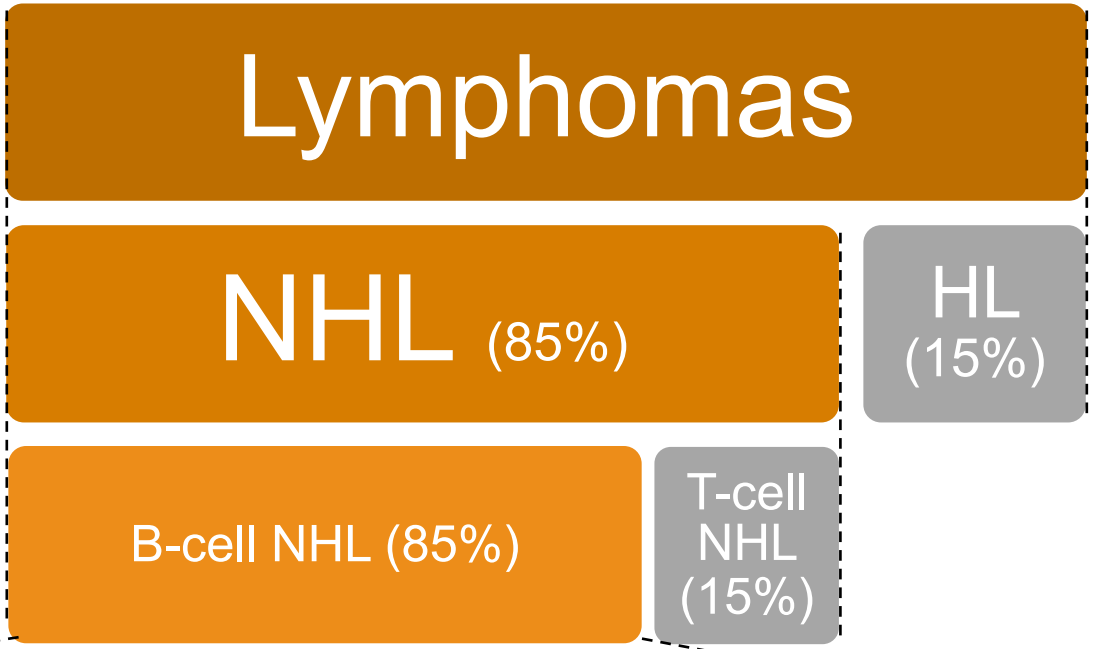
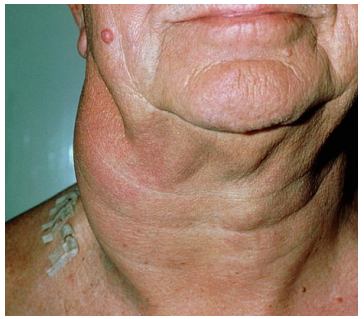
Characterizing Non-Hodgkin's Lymphoma with Patient Level Clinical Trial, Biomarker, and Real World Data

Joseph N. Paulson PhD

Non-Hodgkin Lymphoma: A Beginner's Guide

- **What are NHLs?**
  - A diverse group of cancers that develop in the lymphatic system (a network of tissue which includes the lymph nodes, bone marrow, spleen and thymus)
  - The majority arise from a type of white blood cell called *B cells*
  - NHLs can be classified according to how quickly they progress:
  - DLBCL and FL account for the majority of all NHLs





Indolent
- Follicular lymphoma
- Marginal zone lymphoma

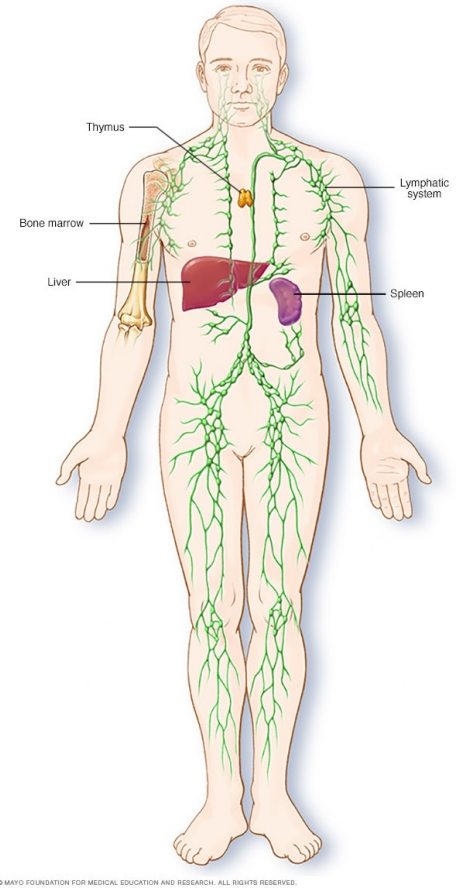
Aggressive
- DLBCL
- Mantle cell

Very aggressive
- Burkitt lymphoma
- B-Lymphoblastic

Edward Bataillard

Non-Hodgkin Lymphoma: An Overview

74,200 cases a year in the USA



Diffuse Large B-Cell Lymphoma (DLBCL) ~ 35%

- R-CHOP
  - 60% → Cure
  - 40% → Relapse
    - 25% → Rescue
    - 75% → Death

Percent Surviving 5 Years: 63.2% (2009-2015)

Follicular Lymphoma (FL) ~

- R-CHOP
  - 90% → Maintain
  - 10% → 10 year PFS

Percent Surviving 5 Years: 88.4% (2009-2015)

Patients who fail 1L (relapsed/refractory) have the worst outcomes, regardless of gene expression.

Diffuse Large B-Cell Lymphoma

- **What factors influence prognosis?**
  - Prognostic scores:
    - **IPI** *(pre-Ritux era)*
    - R-IPI *(post-Ritux era)*
    - NCCN-IPI
  - Biomarkers
    - *MYC* and *BCL2* and/or *BCL6* gene rearrangements ("double/triple hit lymphoma")
    - *MYC* and *BCL2* overexpression ("dual expressors")
    - Cell of origin (COO) according to gene expression profiling
      - Germinal centre B-cell type (**GCB**) → 3yr PFS 75%
      - Activated B-cell type (**ABC**) → 3yr PFS 40%

Risk factors	Risk category	3yr OS[2]
0 - 1	Low	91%
2	Low-intermediate	81%
3	High-intermediate	65%
4 - 5	High	59%

IPI Risk factors
Age > 60 years
Raised LDH
Performance status ≥ 2
Ann Arbor stage ≥ 3
Two or more extranodal sites of disease

- Not widely available in clinical practice
- Immunophenotyping algorithms are used as a surrogate, though with poorer correlation to treatment outcomes

Efforts in 1993 Led to Defining the International Prognostic Index based on Clinical Characteristics

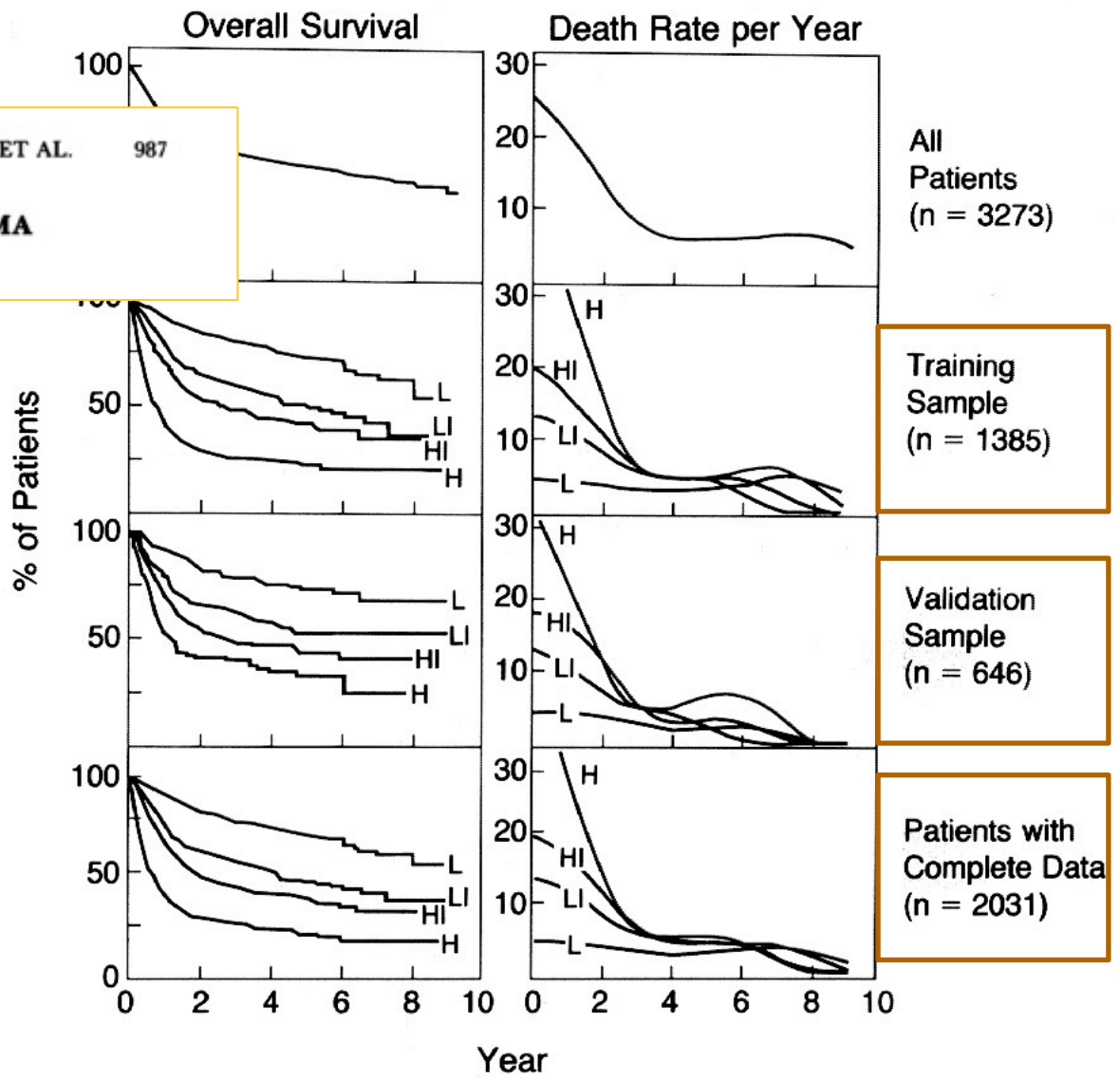
Vol. 329 No. 14 PREDICTIVE MODEL FOR AGGRESSIVE NON-HODGKIN'S LYMPHOMA — SHIPP ET AL. 987

A PREDICTIVE MODEL FOR AGGRESSIVE NON-HODGKIN'S LYMPHOMA

THE INTERNATIONAL NON-HODGKIN'S LYMPHOMA PROGNOSTIC FACTORS PROJECT*

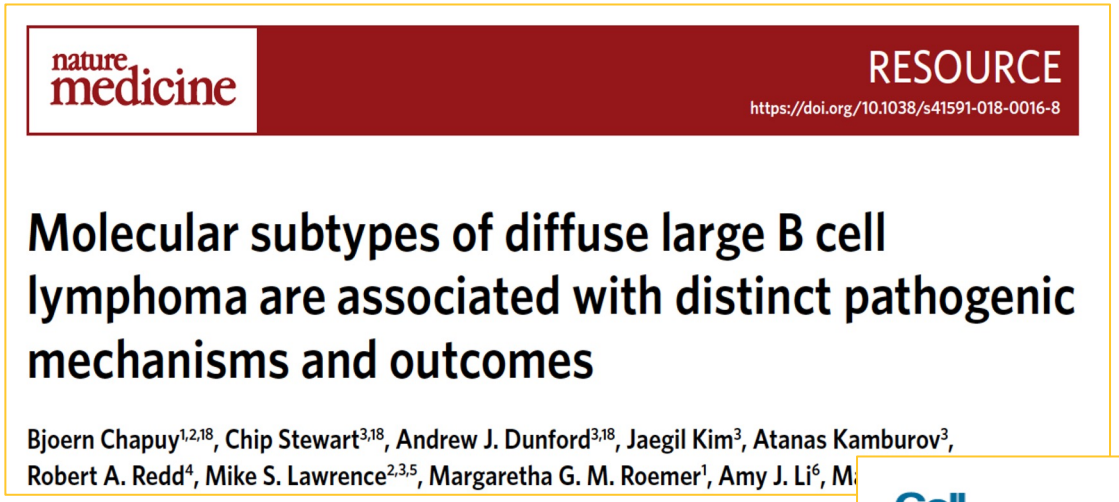
To develop a better prognostic-factor model for aggressive non-Hodgkin's lymphoma **16 institutions and cooperative groups** in the United States, Europe, and Canada participated **pooled together approximately 2,000 patients with complete data**

The step-down regression analysis of overall survival in the training sample evaluated 12 variables.



A Predictive Model for Aggressive Non-Hodgkin's Lymphoma
NEJM 1993

Efforts in 2019 Are Leading Toward Definitions of Novel Prognostic Subgroup Classifications Utilizing Genomics


nature medicine

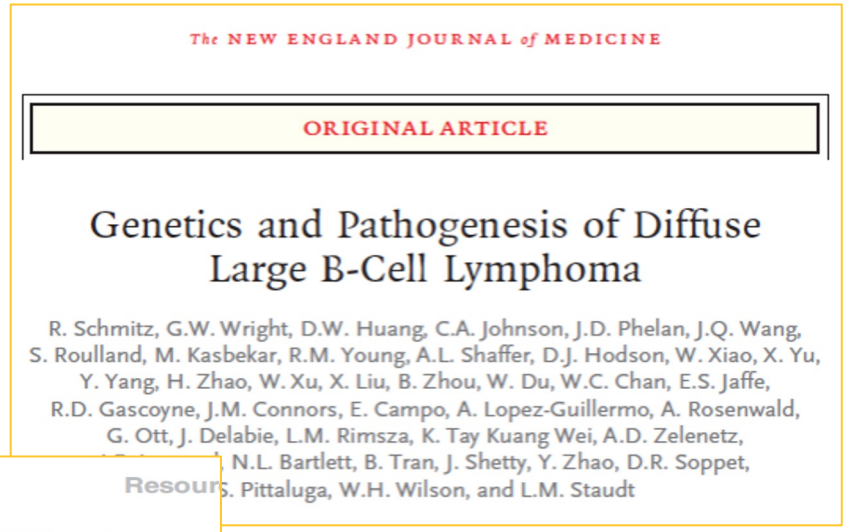
RESOURCE

https://doi.org/10.1038/s41591-018-0016-8

Molecular subtypes of diffuse large B cell lymphoma are associated with distinct pathogenic mechanisms and outcomes

Bjoern Chapuy[1,2,18], Chip Stewart[3,18], Andrew J. Dunford[3,18], Jaegil Kim[3], Atanas Kamburov[3], Robert A. Redd[4], Mike S. Lawrence[2,3,5], Margaretha G. M. Roemer[1], Amy J. Li[6], M

Unsupervised (n=~300)
Non-negative matrix factorization

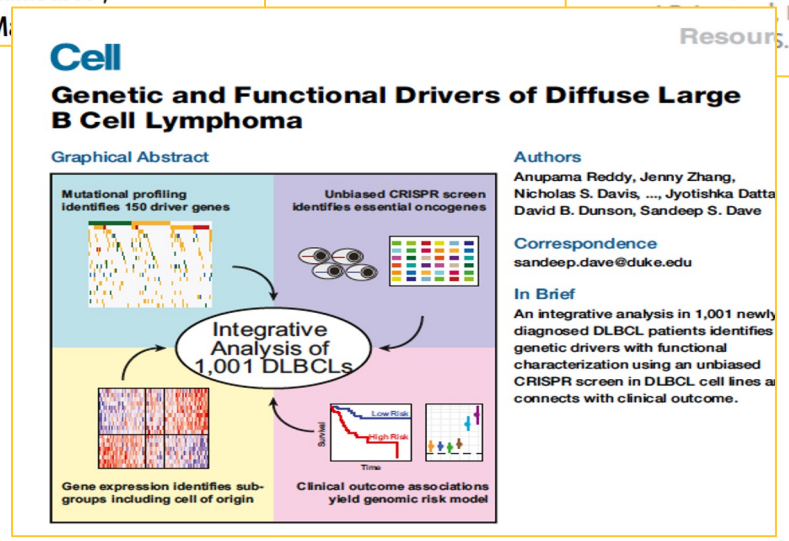

The NEW ENGLAND JOURNAL of MEDICINE

ORIGINAL ARTICLE

Genetics and Pathogenesis of Diffuse Large B-Cell Lymphoma

R. Schmitz, G.W. Wright, D.W. Huang, C.A. Johnson, J.D. Phelan, J.Q. Wang, S. Roulland, M. Kasbekar, R.M. Young, A.L. Shaffer, D.J. Hodson, W. Xiao, X. Yu, Y. Yang, H. Zhao, W. Xu, X. Liu, B. Zhou, W. Du, W.C. Chan, E.S. Jaffe, R.D. Gascoyne, J.M. Connors, E. Campo, A. Lopez-Guillermo, A. Rosenwald, G. Ott, J. Delabie, L.M. Rimsza, K. Tay Kuang Wei, A.D. Zelenetz, N.L. Bartlett, B. Tran, J. Shetty, Y. Zhao, D.R. Soppet, S. Pittaluga, W.H. Wilson, and L.M. Staudt

Semi-supervised (n=~560)
Seed based genetic algorithm


Cell

Genetic and Functional Drivers of Diffuse Large B Cell Lymphoma

Graphical Abstract

Authors
Anupama Reddy, Jenny Zhang, Nicholas S. Davis, ..., Jyotishka Datta, David B. Dunson, Sandeep S. Dave

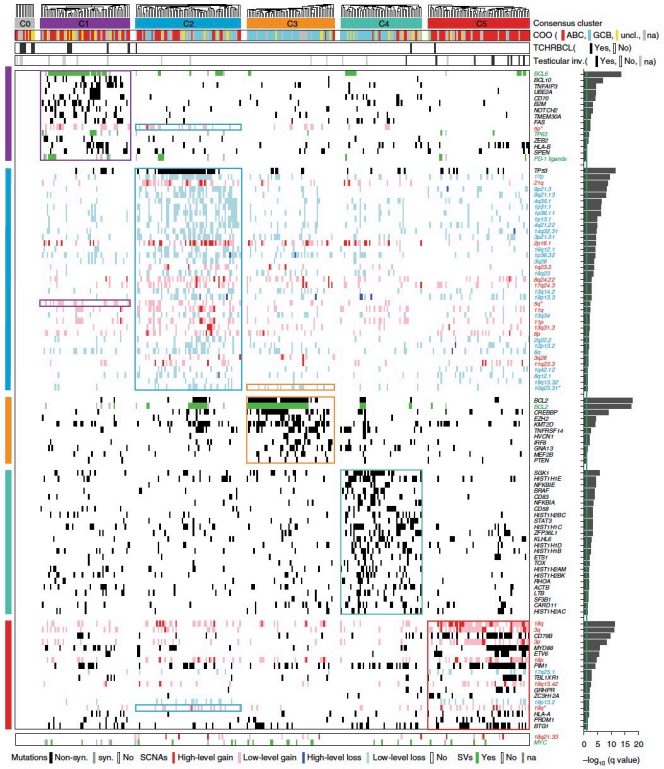
Correspondence
sandeep.dave@duke.edu

In Brief
An integrative analysis in 1,001 newly diagnosed DLBCL patients identifies genetic drivers with functional characterization using an unbiased CRISPR screen in DLBCL cell lines and connects with clinical outcome.

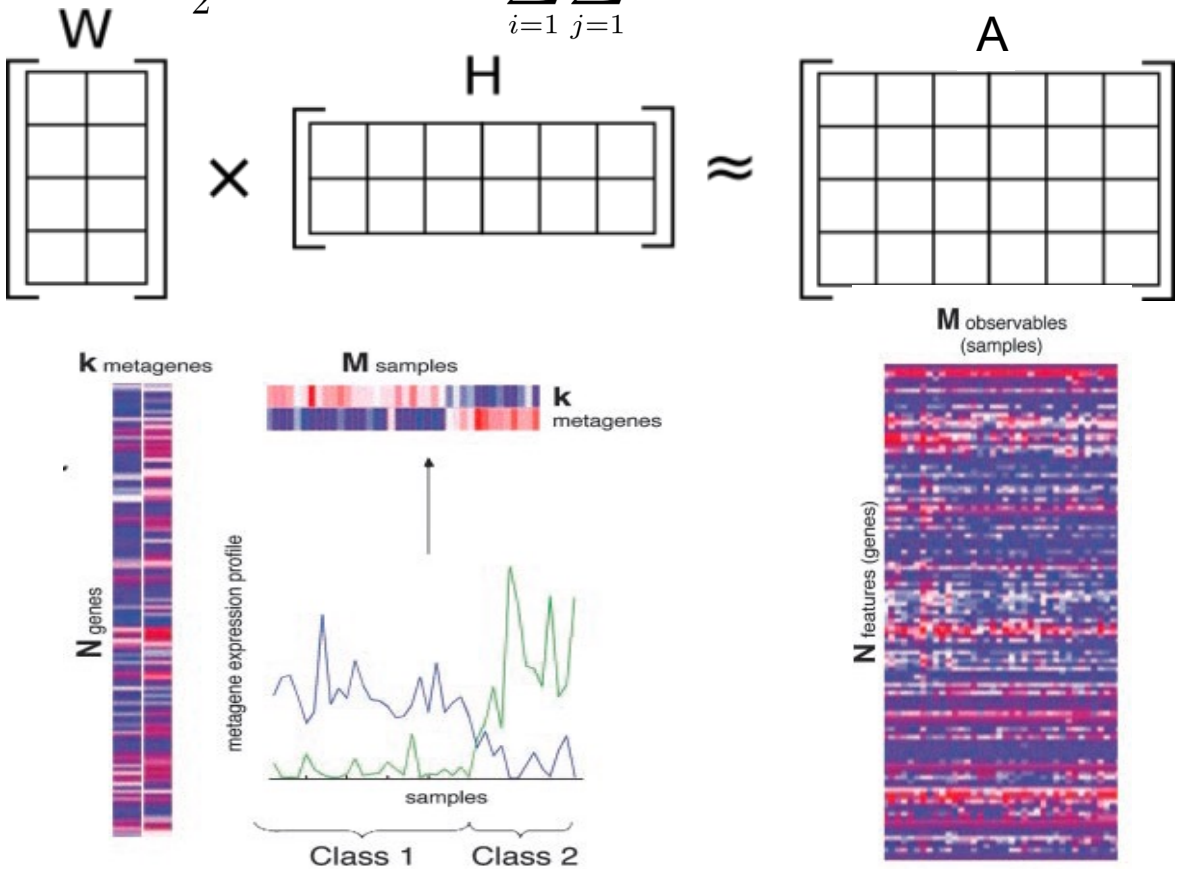
Supervised (n=~1000)
Cox elastic-net regression

Novel Molecular Methods Require Statisticians and Computational Biologists

- Chapuy et al (N = ~300)
  - **Unsupervised** approach utilizing Non-negative Matrix Factorization (NMF)
    - Approximate A ~ W H with a cost function: $\frac{1}{2}||\mathbf{A}-\mathbf{WH}||_{\mathrm{F}}^2 = \sum_{i=1}^{n}\sum_{j=1}^{m}(A_{ij}-(WH)_{ij})^2$
    - Define clusters using basis



Chapuy et al., Cell 2018



Genentech has run a few 1L DLBCL clinical trials…

Question: Can we reproduce the results from these methods and inform new clinical trial development?

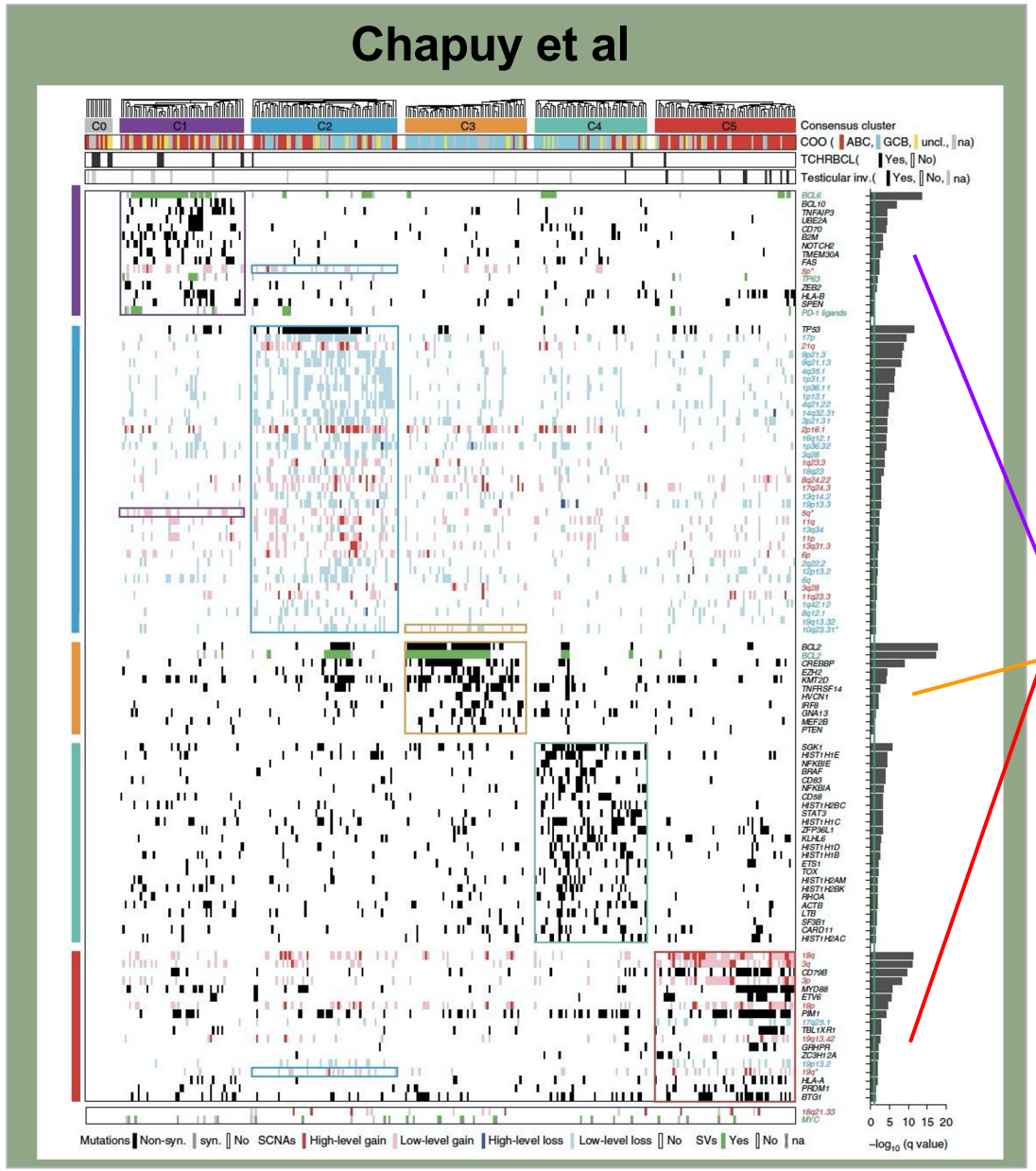
Trial	Phase	Indication	N
MAIN	III	1L DLBCL	787
GOYA	III	1L DLBCL	1418
CAVALLI	IB/II	1L DLBCL	211

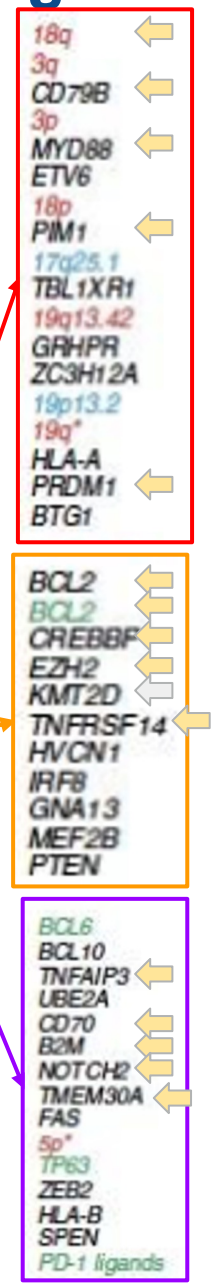
A few of the larger trials we have RNA-Seq or Genomic data on…

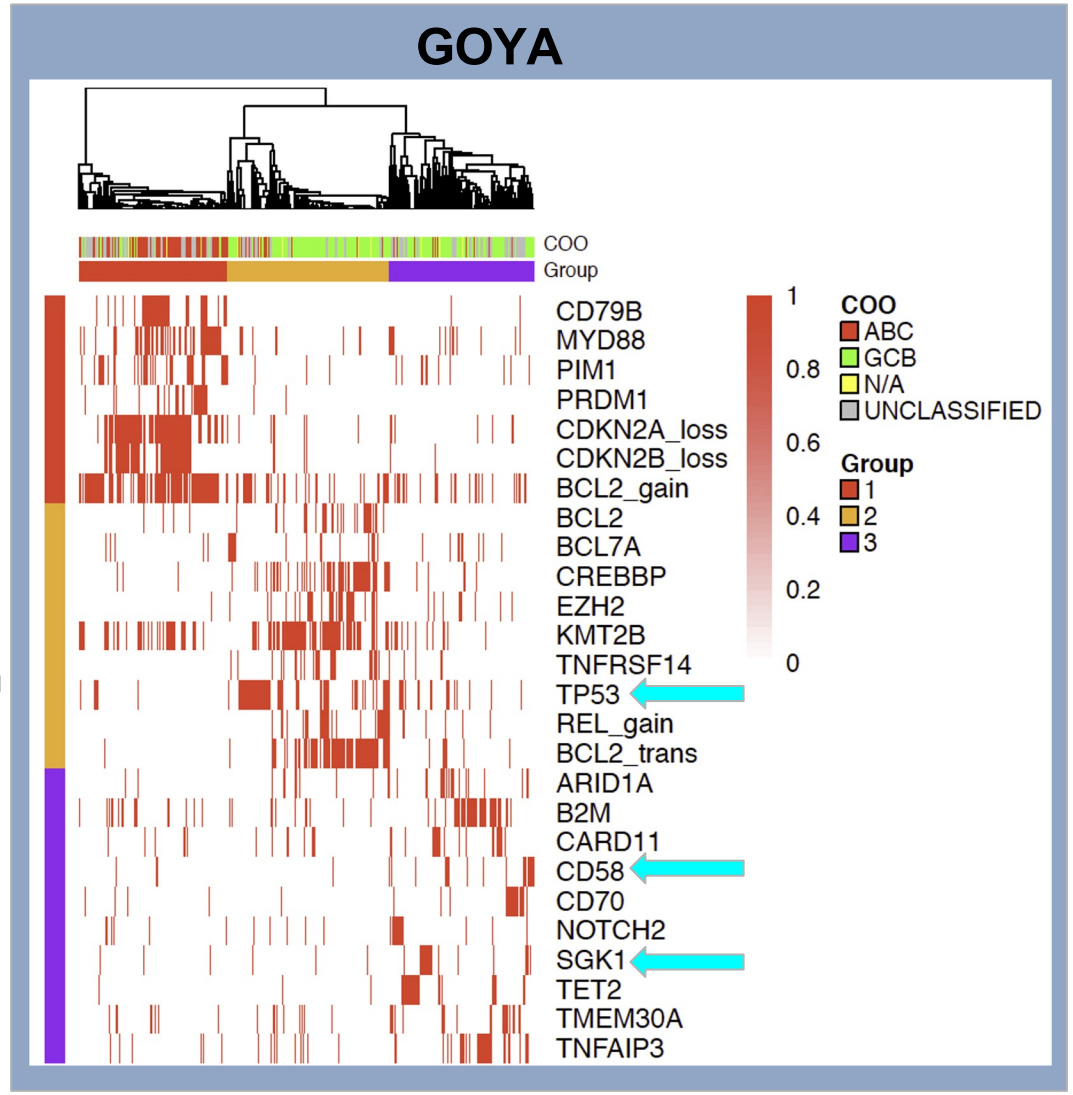
Whole Exome Sequencing vs. FOne Heme

- WES: Assays virtually every protein coding gene
  - Downstream analysis pipelines can capture genomic events other than small mutations (translocations, amplifications)
- FOne Heme: ~465 genes
  - Non-mutation origin genomic events are captured for the same genes, but vary in quality (BCL2 amplifications)

Recapitulation of NMF clusters by using FOne Heme in GOYA







Associated match in Goya

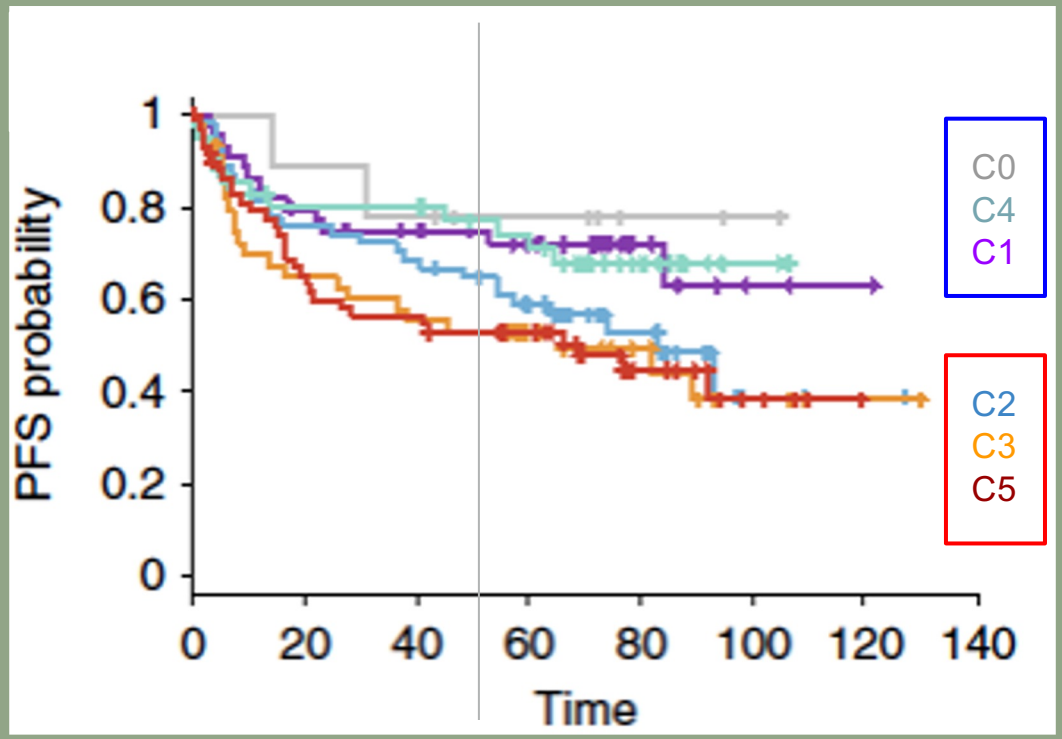
KMT2D (MLL4) in Chapuy, KMT2B (MLL2) in Goya

Notable members of Chapuy clusters 2 and 4 not found in Goya

Eugene Kim

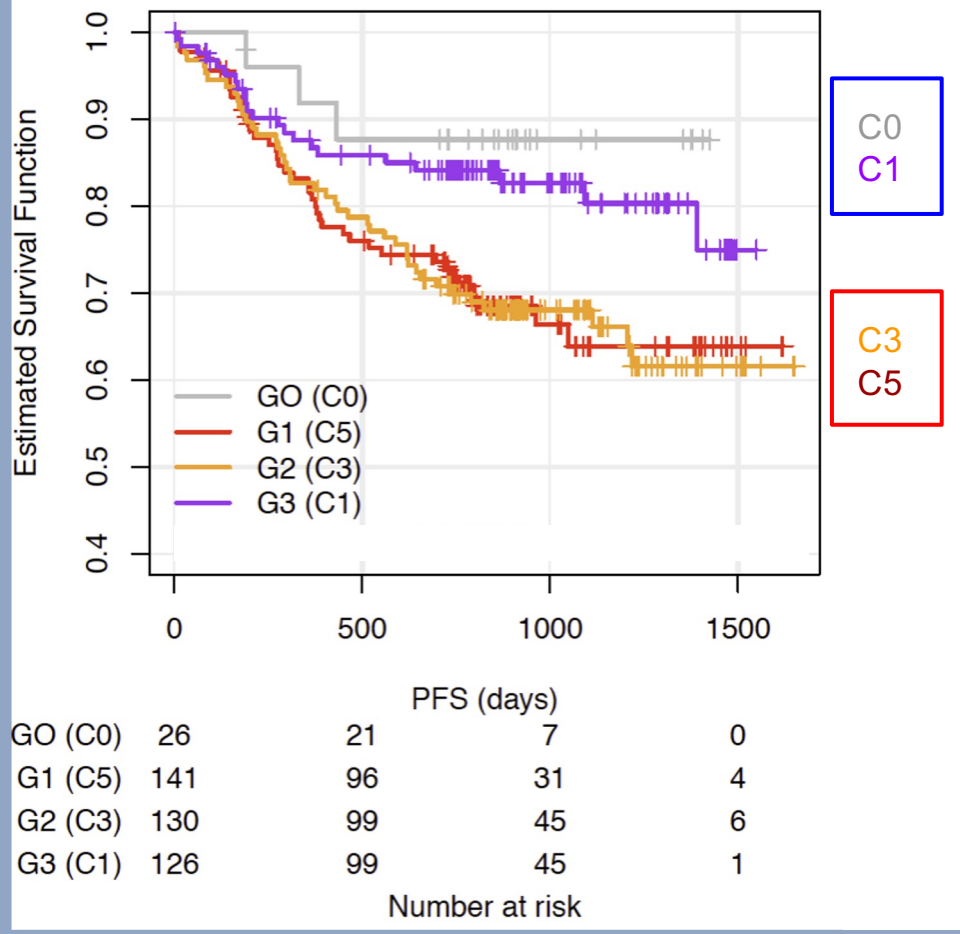
Recapitulation of Survival Profiles (PFS) by NMF Clusters in GOYA

Chapuy et al



Low-Risk Group
High-Risk Group

GOYA

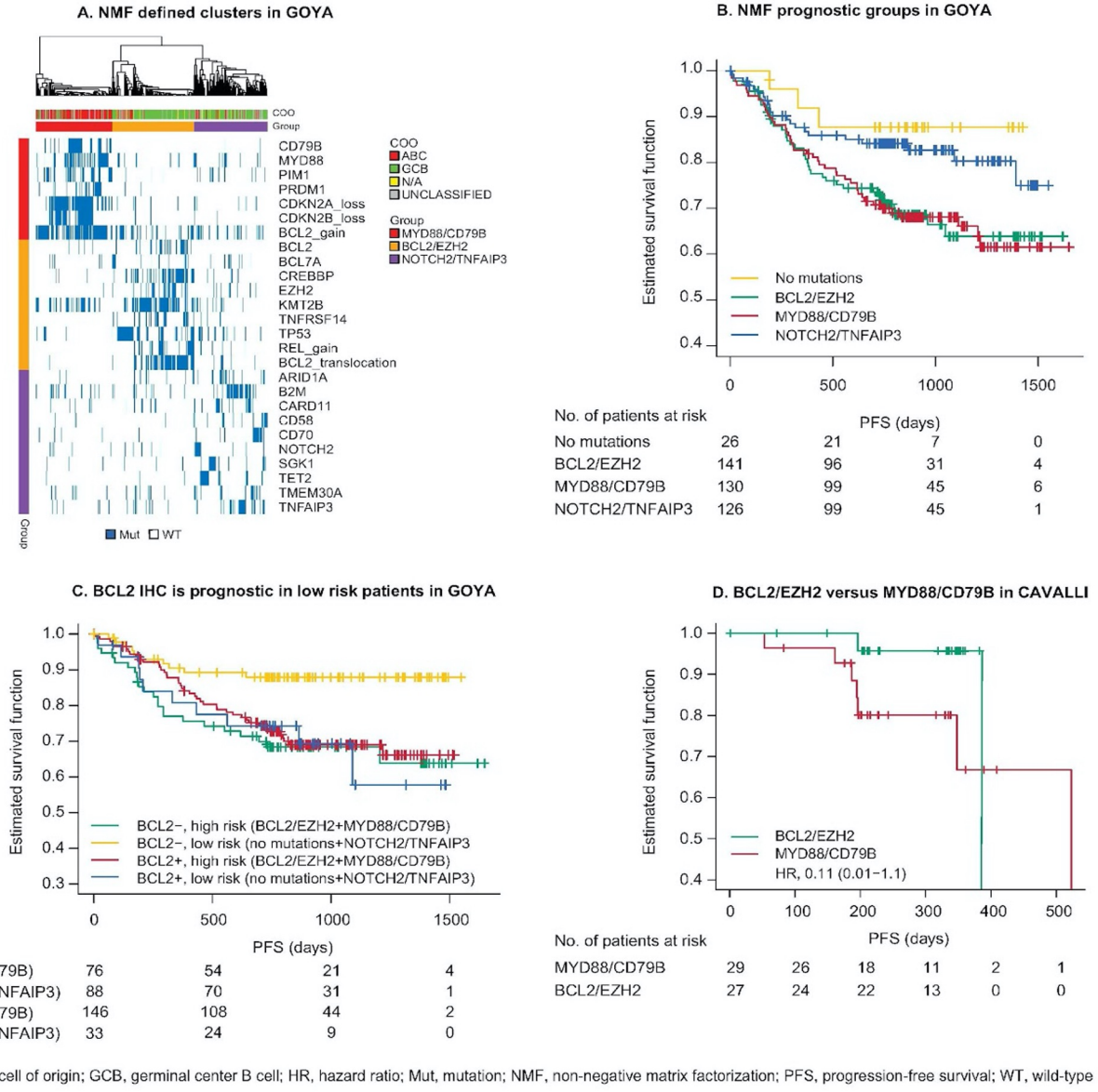


Eugene Kim

Recapitulation of molecularly defined subgroups

1. Utility of targeted mutational data from F1H in recapitulating molecular clusters, previously identified using WES, that further subclassify established prognostic groups such as COO and IPI, but not likely to be used in a trial today
2. Potential for utilizing novel genetic signatures as a means of identifying patients suitable for targeted therapies in the era of personalized health care
3. A trend towards improved survival outcomes for patients in the BCL2/EZH2 cluster upon treatment with the BCL2 inhibitor Venetoclax

Figure 1.
Comparison of prognostic groups in the GOYA and CAVALLI studies



Abbreviations: ABC, activated B cell; COO, cell of origin; GCB, germinal center B cell; HR, hazard ratio; Mut, mutation; NMF, non-negative matrix factorization; PFS, progression-free survival; WT, wild-type

Next steps…

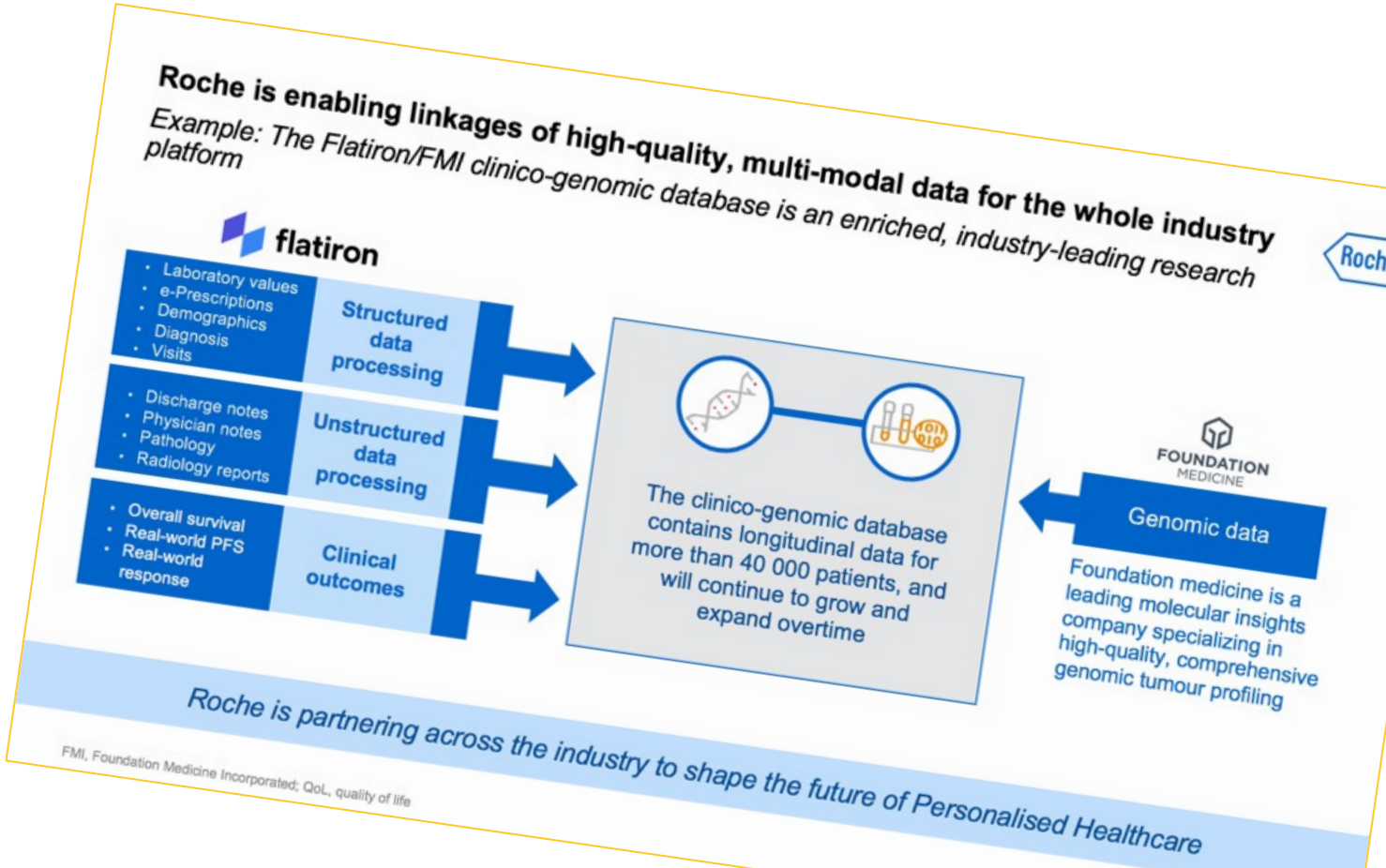
- Finalize evaluation of prognostic benefit for genomic clusters on top of established prognostic clinical and biomarker features

- Apply the NMF algorithm on the clinico-genomic database on 1L DLBCL patients for further validation…

Recapitulation of Prognostic Mutational Subtyping Utilizing F1H in Goya and Cavalli and the Potential to Highlight Benefit of Targeted Therapy in *De Novo* DLBCL

Eugene C Kim, PhD[1*], Jinzhen Fan, PhD[1*], Christopher R Bolen, PhD[1*], Alexandra Bazeos, MD, PhD, FRCPath[2*], Yanwen Jiang, PhD[1*], Sandhya Balasubramanian, MS[1*], Rama Balakrishnan, PhD[1*], Andrea Knapp, PhD[2*], Kathryn Humphrey, BSc[2*], Tina G Nielsen, MD, PhD[2*], Jeffrey M Venstrom, MD, PhD[1*], Genevive Hernandez, PhD[1*] and Joseph N Paulson, PhD[1*]

[1]Genentech, Inc., South San Francisco, CA; [2]F. Hoffmann-La Roche Ltd, Basel, Switzerland



Thankfully – new internal initiatives help with problems of old

Finding data and associated context (e.g. documentation)

Accessing data from different systems

Data sharing culture – “my” data, not “our” data

Inconsistent **storage** approach for managing data

Challenge to find **documentation**

Lack of **identifiers** for re-linking (patient/sample) and properly tracking data (dataset)

Along the way…
we had long term re-use in mind

Workstreams

A Generate Datamart

Create a NHL data mart of legacy clinical trials in a way that will be easily analyzable

Capacity building with historical studies and **genomic** data

B Identify prognostic subgroup(s)

Find subgroups of patients who have High Unmet Medical Need (UMN) with available therapies

C Inform pipeline + disease biology

Test the hypotheses brought forward by that will inform strategic or trial design decisions

D Patient reported outcomes

Evaluate patient reported outcomes from multiple studies

E Safety

Correlation of safety outcomes with baseline factors, patient characteristics, and efficacy.

Patient-centric approach with potential to incorporate **genomic** data

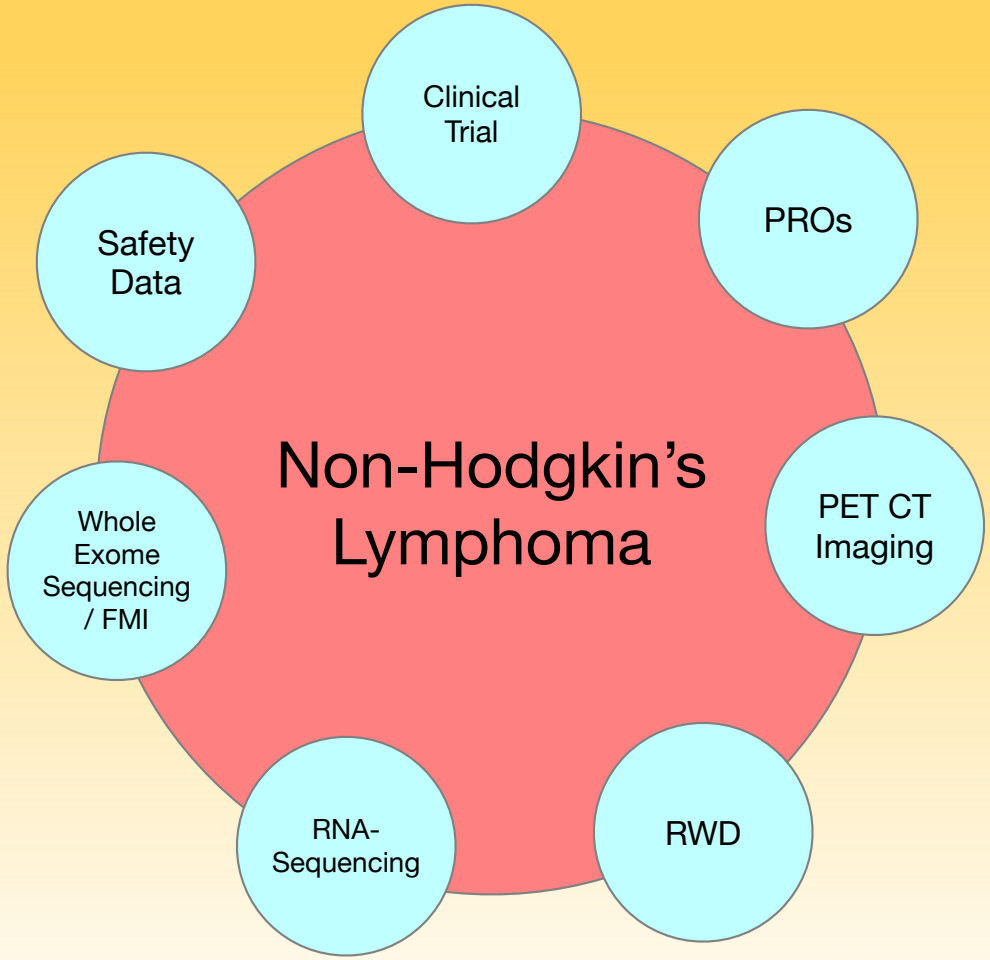
Identify actionable patient subset(s)

Organized a team of statisticians, (computational) biologists, clinicians, curators, integrators, etc…

Statisticians can be the stewards

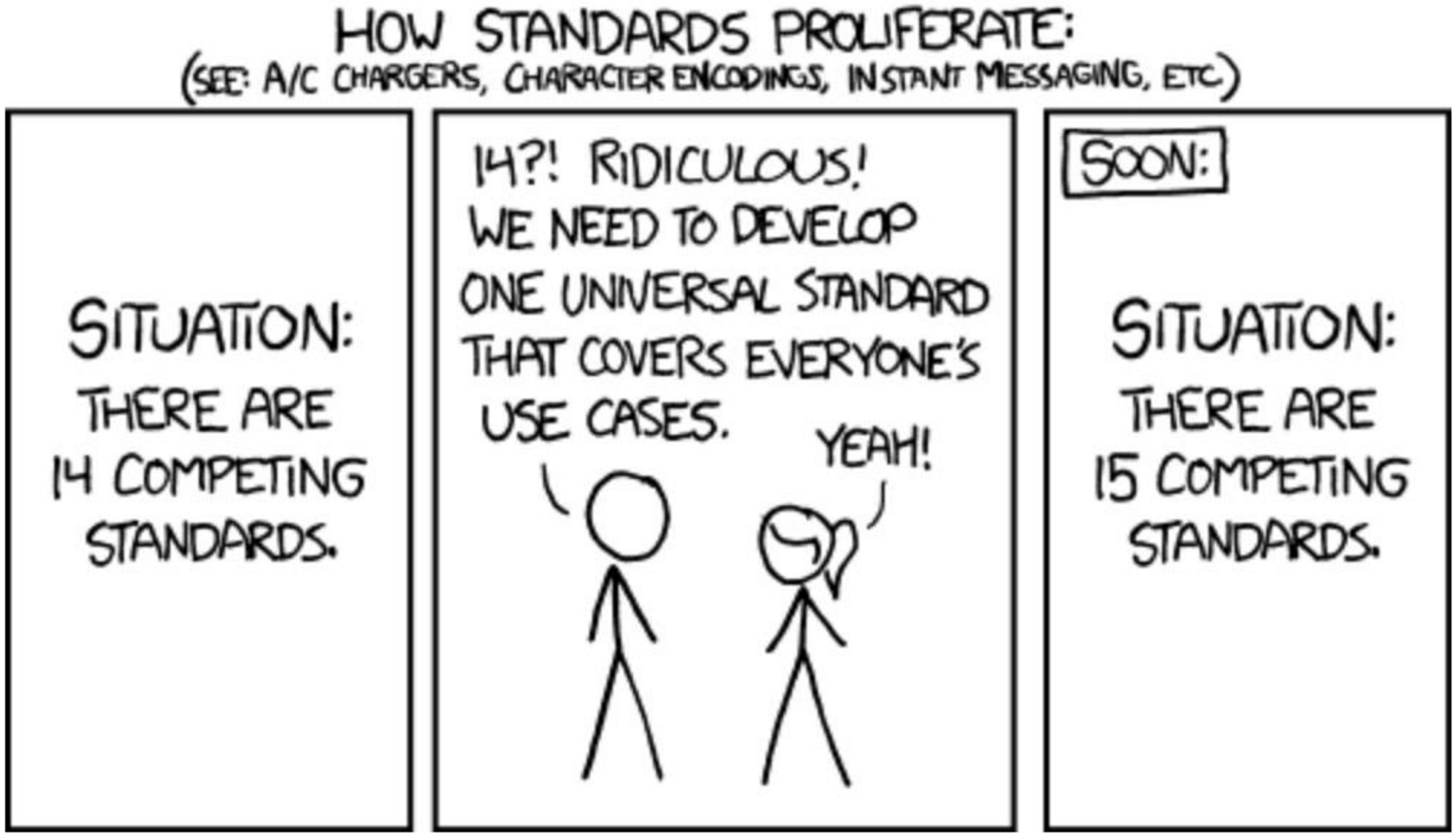
Data is an asset

> 6,000 patients
~1K RNA-Seq samples
~1K F1Heme samples
~2k Imaging PET CT



Trial	Phase	Indication	N
MAIN	III	1L DLBCL	787
GOYA	III	1L DLBCL	1418
GALLIUM	III	1L FL	1400
CAVALLI	IB/II	1L DLBCL	211
CONTRALTO	II	RR FL	163
PRIMA	III	1L FL	1217
SABRINA	III	1L FL	410
GADOLIN	III	RR FL	396

"The best thing about standards is that there are so many to choose from"



xkcd

Acknowledgements

Jane Fridlyand
Tina Nielsen
Eugene Kim
Jemma Fan
Nelson Kinnersley

Doing now what patients need next