By Yuchi Qiu and Guo-Wei Wei

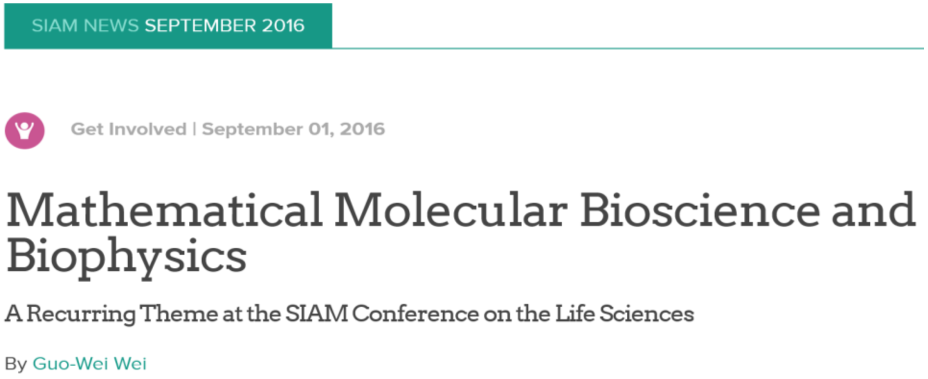


SIAM NEWS BLOG

Research | December 18, 2017

## Mathematics at a Historic Transition in Biology

By Guo-Wei Wei



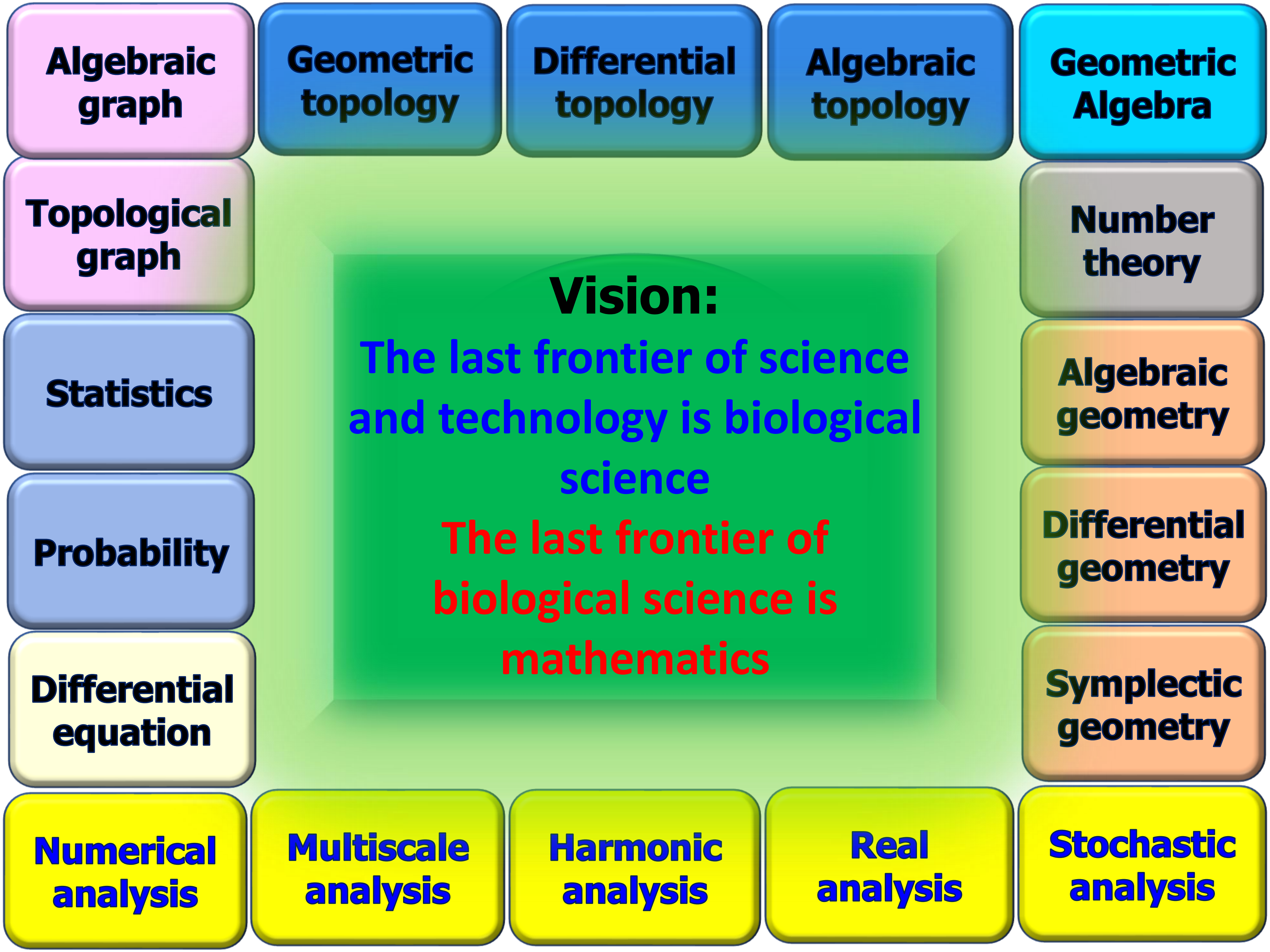
SIAM NEWS SEPTEMBER 2016

Get Involved | September 01, 2016

## Mathematical Molecular Bioscience and Biophysics

A Recurring Theme at the SIAM Conference on the Life Sciences

By Guo-Wei Wei



Algebraic graph

Geometric topology

Differential topology

Algebraic topology

Geometric Algebra

Topological graph

Statistics

Probability

Differential equation

Number theory

Algebraic geometry

Differential geometry

Symplectic geometry

Numerical analysis

Multiscale analysis

Harmonic analysis

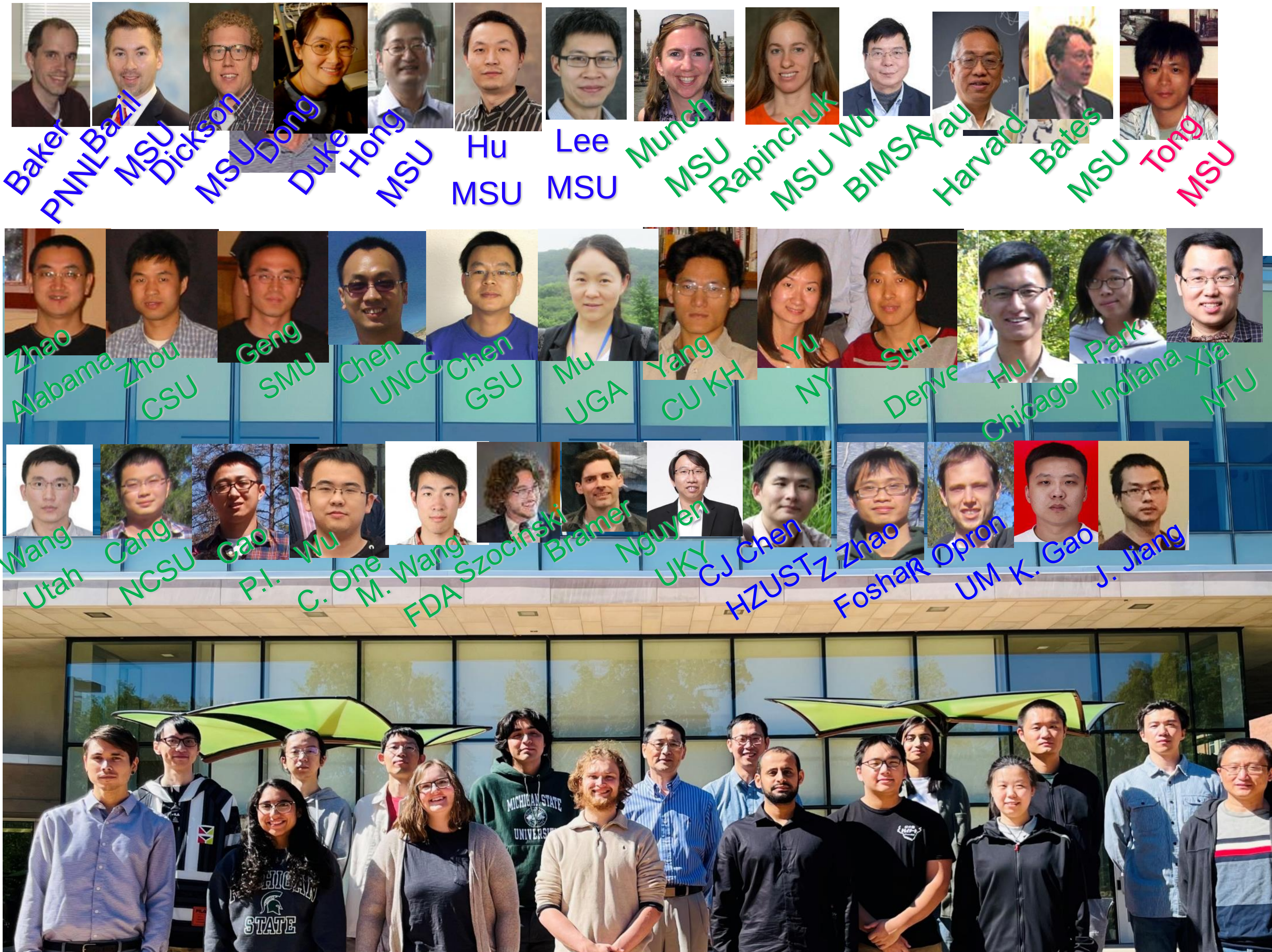
Real analysis

Stochastic analysis

**Vision:**

The last frontier of science and technology is biological science

The last frontier of biological science is mathematics


Baker
PNNL
Bazil
MSU
Dickson
MSU
Dong
Duke
Hong
MSU
Hu
MSU
Lee
MSU
Munch
MSU
Rapinchuk
MSU
Wu
BIMSA
Yau
Harvard
Bates
MSU
Tong
MSU
Zhao
Alabama
Zhou
CSU
Geng
SMU
Chen
UNCC
Chen
GSU
Mu
UGA
Yang
CU KH
Yu
NY
Sun
Denve
Hu
Chicago
Park
Indiana
Xia
NTU
Wang
Utah
Cang
NCSU
Cao
P.I.
Wu
C. One
M. Wang
FDA
Szocinski
Bramer
Nguyen
UKY
CJ Chen
HZUST
Z Zhao
Foshan
R Opron
UM
K. Gao
J. Jiang