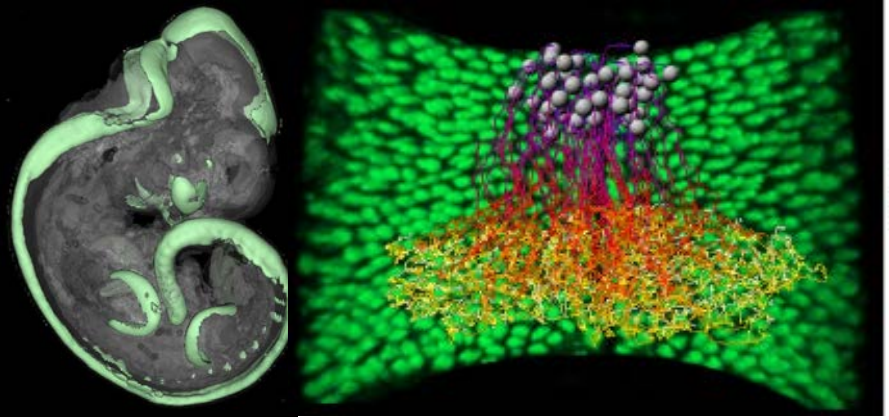
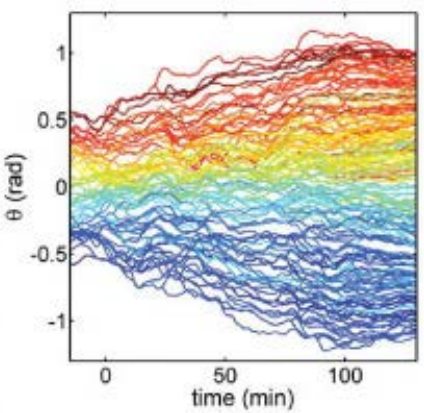
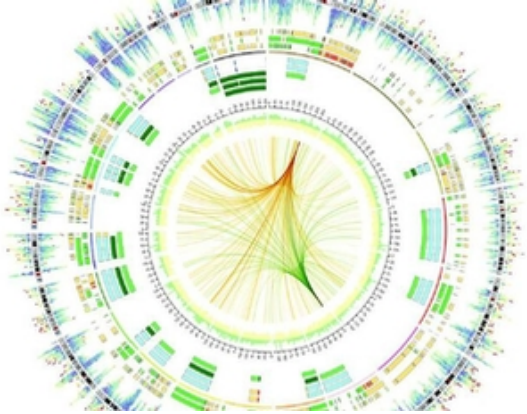




# *Beyond the Bench: What Skills are Needed for Next-Gen Developmental and Reproductive Toxicology?*

**Nisha S. Sipes, PhD**

ORISE Fellow

US EPA /ORD/ National Center for Computational Toxicology

Research Triangle Park, NC

July 1, 2014

The Teratology Society Annual Meeting, Bellevue, WA

Building a Successful Career in Developmental and Reproductive Toxicology: Part 2

http://www.tcd.ie/Zoology/research/research/development/; http://www.bbe.caltech.edu/content/systems-biology-1; http://www.genomequebec.com/32-en/news-new-genome-canada-competition-in-bioinformatics-and-computational-biology.html

# Disclosure

The authors of this research have no financial or other interests which pose a conflict of interest.

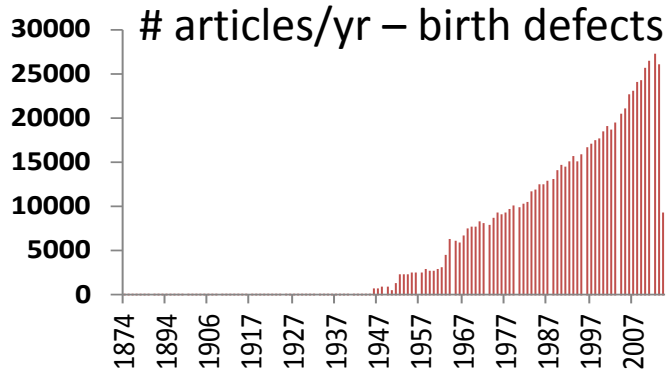
Disclaimer: The views expressed in this presentation are those of the author and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency

Disclaimer: Background -

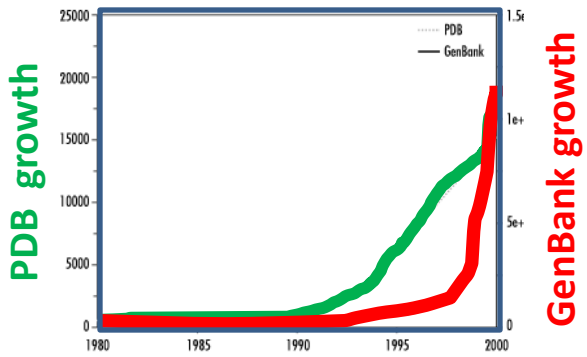
- Midwest (Cincinnati, Ohio)
- B.S. Mechanical Engineering
- M.S. Biomedical Engineering
- Ph.D. Cell and Cancer Biology

# Big Data

- Compiled a mass amount of data over many years
  - **PubMed stats as of June 29, 2014:**
    - Biology (1,110,062)
    - Dev or repro biology (310,816)
    - Toxicology (306,939)
    - Birth defects (775,358)




- New technology allows the collection of a lot of data faster
  - **Next-Gen Sequencing** *applies to genome sequencing, resequencing, transcriptome profiling (RNA-seq), DNA-protein interactions (ChIP-seq), & epigenome character**
  - **High-Throughput Screening** *Using robotics, data processing and control software, liquid handling devices, and sensitive detectors, High-throughput screening allows a researcher to quickly conduct millions of chemical, genetic, or pharmacological tests.** *wikipedia




http://oreilly.com/catalog/bioskills/chapter/ch01.html

# Have you done any of these?

- Next-Gen sequencing
- HTS/HCS data
- Microarray
- 3D imaging stacks
- time-lapse imaging data
- Performed or received sequencing data (can even be on vectors, constructs)

An opportunity obtain a wealth of data and to ask complex questions about the data

# ‘Next-Gen’

intersection of biology and computational sciences dedicated to using big data to find out how living things work

- Computational Biology *science of using biological data to develop algorithms and relations among various biological system**
- Bioinformatics *science of collecting and analyzing complex biological data such as genetic codes**
- Systems Biology *biology-based inter-disciplinary field of study that focuses on complex interactions within biological systems, using a holistic approach (instead of the more traditional reductionism) to biological and biomedical research**
- Systems Toxicology *decoding the toxicological blueprint of active substances that interact with living systems+*



*wikipedia; +Sturla et. al. 2014 CRT 27: 314–329

# What can we do with the data?

- Develop & use tools to order, manage, compile, mine and model the data to find associations & patterns, statistical significance
  - Find new connections
  - computer scientists, mathematicians, engineers, statisticians
  - pitfall: predictions may be made w/out complete understanding of where the biology comes from and what it means
- Biologists & toxicologists
  - bring context & meaning
  - use tools to ask complex questions about the data



*Over generalization, but you get the idea

# Diverse and Collaborative

- moving forward

- Multidisciplinary groups:
  - Biologists, toxicologists, computer scientists, mathematicians, engineers, statisticians, physicists

- Multidisciplinary skill sets:
  - In order to talk with one another & understand, develop, use and update the tools (abilities & limitations)

# What basic skills do we need?

- Core biological knowledge & understanding
- How to ask s
- Formulate h
- Design and c nent to test your hypoth

Guess what?

You already have th eeded to analyze big data (bi nal)

# Huh? I don't know how to program or use computer code

- Developing & using code is just like designing an experiment
  - Think critically
  - Design a process to answer a question
  - Understand what is required to answer the question unambiguously
- Many resources for learning to program or for using computational tools
  - No amount of reading can make you a good programmer
  - Analogous to learning a spoken language
  - You need to practice, practice, practice

# What can I do with my computational skills?

- Develop predictive models (statistical & simulation)

## Predictive signatures

Rat

| Feature | Description | Weight |
|---|---|---|
| RAR | Retinoic Acid receptor | 0.58 |
| GPCR | G-Protein-Coupled Receptors | 0.55 |
| TGFβ | Transforming Growth Factor β | 0.38 |
| MT | Microtubule organization | 0.30 |
| SENS_CYP | Cytochrome P450 (sensitive) | 0.26 |
| AP1 | Activator protein 1 | 0.24 |
| SLCO1B1 | Organic anion transporter 1B1 | 0.11 |
| CYP | CYPs (other) | 0.06 |
| HLA-DR | MHC complex | -0.38 |
| PXR | Pregnane X receptor | -0.24 |
| IL8 | Interleukin 8 | -0.23 |
| PGE2 | Prostaglandin E2 response | -0.18 |

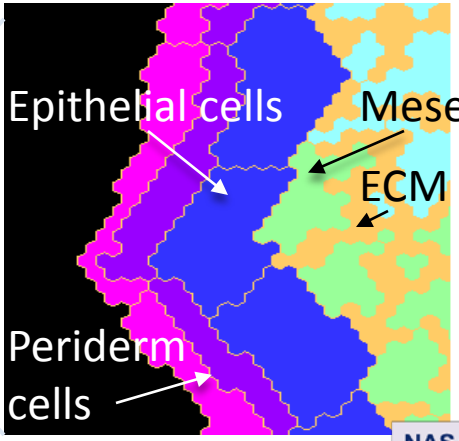
Rabbit

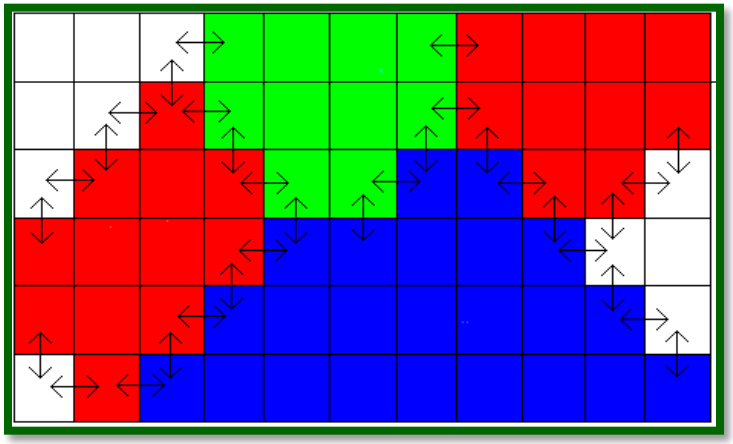
| Feature | Description | Weight |
|---|---|---|
| CCL2 | Chemokine ligand 2 (MCP1) | 1.15 |
| IL | Interleukin (1a and 8) | 0.39 |
| CYP | Cytochrome P450 | 0.24 |
| TGFβ | Transforming Growth Factor β | 0.28 |
| MESC | Mouse ES cells (J1) | 0.13 |
| SULT2A1 | Sulfotransferase | -0.26 |
| PGE2 | Prostaglandin E2 response | -0.15 |

Sipes et. al. 2011 Tox Sci 124:109-27

# Cell *Agent-Based Models* (ABMs)

Palatal fusion



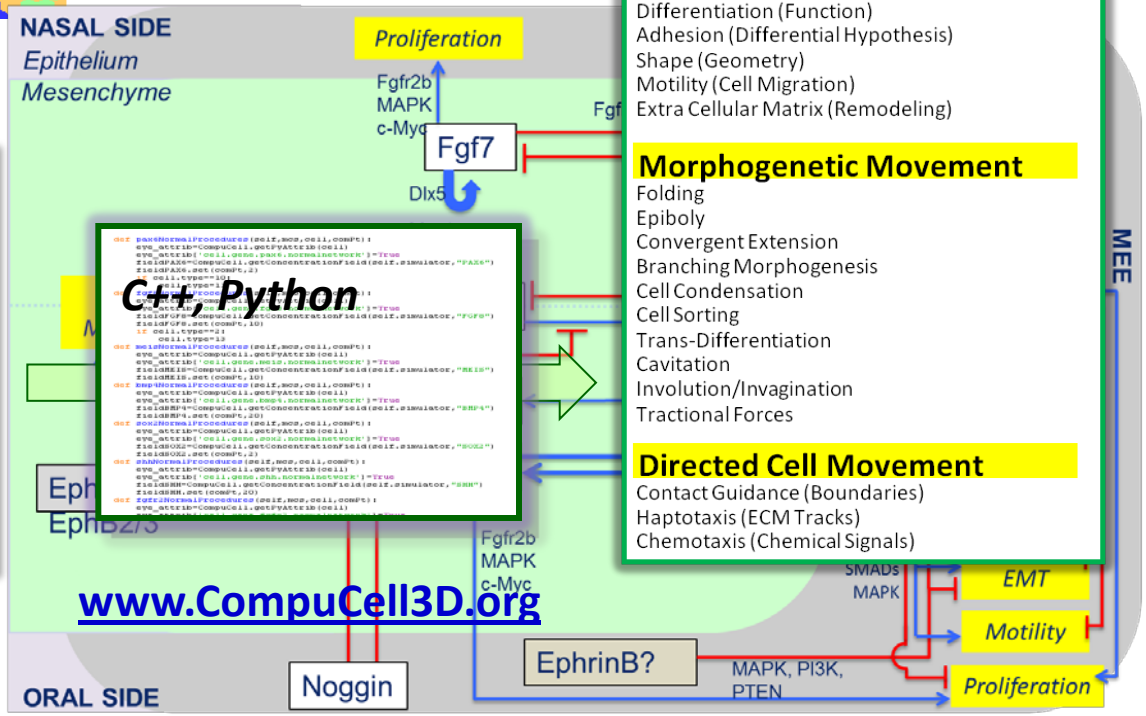



**Core Developmental Processes**
Patterning (Sets up Future Events)
Timing (Clocks and Oscillators)
Differentiation (Cell Diversification)
Morphogenesis (Tissue Organization)

**Cellular Primitives**
Growth (Proliferation)
Growth (Volume Increase)
Death (Apoptosis)
Differentiation (Function)
Adhesion (Differential Hypothesis)
Shape (Geometry)
Motility (Cell Migration)
Extra Cellular Matrix (Remodeling)

**Morphogenetic Movement**
Folding
Epiboly
Convergent Extension
Branching Morphogenesis
Cell Condensation
Cell Sorting
Trans-Differentiation
Cavitation
Involution/Invagination
Tractional Forces

**Directed Cell Movement**
Contact Guidance (Boundaries)
Haptotaxis (ECM Tracks)
Chemotaxis (Chemical Signals)



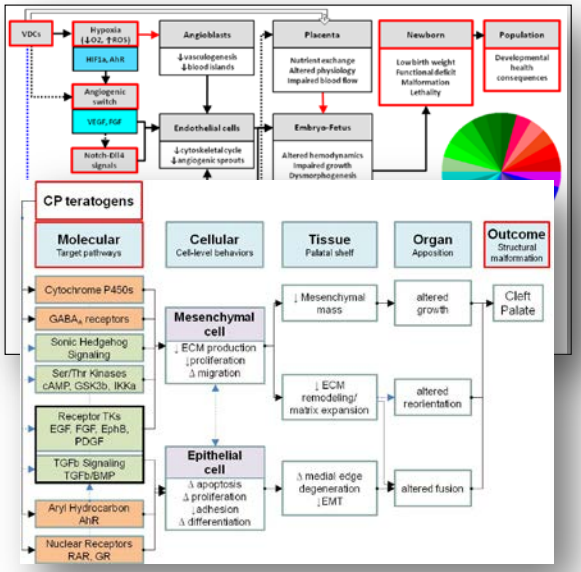

www.CompuCell3D.org

*How can we model the underlying biology and cellular dynamics of morphogenetic fusion?*

# What can I do with my computational skills?

- Find relevant relationships among mass amount of data (e.g., genetic, HTS, literature)

Molecular targets & pathways




Knudsen & Kleinstreuer 2011 BDRC 93: 312-323.;
Sipes et al manuscript in preparation

# Other valuable skills

- **Data mining** to extract data from multiple resources is invaluable
- **Data visualization skills** will help you take complex data and interpret it
- **Familiarity with bioinformatic resources** such as the UCSC genome browser and Entrez, NCBI, ToxCast
- **Familiarity with bioinformatic tools** such as Blast, DAVID, Cytoscape for sequence analysis, clustering & visualization

# What soft skills do we need?

- **Communication & Teamwork** will be working on teams with people who have diverse backgrounds & differing areas of expertise
- **Multi-task** learning while maintaining workload; taking advantage of downtime
- **Risk-taking** getting out of comfort zone and trying something new
- **Drive & persistence** it's not going to be simple and takes time to learn & get results (just like wet-lab experiments)

# Careers

## There are many types of positions, no one background is ideal for all of them

Scientific Curator

Gene/protein Analyst

Developmental/Reproductive Scientist

Research Scientist

Database programmer

Bioinformatics software developer

Computational Biologist/Toxicologist

Systems Biologist/Toxicologist

Structural Analyst

Molecular Modeler

Biostatistician, Database programmer, Cheminformatician, Pharmacogenetician, Pharmacogenomics, Toxicologist

Because the field is so new almost everyone in it did something else before

# Careers

There are distinct categories of professionals that the industry needs:

1. Computer Programmers, Mathematicians and people trained in Physics, Statistics etc. who develop software tools and applications for biotechnology and life science companies. They are cross trained in life sciences, such as molecular biology, DNA sequence analysis and in addition that they would need skills in writing algorithms and codes for developing such programs. A very specific training is required for such professionals to meet the need of life sciences companies.

2. People with a background in life sciences who are the end users of such programs and packages and they use these tools to translate the information into tangible products such as new molecules, drugs, enzymes etc. They can conduct their R&D program more effectively if they are cross-trained in computing skills. They can also be Business Analysts for life science companies.

http://bioinformaticsweb.net/carrier.html

# Be **ACTIVE** in your career!

- If you are interested in this field find the appropriate resources and learn

- Resources @ schools/institutions
  - Courses (audit)
  - Departments
  - Seminars
  - Faculty & Staff

# Teratology Society student/postdoc member institutions

Boston University
Brown University
California State University Northridge
Central Michigan University Coll of Med
Children's Hospital of Philadelphia
Creighton University School of Medicine
Emporia State University
Harvard University
Health Canada
Hospital for Sick Children
Johns Hopkins University
Mashhad University of Medical Sciences
MassGeneral Hospital for Children
McGill University
NCTR/FDA
NICHD
NIEHS
Northeastern University
Queen's University
RIVM
Sahlgrenska Academy at Göteborg Univ
Simmons College
South Dakota State University
Susquehanna University
Texas A&M Health Science Center
Thomas Jefferson University
Tulane University
University of Adelaide
University of Alabama
University of Alabama Birmingham
University of British Columbia
University of California, San Diego
University of California, San Francisco
University of Colorado
University of Florida
University of Georgia
University of Iowa
University of Lagos
University of Manchester
University of Maryland School of Medicine
University of Montreal
University of New Mexico
University of North Carolina, Chapel Hill
University of Rochester
University of Toronto
University of Washington
US EPA
Vanderbilt University Medical Center
Virginia Maryland Regional Coll of Vet Med
Virginia Tech
Yale University

THE TERATOLOGY SOCIETY

Teratology Society Members can connect via BDR connection

http://www.teratology.org/connection.asp

# Listen to presentations

- http://www.epa.gov/ncct/communities_of_practice.html
- http://www.toxicology.org/AI/ce/ce_video/index.asp#search (free for SOT student & postdoc members)

# Free online databases & tools

- http://www.ncbi.nlm.nih.gov/tools/gbench/
- https://genome.ucsc.edu/
- http://david.abcc.ncifcrf.gov/
- http://www.cytoscape.org/
- http://www.r-project.org/
- http://guides.library.duke.edu/datavis/
- http://epa.gov/ncct/toxcast/data.html
- http://actor.epa.gov/actor/faces/CSSDashboardLaunch.jsp

# Conferences

- http://www.issb.org/
- http://www.icsb14.com/pages/about-us.php
- http://www.conference-service.com/conferences/mathematical-biology.html
- http://integrativebiology2014.conferenceseries.net/

# Training

- https://www.systemsbiology.org/advanced-courses
- http://bioinformatics.ca/workshops/2014
- http://training.bioinformatics.ucdavis.edu/
- http://hermes.mbl.edu/education/courses/special_topics/index.html
- http://www.broadinstitute.org/collaboration/cegs/outreach_workshop

# Websites

- http://www.systemscenters.org/
- http://ccbs.uci.edu/
- http://csbi.mit.edu/
- http://www.systemsbiology.org/
- http://genomesciences.wustl.edu/
- http://www.nyas.org/whatwedo/fos/systemsbio.aspx
- http://www.systemscenters.org/outreach-education/research-positions/faculty-development/
- http://www.epa.gov/risk/nexgen/workshops.htm

# Papers, books, reading materials

- http://www.ploscompbiol.org/article/info%3Adoi%2F10.1371%2Fjournal.pcbi.1000589
- http://oreilly.com/catalog/bioskills/chapter/ch01.html

# Computational tools make it possible to make potentially important discoveries

"We can't overstate the importance of understanding the limitations of these tools. But once you gain that understanding and become an intelligent consumer of bioinformatics methods, the speed at which your research progresses can be truly amazing."



http://oreilly.com/catalog/bioskills/chapter/ch01.html

# Thank You!

Additional questions?

Email me! nisha.sipes@gmail.com

Want to connect?

LinkedIn me!