



CPW COMPUTATIONAL
PATHOLOGY WORKSHOP
SECOND EDITION
SEPTEMBER 9, 2018, ATHENS, GREECE
www.cpw2018.org

Digital Pathology and its importance as an omics data layer

Yves Sucaet, PhD

A word from our sponsors

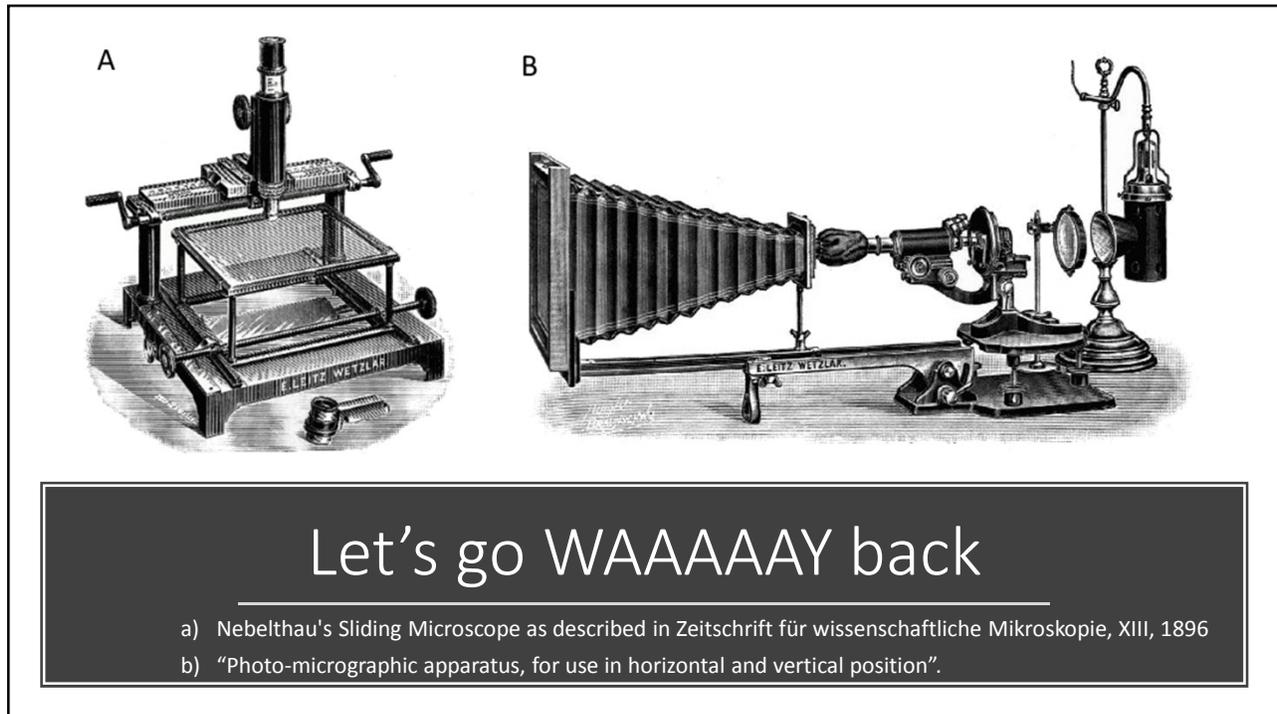


Personal financial disclosure

- I'm a co-founder and shareholder in Pathomation.
- I'm affiliated with HistoGeneX in the role of data scientist

Permit me a
small trip down
memory lane





Let's go WAAAAAY back

- a) Nebelthau's Sliding Microscope as described in Zeitschrift für wissenschaftliche Mikroskopie, XIII, 1896
 b) "Photo-micrographic apparatus, for use in horizontal and vertical position".



My educational background

- BS. Computer science
- MS. Biological sciences
 - Replication between transcription termination and replication initiation through in vivo and in silico study of Autonomous Replication Sequences
 - 2003: Keynote by Manolis Kellis at YGMB Göteborg
- PhD. Bioinformatics
 - Network biology with plant modeling systems
 - In silico pathway integration of heterogeneous datasets: TAIR, AraCyc, atPID

After graduation



Joined an unknown CRO in Belgium: HistoGeneX

2010: About 30 people in Antwerp
Today: 110+ people in Belgium, US, planning a third site in China



The challenge: develop complementary bioinformatics activities to a mostly wet-lab based product portfolio

Veerle was there when I was hired ☺



Data Fusion and & 21st Century Diagnostics

Fused Diagnostics
Integrated Diagnostics
Advanced Analytics
Computational Diagnostics
Personalized Predictive Modeling

Pathology Visions (DPA)
2012, Baltimore

I want pixels

Many scanners

Much frustration

But wait, how is this possible at scale?

2013: the birth of Pathomation

2014: Digital pathology is introduced at VUB



2016: First CPW SIG at ECCB, The Hague



ISMB 2017, Prague

In his keynote, Dr. Bock states that "Imaging is the new omics"

2017 ISCB OVERTON PRIZE
CHRISTOPH BOCK



2018: Digital pathology is operational at VUB

Enhanced network and local cloud infrastructure

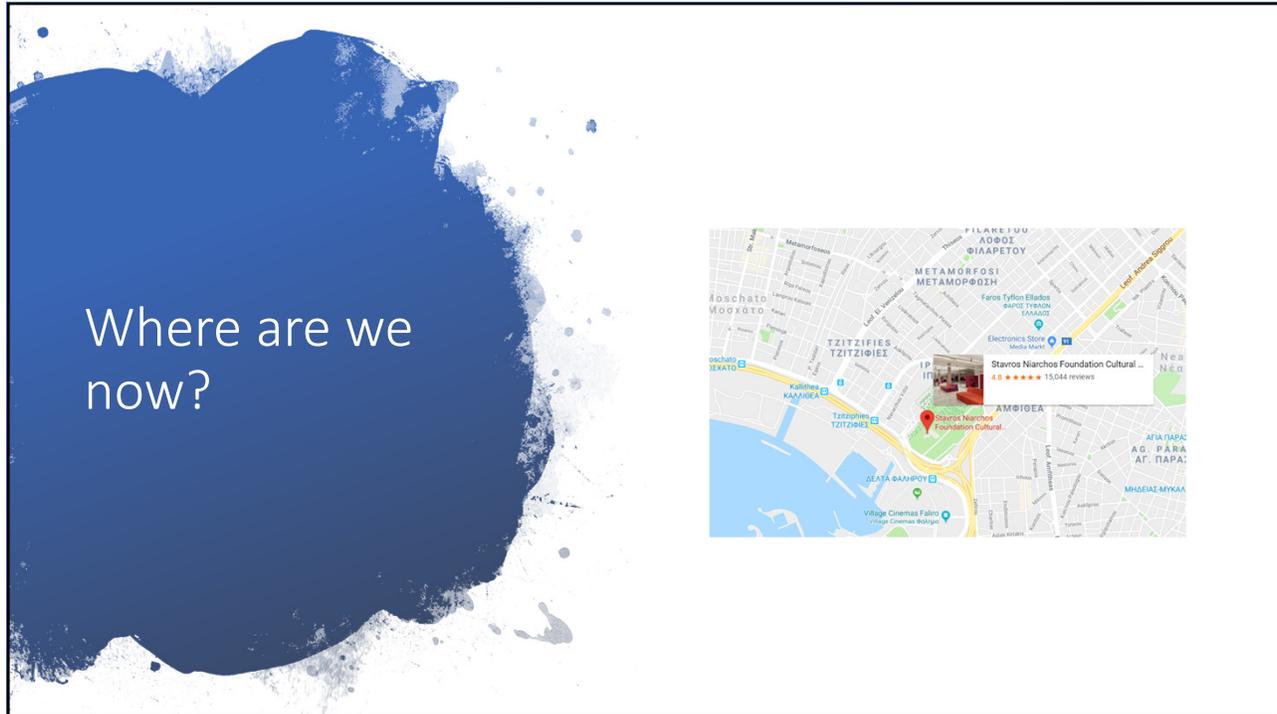
Education

Research

Biobanking

<http://www.dabnetbiobank.org>





Bioinformatics + digital pathology = computational pathology

- The term was coined first by Dr. Thomas Fuchs
- Other terms used: augmented pathology, integrated / integrative pathology, histomics, ...
- Many workshops now on the subject, at API, DPA, Global Engage...
- CPW is the only event organized at the bioinformatics community level!

Computational Pathology Analysis of Tissue Microarrays Predicts Survival of Renal Clear Cell Carcinoma Patients

February 2008

DOI: 10.1007/978-3-540-85990-1_1

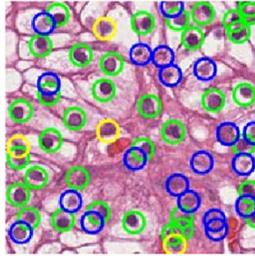
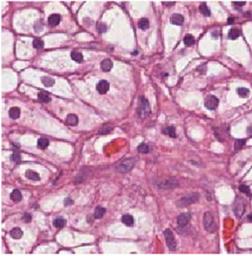
Source · [PubMed](#)

Conference: Proceedings of the 11th International Conference on Medical Image Computing and Computer-Assisted Intervention, Part II

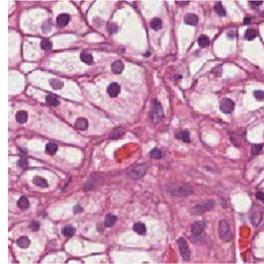
Thomas J Fuchs · Peter J Wild · Holger Moch · Joachim M Buhmann

Bioinformatics vs. Pathology

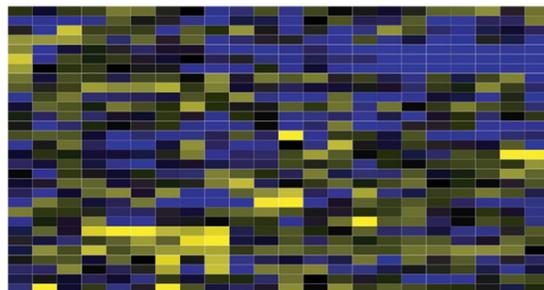
Pathology: Tissue is the issue



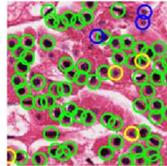
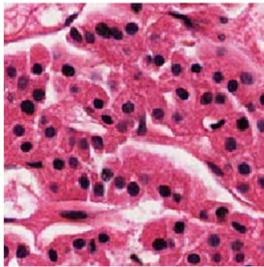
Bioinformatics: high-throughput



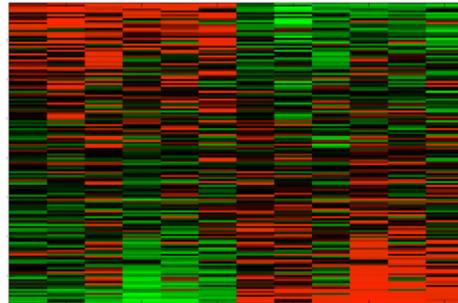
The challenge: bridging experiments



Combining the toolboxes from both fields



- Diseased vs. healthy tissue, patient stratification, better predictive markers



A concrete example

Bioinformatics, 32(1), 2016, 122–129
doi: 10.1093/bioinformatics/btv542
Advance Access Publication Date: 11 September 2015
Original Paper



Bioimage informatics

CD30 cell graphs of Hodgkin lymphoma are not scale-free—an image analysis approach

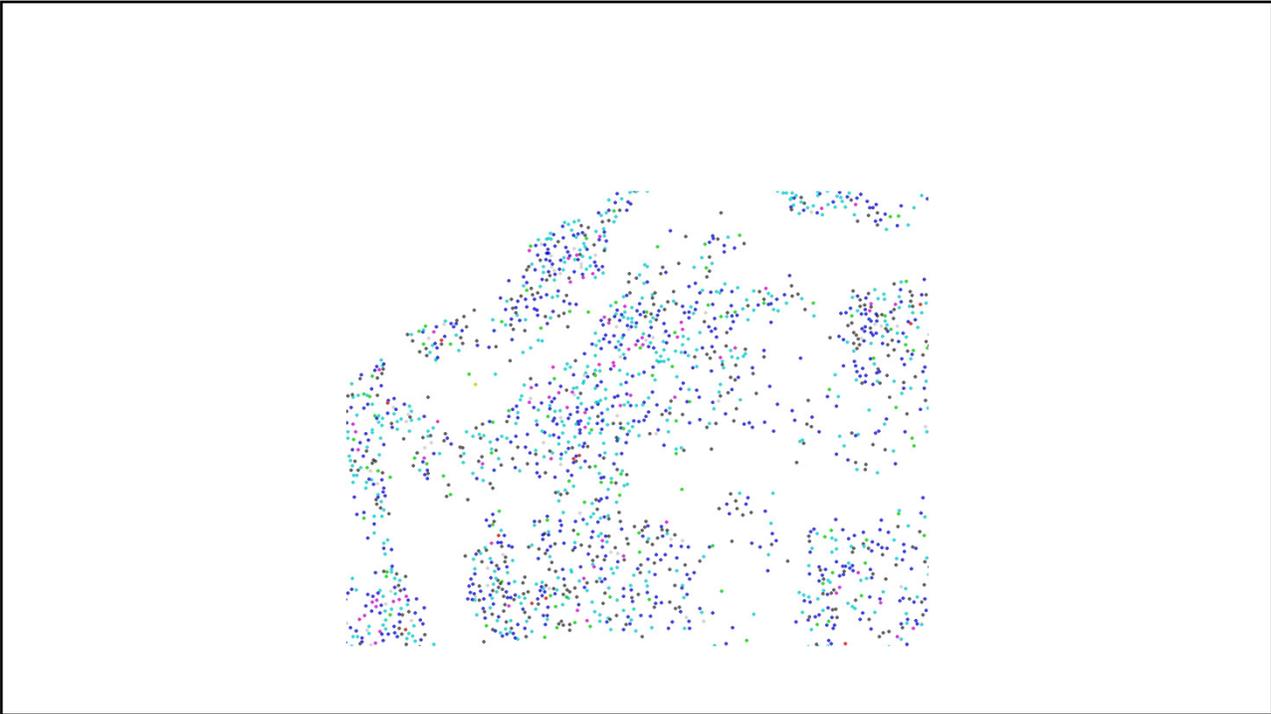
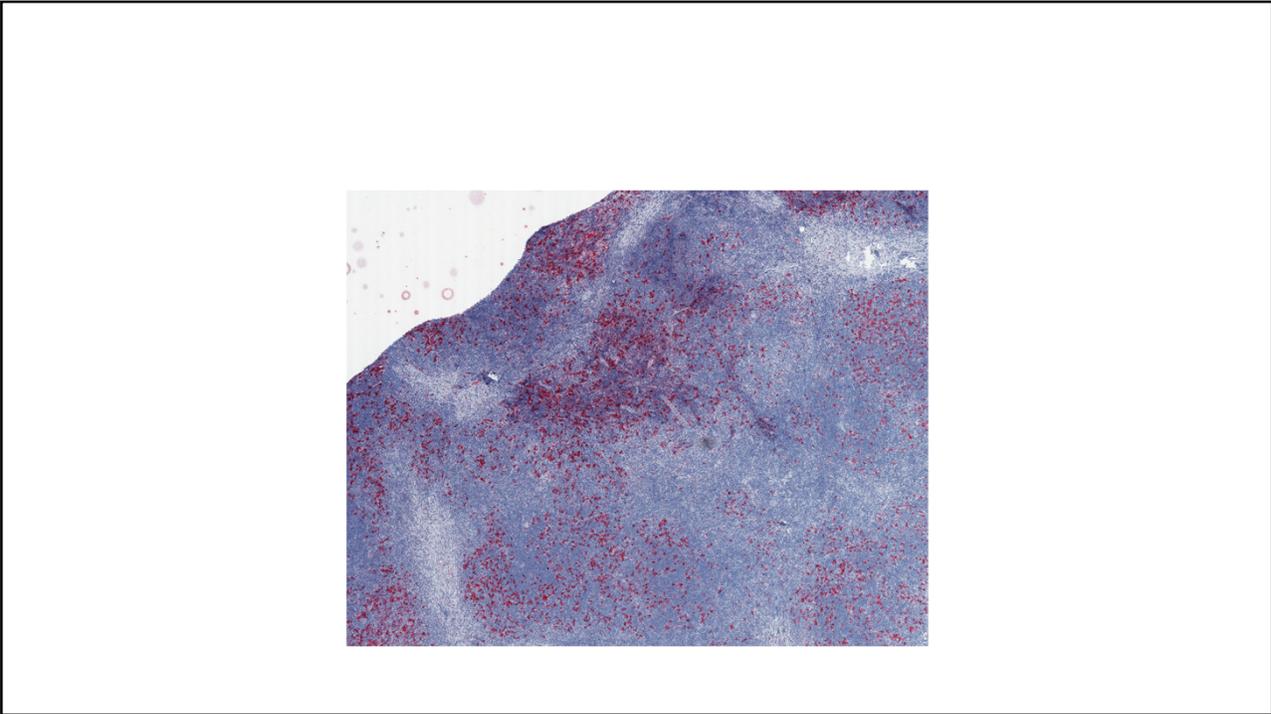
Hendrik Schäfer^{1,†}, Tim Schäfer^{1,†}, Jörg Ackermann¹, Norbert Dichter¹,
Claudia Döring², Sylvia Hartmann², Martin-Leo Hansmann²
and Ina Koch^{1,*}

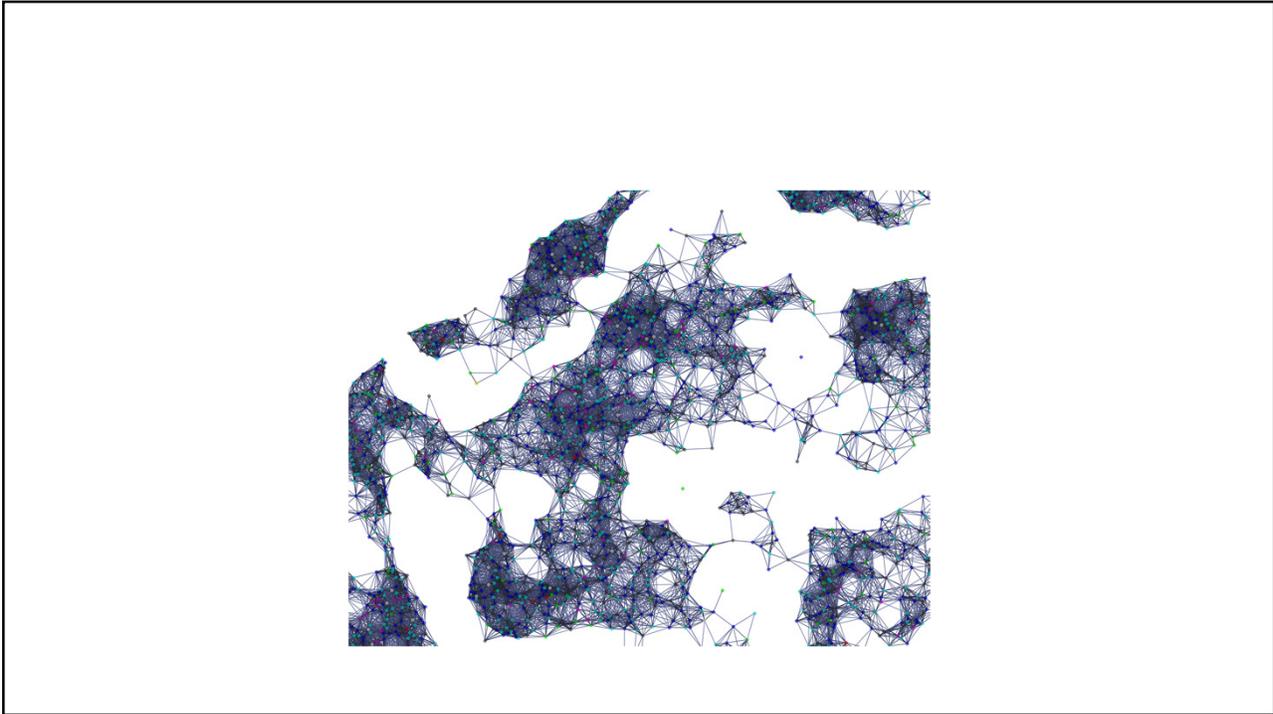
¹Department of Molecular Bioinformatics, Institute of Computer Science, Cluster of Excellence Macromolecular Complexes, Johann Wolfgang Goethe-University Frankfurt am Main, Robert-Mayer-Str. 11-15, 60325 Frankfurt am Main and ²Dr. Senckenbergisches Institut für Pathologie, Universitätsklinikum Frankfurt, Theodor-Stern-Kai 7, 60590 Frankfurt am Main, Germany

*To whom correspondence should be addressed.

[†]The authors wish it to be known that, in their opinion, the first two authors should be regarded as Joint First Authors.
Associate Editor: Robert Murphy

Received on May 7, 2015; revised on September 7, 2015; accepted on September 8, 2015





So why not leverage Cytoscape?

A screenshot of the Cytoscape App Store interface. On the left, there is a sidebar with 'Categories' including 'network analysis', 'data visualization', 'network generation', and 'graph analysis'. The main area is titled 'Network analysis' and contains a grid of application cards. Each card features an icon, the app name, a brief description, and a version number. The apps listed include ARACNE, Cytoscape, Cytoscape 3.8.2, Cytoscape 3.8.1, Cytoscape 3.8.0, Cytoscape 3.7.5, Cytoscape 3.7.4, Cytoscape 3.7.3, Cytoscape 3.7.2, Cytoscape 3.7.1, Cytoscape 3.6.1, Cytoscape 3.6.0, Cytoscape 3.5.4, Cytoscape 3.5.3, Cytoscape 3.5.2, Cytoscape 3.5.1, Cytoscape 3.4.0, Cytoscape 3.3.0, Cytoscape 3.2.0, Cytoscape 3.1.0, Cytoscape 3.0.0, Cytoscape 2.9.0, Cytoscape 2.8.0, Cytoscape 2.7.0, Cytoscape 2.6.0, Cytoscape 2.5.0, Cytoscape 2.4.0, Cytoscape 2.3.0, Cytoscape 2.2.0, Cytoscape 2.1.0, Cytoscape 2.0.0, Cytoscape 1.9.0, Cytoscape 1.8.0, Cytoscape 1.7.0, Cytoscape 1.6.0, Cytoscape 1.5.0, Cytoscape 1.4.0, Cytoscape 1.3.0, Cytoscape 1.2.0, Cytoscape 1.1.0, Cytoscape 1.0.0. The cards are arranged in a grid and are separated by thin lines.

Another opportunity for bioinformatics?

cytoscape-helpdesk
Inferring edges from positional information
 3 posts by 3 authors

Yves Sucaet
 ☆ Other recipients: Jordy.St., @vub.be

Hello,

I have a datfile that looks like this:

```
NODE_POS_X_POS_Y
node01,15,20
node02,12,25
node03,14,22
etc...
```

I'm able to import this network into Cytoscape and map the X- and Y-coordinates, to present me with a network that shows the relative positions of the nodes to each other.

How can I now automatically add edges to this network, based on some distance metric calculated from the positional node information?

Further background: the network contains about 2000-3000 nodes currently. The math is straightforward, and I clearly could just transform my input-file, but something tells me someone else must have thought about this before me and probably already wrote a plugin for it. I would probably be that it would be v... between the nodes vary based on th... an edge when nodeA and nodeB are...

Any suggestions?
 Thanks in advance,
 Yves Sucaet

Scotter Morris
 ☆ Hi Yves,
 I'm not aware of anyone who has written such a plugin. Shouldn't be hard, though.
 --scotter
 (show quoted text)

Building a bridge to Cytoscape

The screenshot displays the Cytoscape interface. The main window shows a dense network of nodes and edges, rendered in a circular layout. On the left, the 'Control Panel' is visible, showing various settings for the network, including 'Node Size', 'Node Color', and 'Edge Width'. A 'Results Panel' is open in the foreground, displaying network statistics for 'Positional and spring'.

Results Panel: Network Statistics of Positional and spring (Undirected)

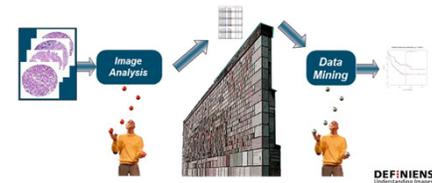
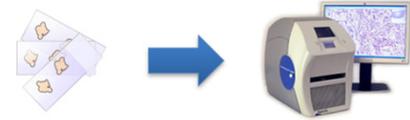
Simple Parameter	Node Degree Distribution	Betweenness Centrality	Closeness Centrality	Stress Centrality Distribution
Connected components	3032	Network Centrality: Undefined	Network Centrality: Undefined	Network Centrality: Undefined
Network centralization	0.0	Network Centrality: Undefined	Network Centrality: Undefined	Network Centrality: Undefined
Shortest paths	0 (0%)	Network Centrality: Undefined	Network Centrality: Undefined	Network Centrality: Undefined
Avg. number of neighbors	0.0	Network Centrality: Undefined	Network Centrality: Undefined	Network Centrality: Undefined
Number of nodes	2932	Network Centrality: Undefined	Network Centrality: Undefined	Network Centrality: Undefined
Number of edges	265,791	Network Centrality: Undefined	Network Centrality: Undefined	Network Centrality: Undefined
Network density	0.0	Network Centrality: Undefined	Network Centrality: Undefined	Network Centrality: Undefined

At the bottom right, a table shows node statistics:

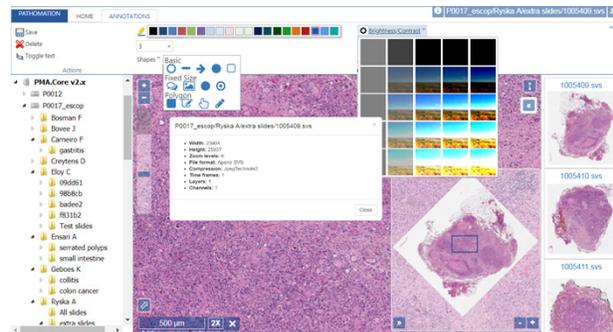
Statistics	name	Node (E-analysis 02_1)	Node (E-analysis 02_1)	Node (E-analysis 02_1)	Node (E-analysis 02_1)
	230	1	13	9	28
	231	2	9	3	14

Are we there yet?

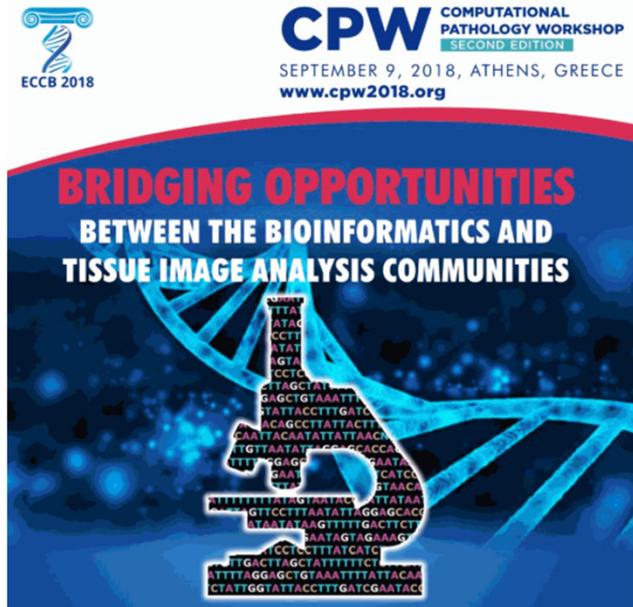
- Anno 2018, big divides remain, adaptation is slower than expected, and bridging communities is more necessary than ever
- Beacons of hope: “students don’t begin from scratch but enjoy a wealth of tools and code written before them. Most of our students likely don’t even realize that they’re using openslide or matlab as a backend because we’ve already written appropriate wrappers for our common tasks. In that context as well, us with our collaborators have established working protocols (e.g., file formats, scanners, etc) which our code is based around.”
- But (much) more work is definitely needed!



How can we attempt to address these issues?



Attend a
workshop



1) Don't build software for a single scanner

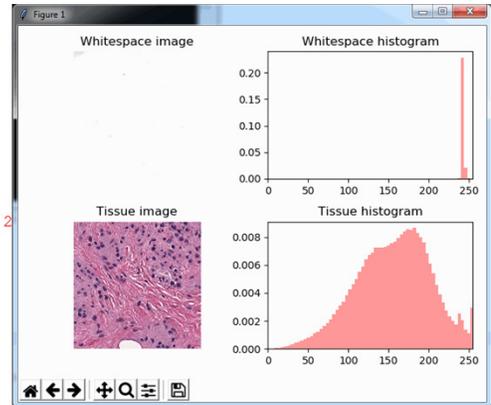
- It's (still) about the file formats
- Build your software on a digital pathology abstraction layer
 - PMA.start, OpenSlide, OMERO,...

PMA.start, Python, and OpenCV

```

1 from pma_python import pma
2 import numpy as np
3
4 def is_tissue(tile):
5     pixels = np.array(tile).flatten() # np refers to numpy
6     mean_threshold = np.mean(pixels) < 192 # 75th percentile
7     std_threshold = np.std(pixels) < 75
8     return mean_threshold == True and std_threshold == True
9
10 slide = "C:/my_slides/CMU-1.svs"
11 max_zl = pma.get_max_zoomlevel(slide)
12 dims = pma.get_zoomlevels_dict(slide)[max_zl]
13
14 tile_0_0 = pma.get_tile(slide, x=0, y=0, zoomlevel=max_zl)
15 print(is_tissue(tile_0_0))
16
17 tile_x_y = pma.get_tile(slide, x=int(dims[0] / 4), y=int(dims[1] / 2))
18 print(is_tissue(tile_x_y))

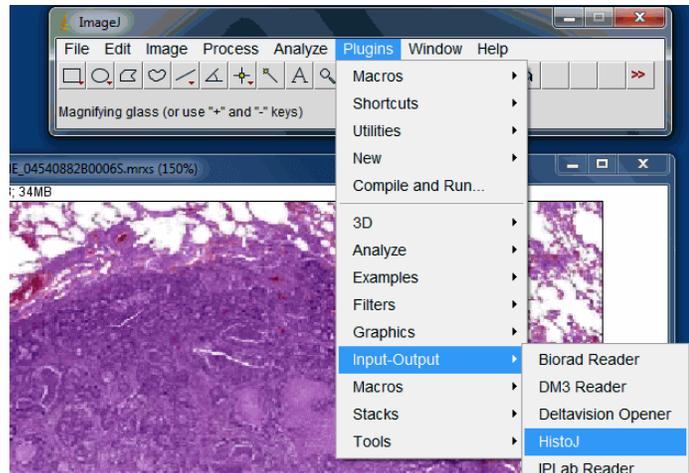
```



2) Don't reinvent the wheel

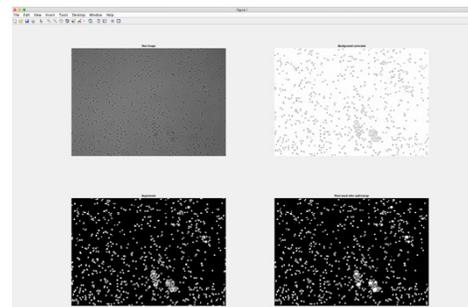
- But you can still improve the cars!
- Network analysis:
 - Cytoscape plugins for digital pathology?
- Expand AI environments so they become friendlier for image analysis:
 - AzureML Studio
- Contribute content, tutorials, presentations:
 - a DataCamp or Coursea MOOC on digital pathology?
 - Talk to Pavel Pevzner for a new section in the UCSD Bioinformatics curriculum?

Bringing WSI pixels into ImageJ

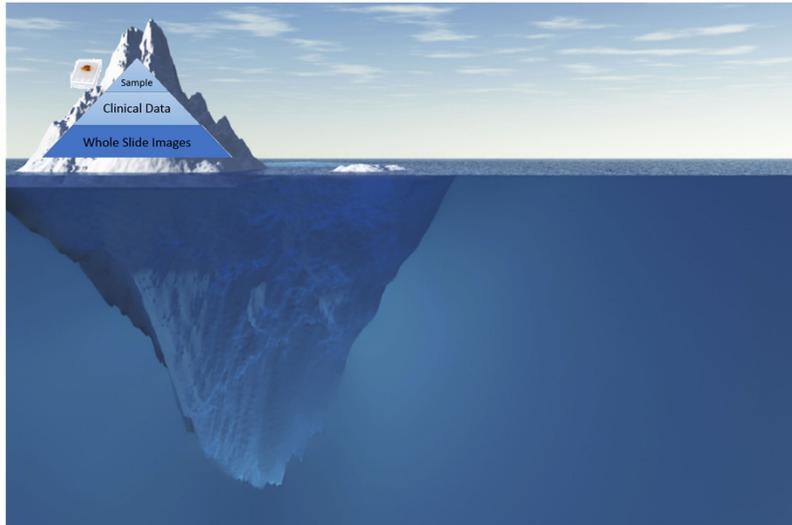


3) Usability and replicability are not the same

- Test your software / protocols / algorithms on data from others
 - People in your own lab (or even building) don't count
- This is not just about user-friendliness, it's about getting similar results on data generated on different imaging platforms
- Your segmentation algorithm may only work for you; what about others?



Tremendous opportunities still abound!



A banner for the Computational Pathology Workshop (CPW) 2018 in Athens, Greece. The background is a photograph of the Parthenon on the Acropolis. In the upper center, there is a logo for CPW (Computational Pathology Workshop) Second Edition, with the text 'SEPTEMBER 9, 2018, ATHENS, GREECE' and the website 'www.cpw2018.org'. Below the logo are three portrait photographs of men. At the bottom of the banner, there are two logos: 'PATHOMATION' on the left and 'HistoGeneX' on the right. A white text box at the bottom center of the banner contains the text 'Welcome to Athens!'.