OMICS Data Integration: Why and How?

Ali Shojaie
Department of Biostatistics
University of Washington
faculty.washington.edu/ashojaie

JSM 2019
Statistics in Genomics and Genetics Roundtable

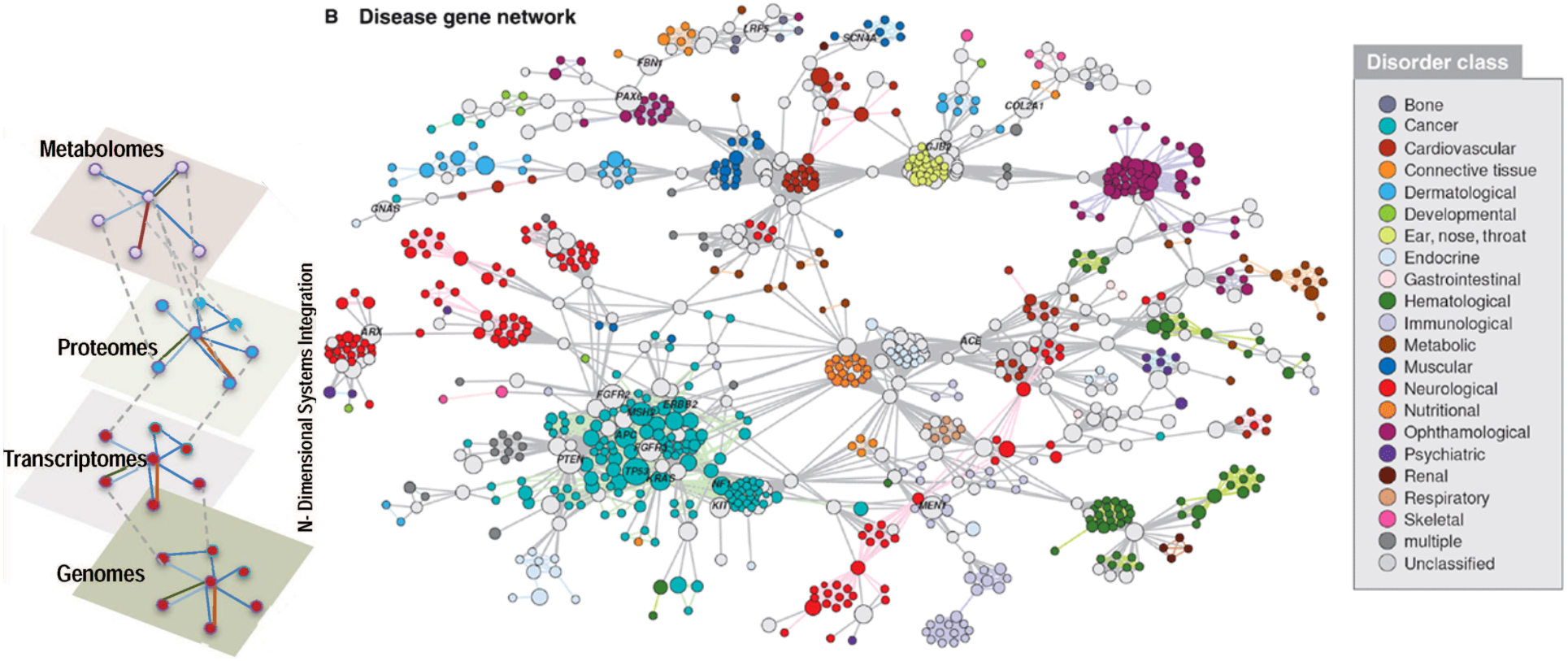
Part I: WHY?

Why Integrate Omics Data?

Biology is complex, heterogenous and structured!

Why Integrate Omics Data?

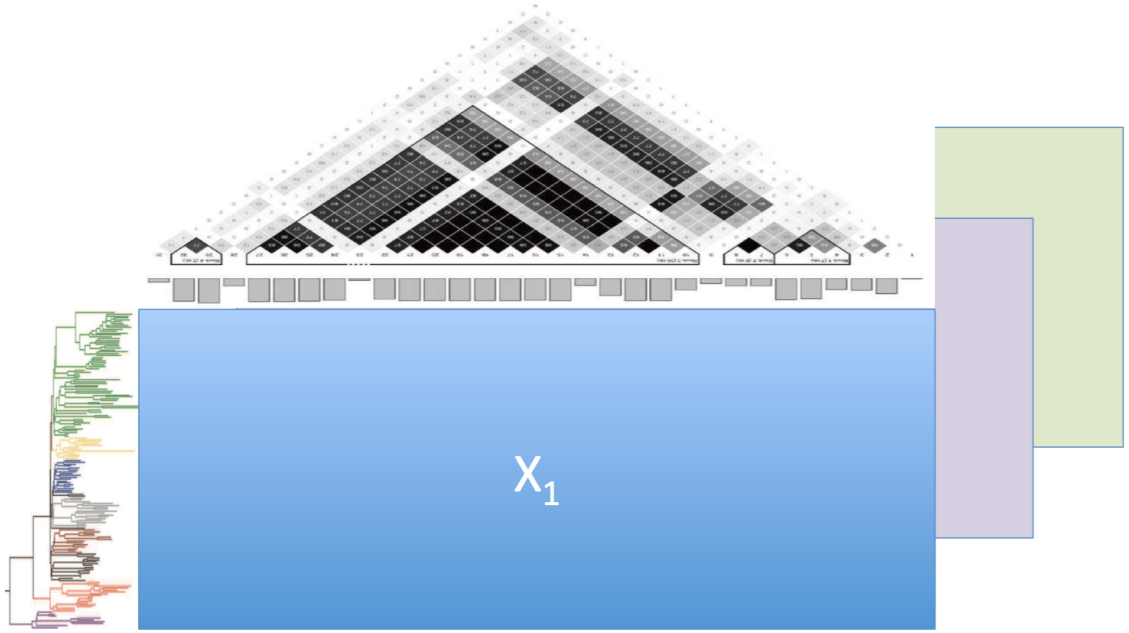
Biology is complex, heterogenous and structured!



Why Integrate Omics Data?

Comprehensive view of biology requires looking at multiple types of *omics* data (TCGA, ENCODE, etc)

⇒ integrative analysis of multiple structured omics data



Why Integrate Omics Data?

Possible reasons:

Why Integrate Omics Data?

Possible reasons:

1. To confirm or narrow down omics signals

Why Integrate Omics Data?

Possible reasons:

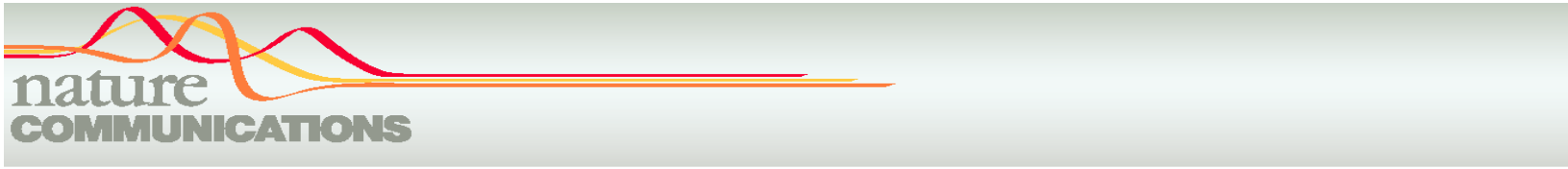
1. To confirm or narrow down omics signals
2. To complement or boost omics signals

Why Integrate Omics Data?

Possible reasons:

1. To confirm or narrow down omics signals
2. To complement or boost omics signals
3. To glean systems perspective

Omics data integration can lead to new discoveries...



ARTICLE

Received 19 Feb 2013 | Accepted 16 Sep 2013 | Published 18 Oct 2013

DOI: 10.1038/ncomms3617

Reconstructing targetable pathways in lung cancer by integrating diverse omics data

O. Alejandro Balbin[1,2,3], John R. Prensner[1,2], Anirban Sahu[1,2], Anastasia Yocum[1,2], Sunita Shankar[1,2], Rohit Malik[1,2], Damian Fermin[2], Saravana M. Dhanasekaran[1,2], Benjamin Chandler[1], Dafydd Thomas[2], David G. Beer[2], Xuhong Cao[1,2], Alexey I. Nesvizhskii[1,2,3] & Arul M. Chinnaiyan[1,2,3]

Global 'multi-omics' profiling of cancer cells harbours the potential for characterizing the signalling networks associated with specific oncogenes. Here we profile the transcriptome, proteome and phosphoproteome in a panel of non-small cell lung cancer (NSCLC) cell lines in order to reconstruct targetable networks associated with KRAS dependency. We develop a two-step bioinformatics strategy addressing the challenge of integrating these disparate data sets. We first define an 'abundance-score' combining transcript, protein and phospho-protein abundances to nominate differentially abundant proteins and then use the Prize Collecting Steiner Tree algorithm to identify functional sub-networks. We identify three modules centred on KRAS and MET, LCK and PAK1 and β-Catenin. We validate activation of these proteins in KRAS-dependent (KRAS-Dep) cells and perform functional studies defining *LCK* as a critical gene for cell proliferation in KRAS-Dep but not KRAS-independent NSCLCs. These results suggest that LCK is a potential druggable target protein in KRAS-Dep lung cancers.

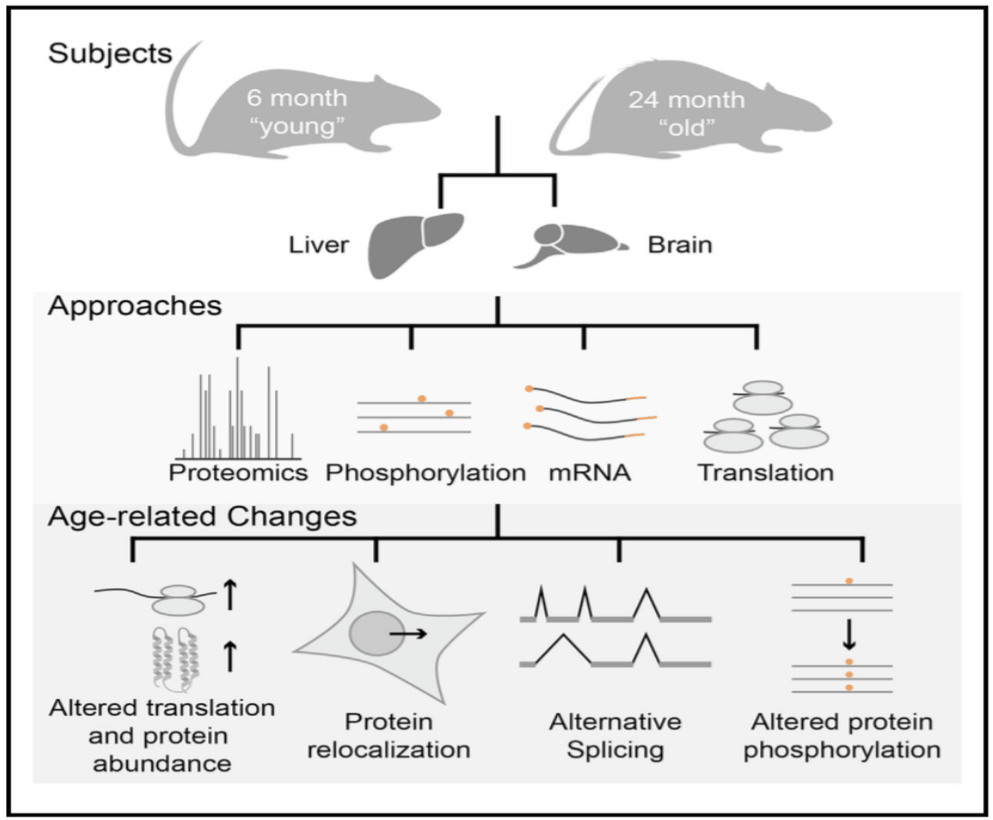
Omics data integration can lead to new discoveries...

Article

Cell Systems

Integrated Transcriptome and Proteome Analyses Reveal Organ-Specific Proteome Deterioration in Old Rats

Graphical Abstract



Highlights

Authors

Alessandro Ori, Brandon H. Toyama, Michael S. Harris, ..., Nicholas T. Ingolia, Martin W. Hetzer, Martin Beck

Correspondence

ingolia@berkeley.edu (N.T.I.), hetzer@salk.edu (M.W.H.), mbeck@embl.de (M.B.)

In Brief

Ori et al. quantified the molecular alterations that occur between young and old rats in two organs: brain and liver. By integrating genomic and proteomic measurements, the authors were able to reveal that changes in translation are the primary cause of protein level alterations during aging. However, they also identified other levels of regulation such as protein localization and phosphorylation that co-participate in modifying the proteome in old animals.

Omics data integration can lead to new discoveries...

SCIENTIFIC REPORTS

OPEN

Received: 30 June 2015
Accepted: 16 November 2015
Published: 11 December 2015

Integrated analysis of global proteome, phosphoproteome, and glycoproteome enables complementary interpretation of disease-related protein networks

Jong-Moon Park[1,*], Ji-Hwan Park[2,*], Dong-Gi Mun[3,*], Jingi Bae[3,*], Jae Hun Jung[4], Seunghoon Back[3], Hangyeore Lee[3], Hokeun Kim[3], Hee-Jung Jung[6], Hark Kyun Kim[5], Hookeun Lee[1], Kwang Pyo Kim[4], Daehee Hwang[2,6] & Sang-Won Lee[3]

Multi-dimensional proteomic analyses provide different layers of protein information, including protein abundance and post-translational modifications. Here, we report an integrated analysis of protein expression, phosphorylation, and N-glycosylation by serial enrichments of phosphorylation and N-glycosylation (SEPG) from the same tissue samples. On average, the SEPG identified 142,106 unmodified peptides of 8,625 protein groups, 18,846 phosphopeptides (15,647 phosphosites), and 4,019 N-glycopeptides (2,634 N-glycosites) in tumor and adjacent normal tissues from three gastric cancer patients. The combined analysis of these data showed that the integrated analysis additively improved the coverages of gastric cancer-related protein networks; phosphoproteome and N-glycoproteome captured predominantly low abundant signal proteins, and membranous or secreted proteins, respectively, while global proteome provided abundances for general population of the proteome. Therefore, our results demonstrate that the SEPG can serve as an effective approach for multi-dimensional proteome analyses, and the holistic profiles of protein expression and PTMs enabled improved interpretation of disease-related networks by providing complementary information.

Omics data integration can lead to new discoveries...

Cancer Cell
Article

Integrated Analyses Identify a Master MicroRNA Regulatory Network for the Mesenchymal Subtype in Serous Ovarian Cancer

Da Yang,[1,11] Yan Sun,[1,7,11] Limei Hu,[1,11] Hong Zheng,[8,11] Ping Ji,[1] Chad V. Pecot,[6] Yanrui Zhao,[8] Sheila Reynolds,[9] Hanyin Cheng,[1,12] Rajesha Rupaimoole,[2] David Cogdell,[1] Matti Nykter,[10] Russell Broaddus,[1] Cristian Rodriguez-Aguayo,[4] Gabriel Lopez-Berestein,[4,5] Jinsong Liu,[1] Ilya Shmulevich,[9] Anil K. Sood,[2,3,5,*] Kexin Chen,[8,*] and Wei Zhang[1,5,*]
[1]Department of Pathology
[2]Department of Gynecologic Oncology and Reproductive Medicine
[3]Department of Cancer Biology
[4]Department of Experimental Therapeutics
[5]Center for RNAi and Non-Coding RNA
[6]Division of Cancer Medicine
The University of Texas MD Anderson Cancer Center, Houston, TX 77030, USA
[7]Department of Pathology
[8]Department of Epidemiology and Biostatistics
Tianjin Medical University Cancer Institute and Hospital, Tianjin 300060, China
[9]Institute for Systems Biology, Seattle, WA 98103, USA
[10]Tampere University of Technology, Tampere 33101, Finland
[11]These authors contributed equally to this work
[12]Present address: Department of Cancer Biology, Kimmel Cancer Center, Thomas Jefferson University, Philadelphia, PA
*Correspondence: asood@mdanderson.org (A.K.S.), chenkexin@tijmu.edu.cn (K.C.), wzhang@mdanderson.org (W.Z.)
http://dx.doi.org/10.1016/j.ccr.2012.12.020

SUMMARY

Integrated genomic analyses revealed a miRNA-regulatory network that further defined a robust integrated mesenchymal subtype associated with poor overall survival in 459 cases of serous ovarian cancer (OvCa) from The Cancer Genome Atlas and 560 cases from independent cohorts. Eight key miRNAs, including miR-506, miR-141, and miR-200a, were predicted to regulate 89% of the targets in this network. Follow-up functional experiments illustrate that miR-506 augmented E-cadherin expression, inhibited cell migration and invasion, and prevented TGFβ-induced epithelial-mesenchymal transition by targeting *SNAI2*, a transcriptional repressor of E-cadherin. In human OvCa, miR-506 expression was correlated with decreased SNAI2 and VIM, elevated E-cadherin, and beneficial prognosis. Nanoparticle delivery of miR-506 in orthotopic OvCa mouse models led to E-cadherin induction and reduced tumor growth.

Part II: HOW?

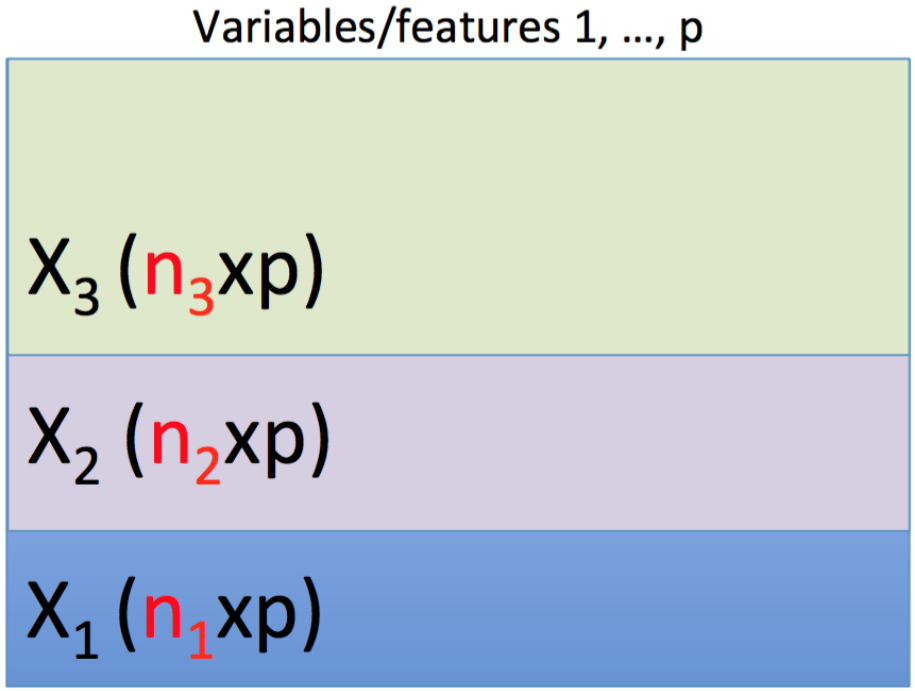
More on Omics Data Integration

Broadly, two existing integration approaches:

More on Omics Data Integration

Broadly, two existing integration approaches:

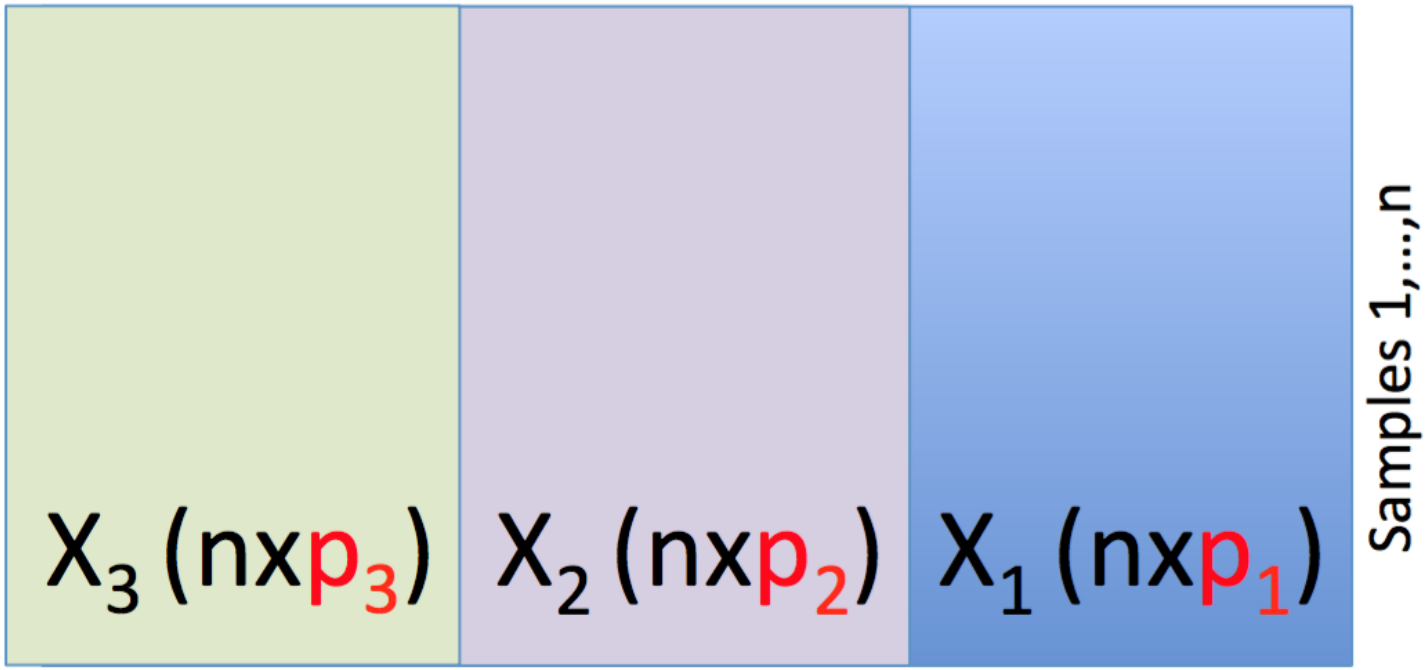
- Horizontal integration (same variables, different studies/subjects)



More on Omics Data Integration

Broadly, two existing integration approaches:

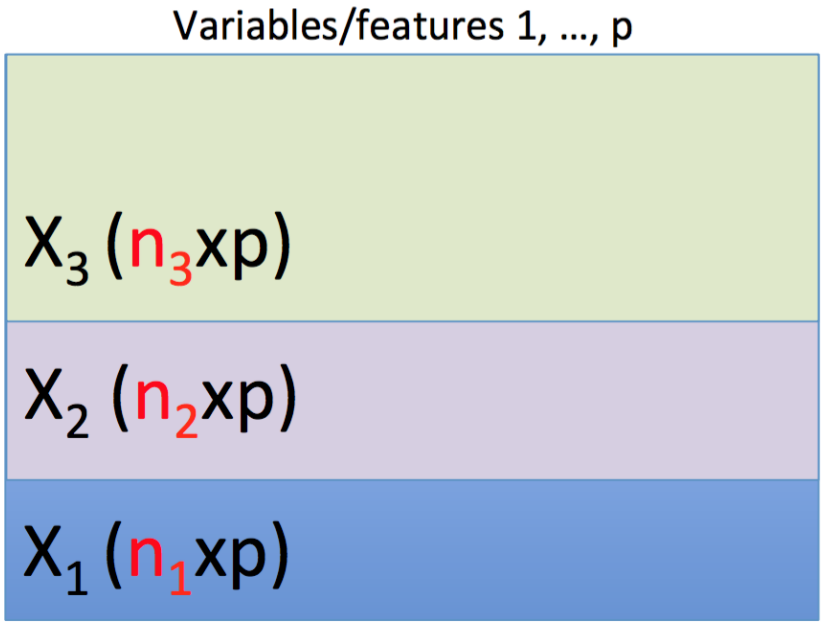
- Vertical integration (different platforms/variables, same subjects)



Existing Approaches: Meta Analysis

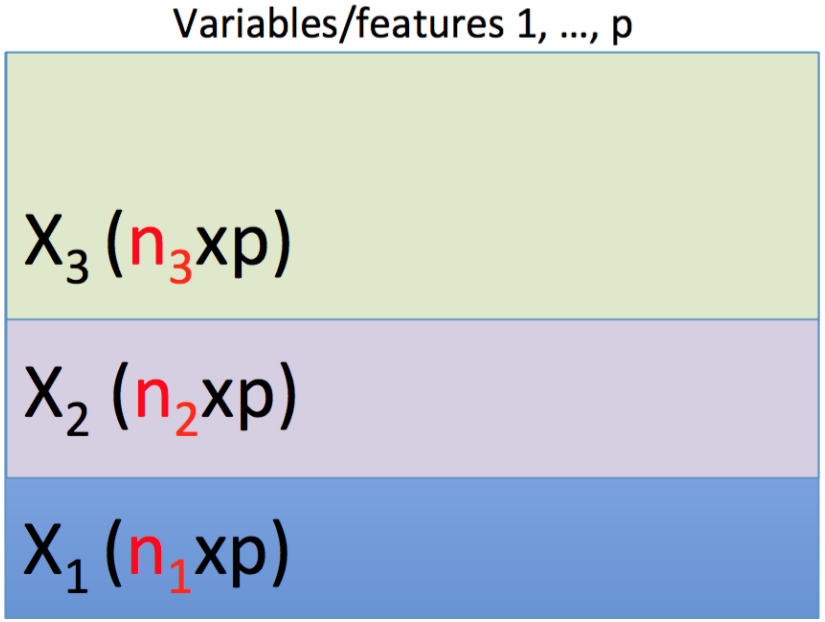
Existing Approaches: Meta Analysis

Horizontal integration



Existing Approaches: Meta Analysis

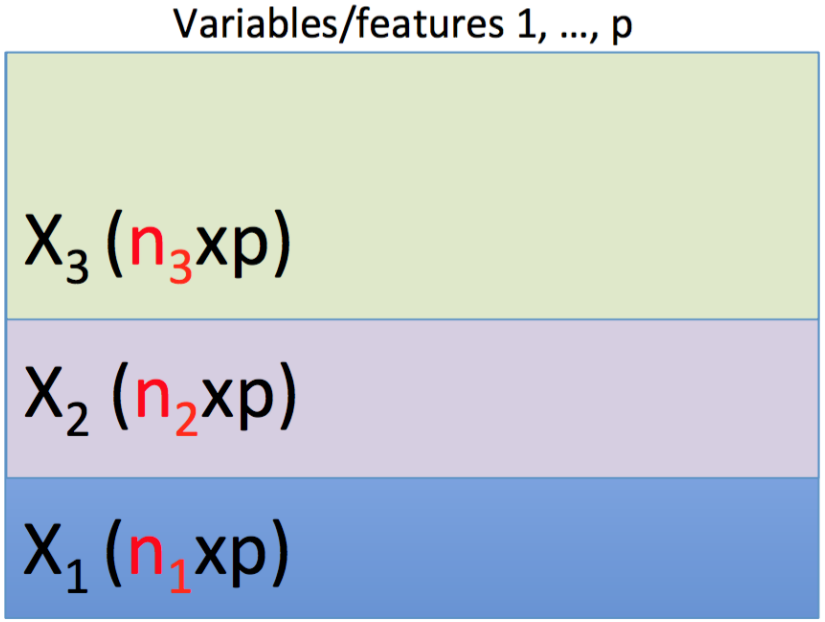
Horizontal integration



- Used extensively in GWAS (especially in consortiums)

Existing Approaches: Meta Analysis

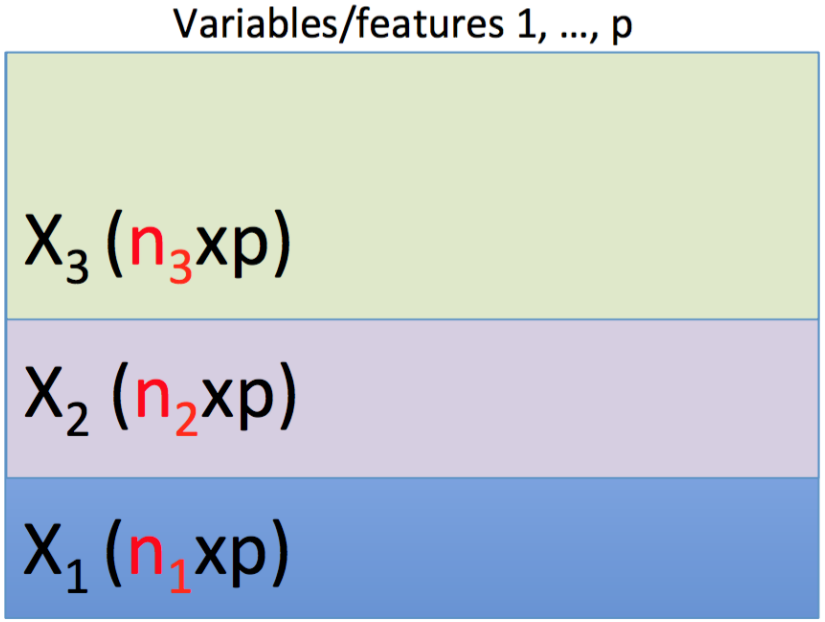
Horizontal integration



- Used extensively in GWAS (especially in consortiums)
- Used to boost the signal (larger sample size)

Existing Approaches: Meta Analysis

Horizontal integration



- Used extensively in GWAS (especially in consortiums)
- Used to boost the signal (larger sample size)
- Used to confirm previous findings (reproducibility)

Existing Approaches: Meta Analysis

Table 3 | **Summary of methods for meta-analysis of genome-wide data**

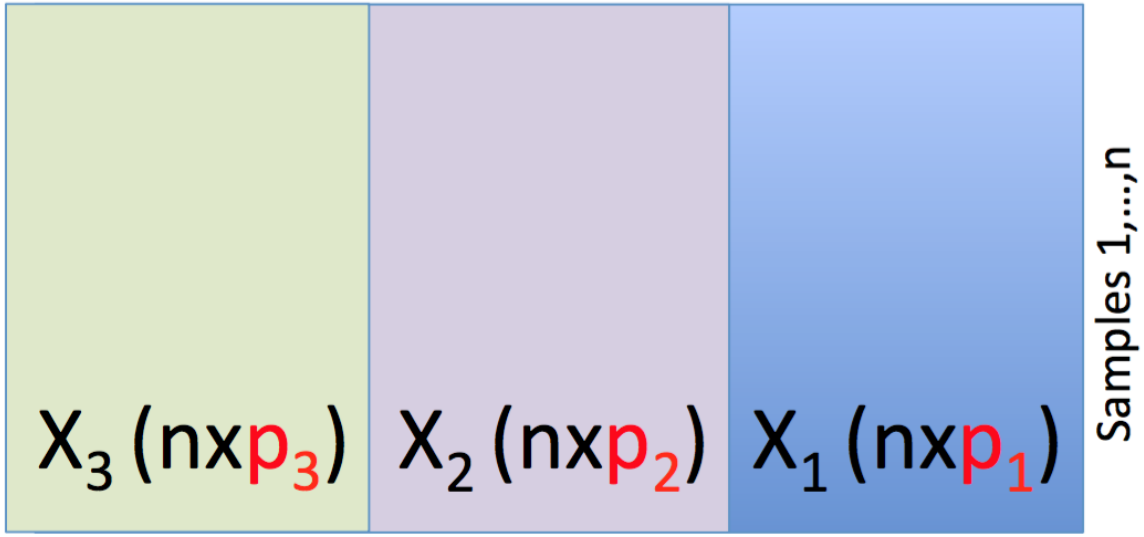
Method	Description	Advantages	Disadvantages	Main software used
P value meta-analysis	Simplest meta-analytical approach	Allows meta-analysis when effects are not available	Direction of effect is not always available; inability to provide effect sizes; difficulties in interpretation	METAL, GWAMA, R packages
Fixed effects	Synthesis of effect sizes. Between-study variance is assumed to be zero	Effects readily available through specialized software	Results may be biased if a large amount of heterogeneity exists	METAL, GWAMA, R packages
Random effects	Synthesis of effect sizes. Assumes that the individual studies estimate different effects	Generalizability of results	Power deserts in discovery efforts; may yield spuriously large summary effect estimates when there are selection biases	GWAMA, R packages
Bayesian approach	Incorporates prior assessment of the genetic effects	Most direct method for interpretation of results as posterior probabilities given the observed data	Methodologically challenging; GWAS-tailored routine software not available; subjective prior information used	R packages
Multivariate approaches	Incorporates the possible correlation between outcomes or genetic variants	Increased power can identify variants that conventional meta-analysis do not reveal using the same data sets	Computationally intensive; software not available for all analyses; some may require individual-level data	GCTA for multi-locus approaches
Other extensions	A set of different approaches that allows for the identification of multiple variants across different diseases	Summary results of previous meta-analyses can be used	May need additional exploratory analyses for the identification of variants; prone to systematic biases	Software developed by the authors of the proposed methodologies

Evangelou & Ioannidis (2012) Nat. Rev. Gen.

Existing Approaches: Direct Integration

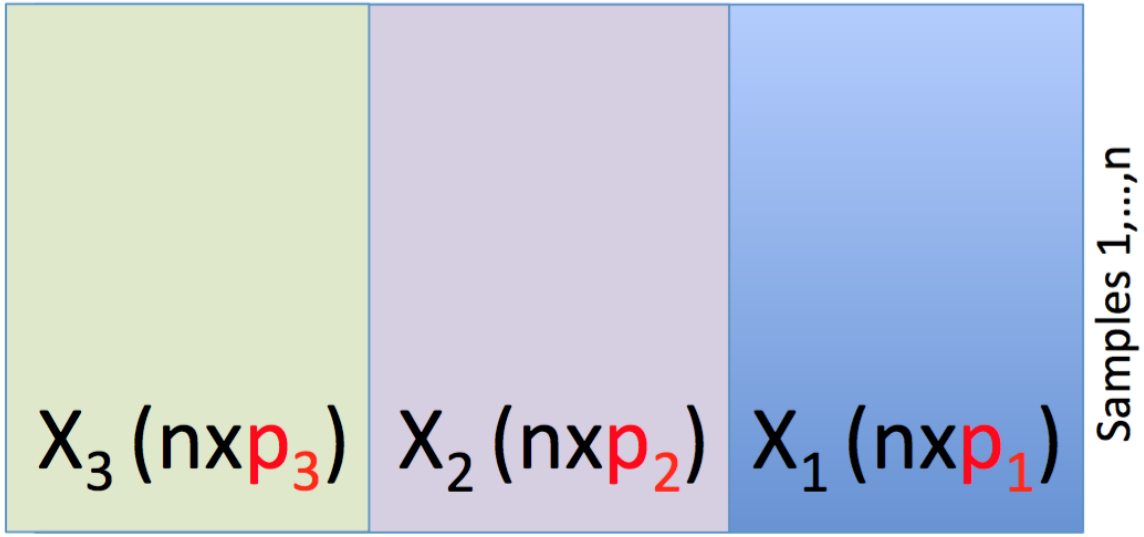
Existing Approaches: Direct Integration

Vertical integration



Existing Approaches: Direct Integration

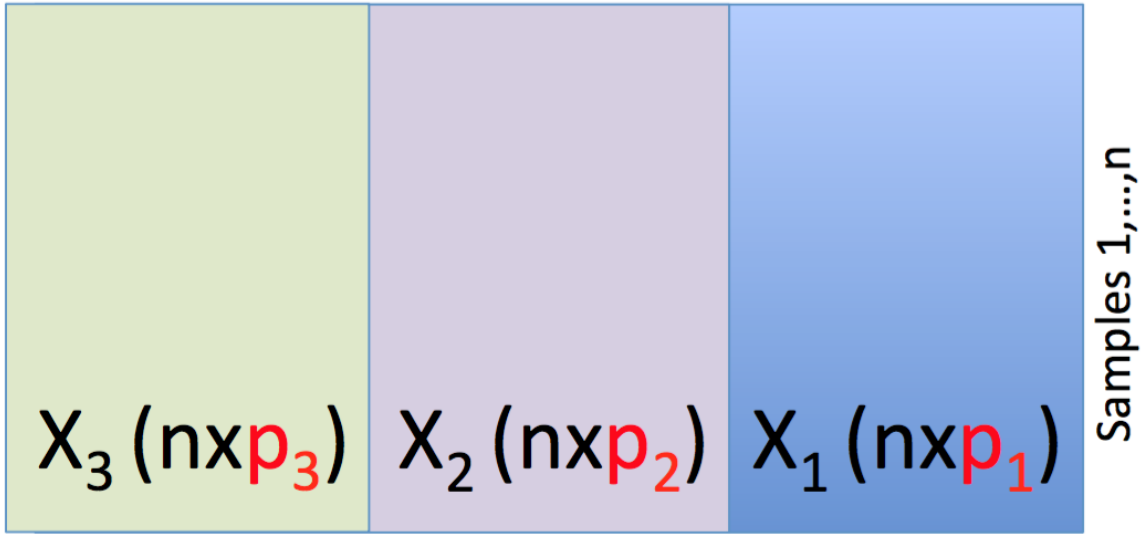
Vertical integration



The simplest approach: concatenate the variables!!

Existing Approaches: Direct Integration

Vertical integration



The simplest approach: concatenate the variables!!

- Can result in way-too-many variables
- Can discern conditional associations with phenotype *y*

Existing Approaches: Kernel-Based Methods

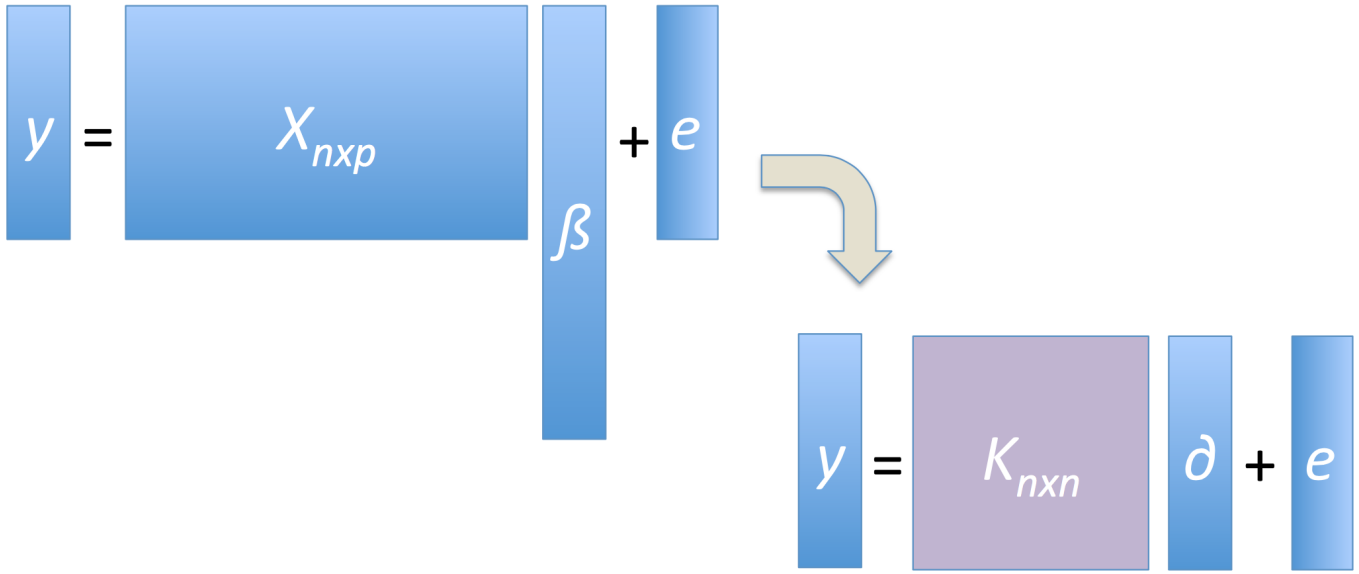
Existing Approaches: Kernel-Based Methods

Vertical integration using kernel regression

Existing Approaches: Kernel-Based Methods

Vertical integration using kernel regression

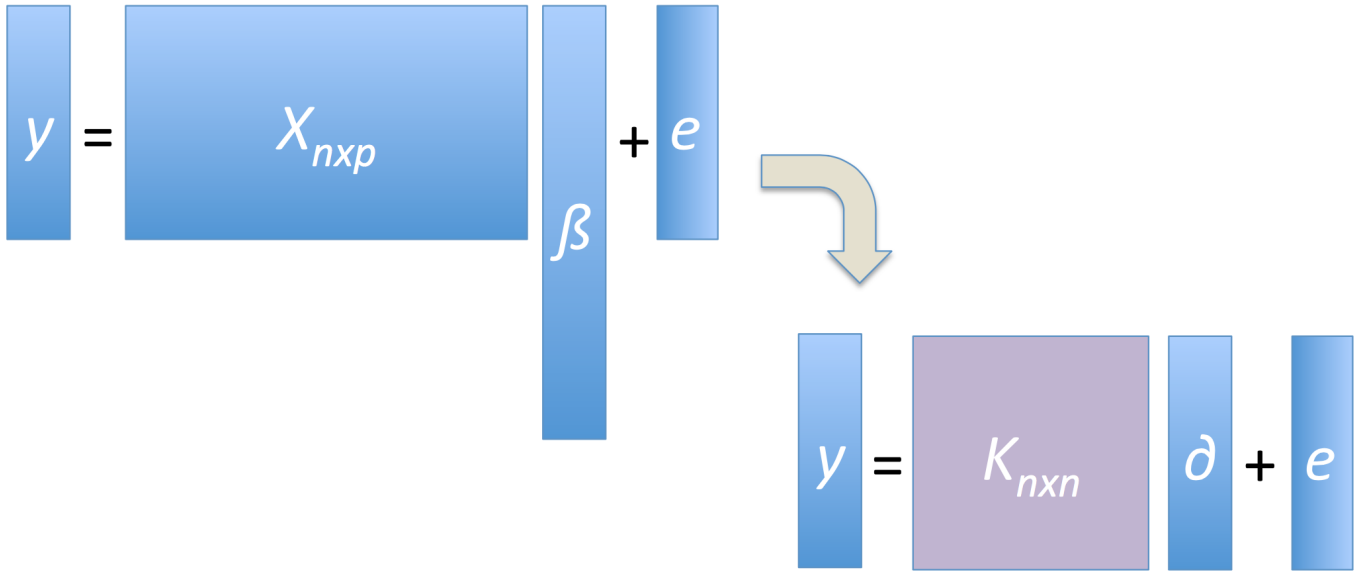
Kernel Regression



Existing Approaches: Kernel-Based Methods

Vertical integration using kernel regression

Kernel Regression



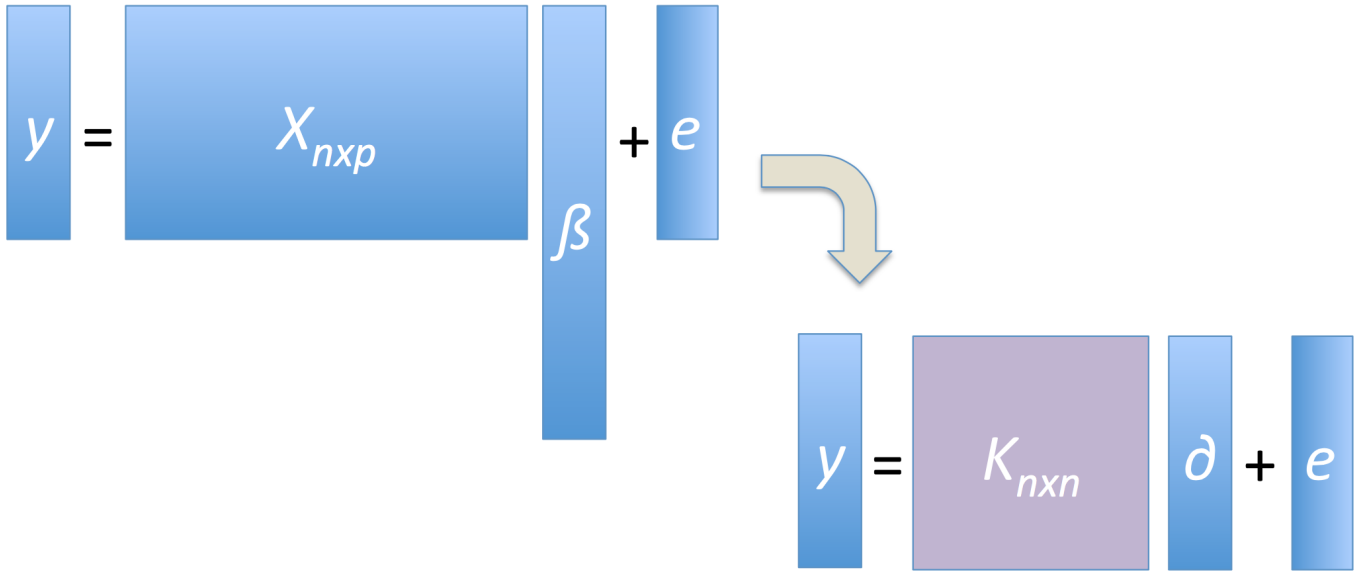
- Penalized regression in terms of a kernel K

$$\hat{\delta} = \underset{\delta \in \mathbb{R}^n}{\operatorname{argmin}} \, \| y - K\delta \|_2^2 + \lambda \| \delta \|_K^2$$

Existing Approaches: Kernel-Based Methods

Vertical integration using kernel regression

Kernel Regression



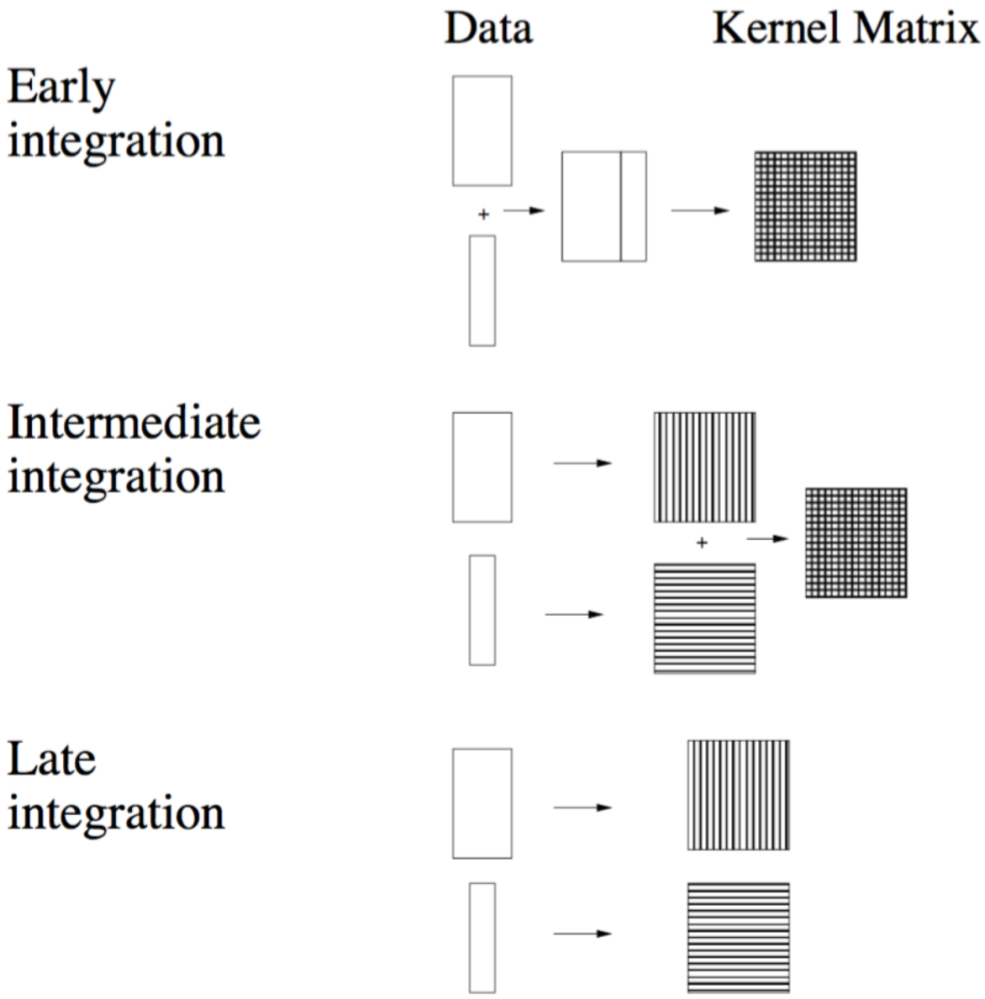
- Penalized regression in terms of a kernel K

$$\hat{\delta} = \underset{\delta \in \mathbb{R}^n}{\operatorname{argmin}} \, \| y - K\delta \|_2^2 + \lambda \| \delta \|_K^2$$

- Ideal for predicting y in the dual space
- Also used to test for association between y and X (SKAT)

Existing Approaches: Kernel-Based Methods

Can define different kernels (or feature maps) for different omics data types

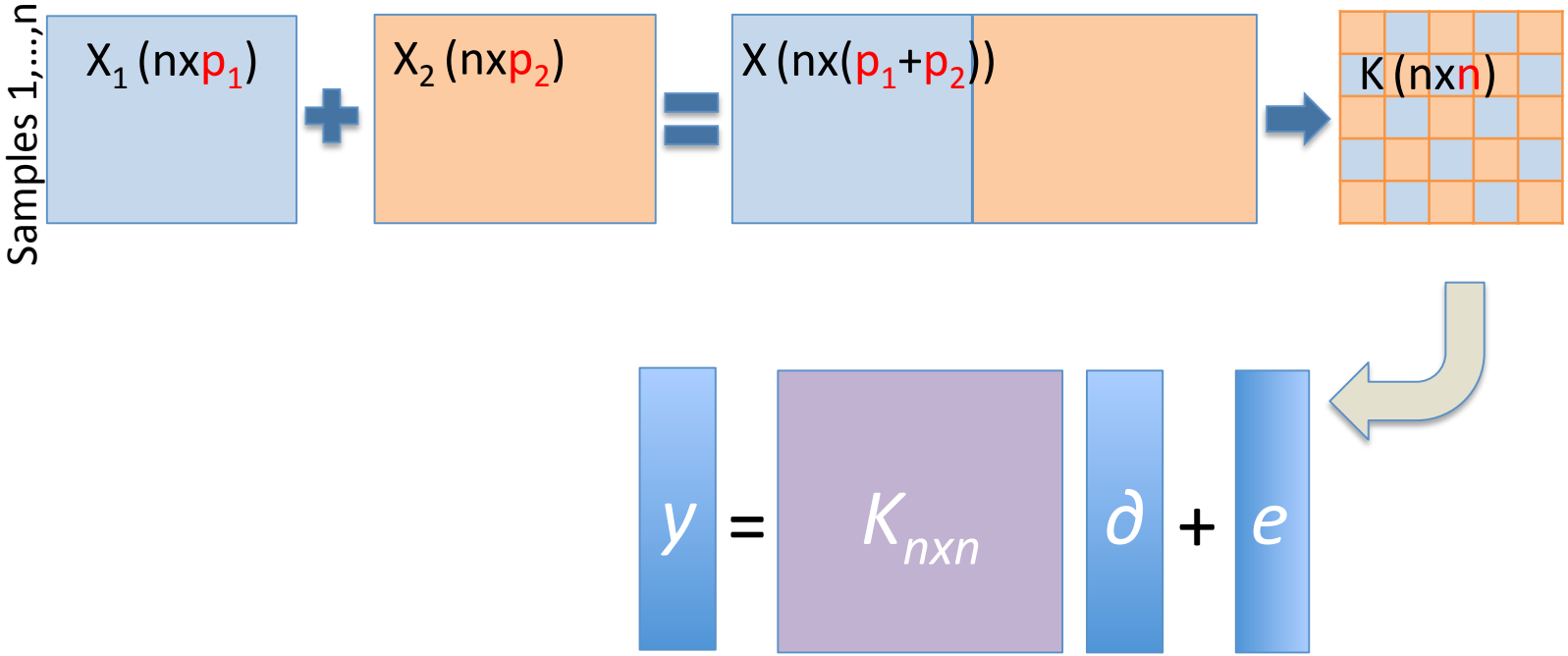


Pavlidis et al (2001, 2002)

Existing Approaches: Kernel-Based Methods

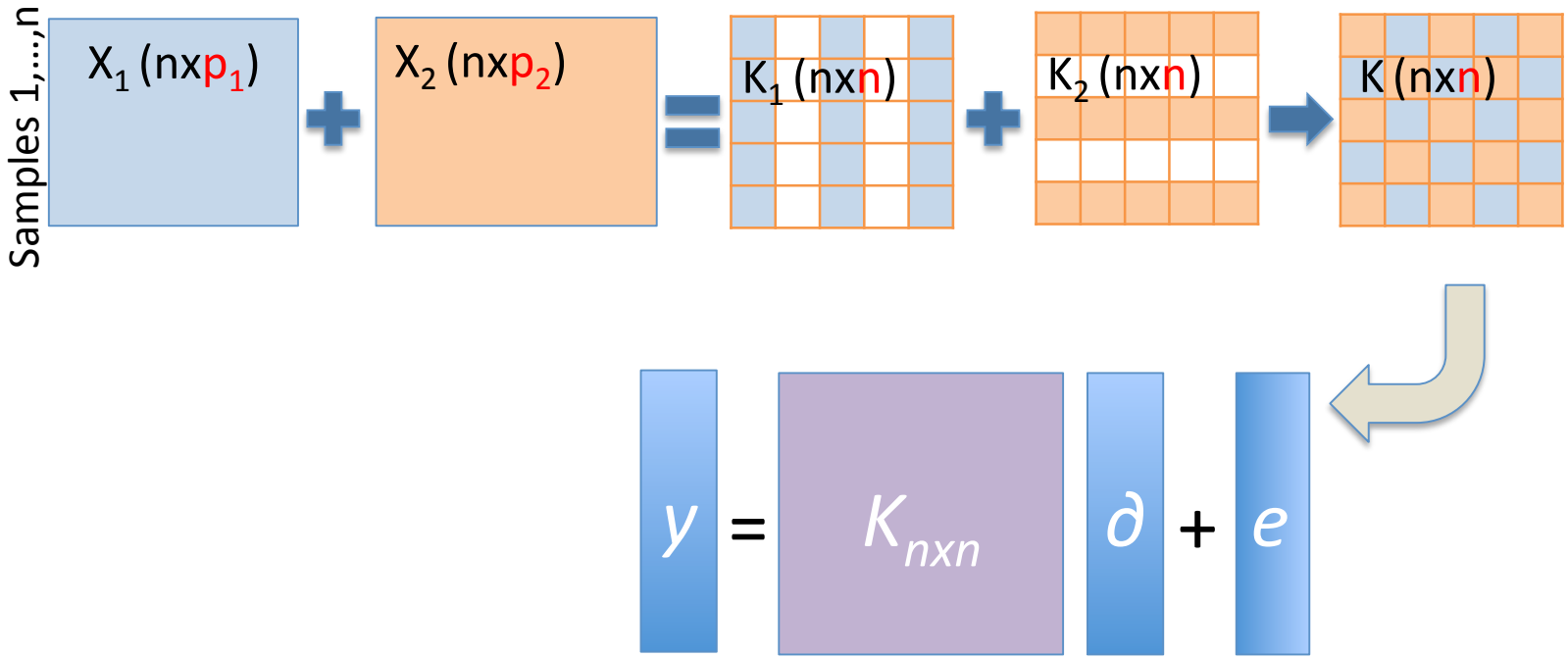
Existing Approaches: Kernel-Based Methods

- Early integration



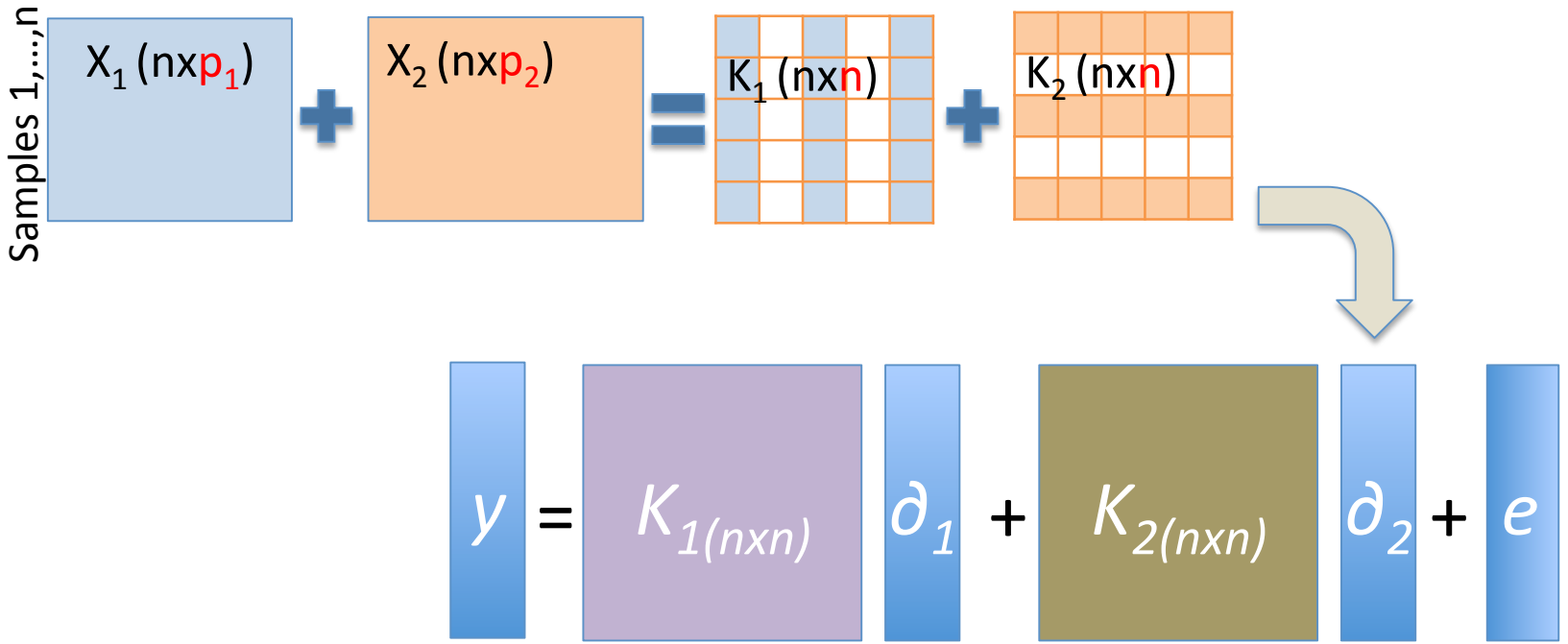
Existing Approaches: Kernel-Based Methods

- Intermediate integration



Existing Approaches: Kernel-Based Methods

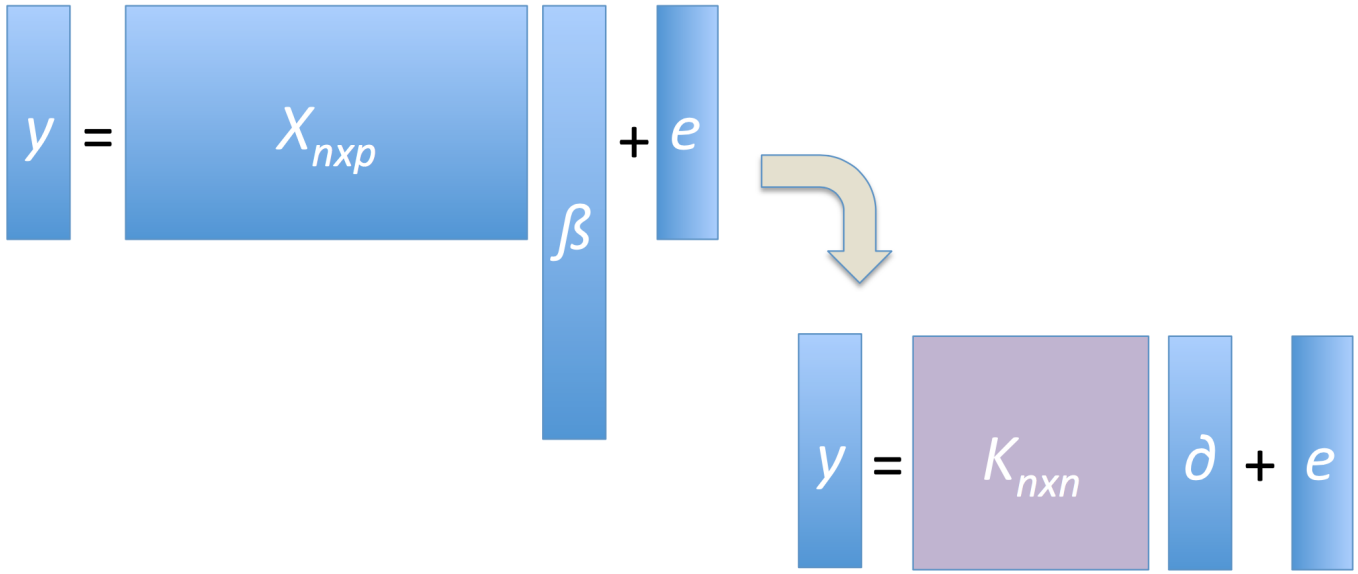
- Late integration



Kernel-Penalized Regression[1]

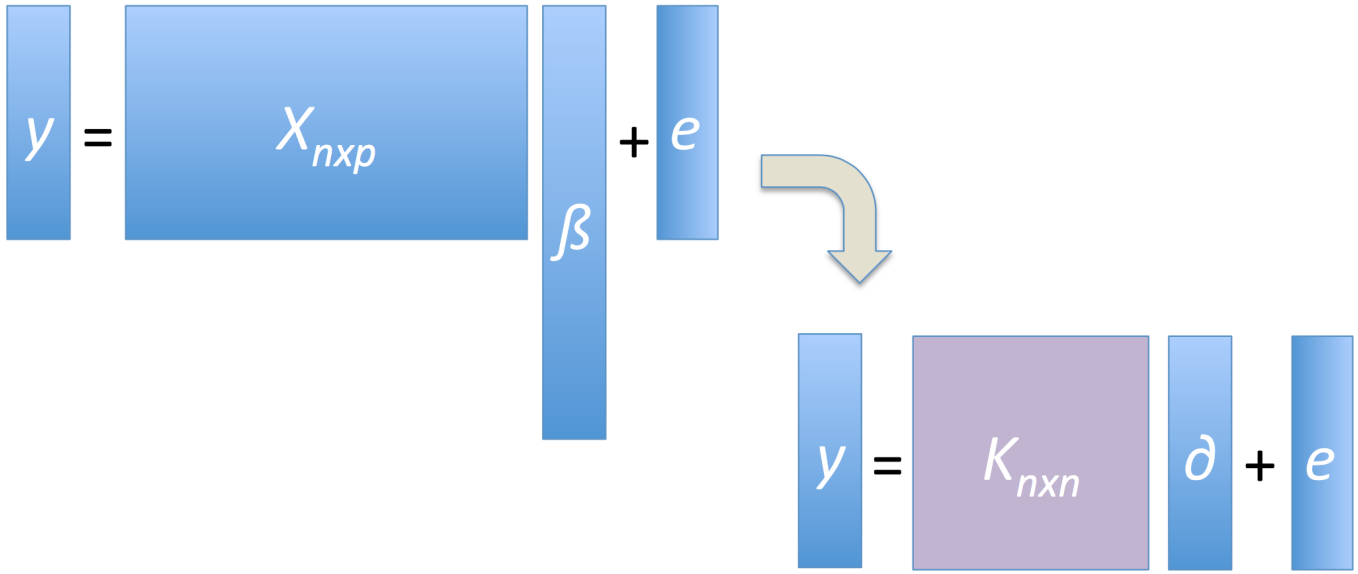
[1] Randolph et al (2018)

Kernel-Penalized Regression[1]



[1] Randolph et al (2018)

Kernel-Penalized Regression[1]



- How can we incorporate network information?
- How can we evaluate association of individual omics measures and the response (biomarker discovery)?

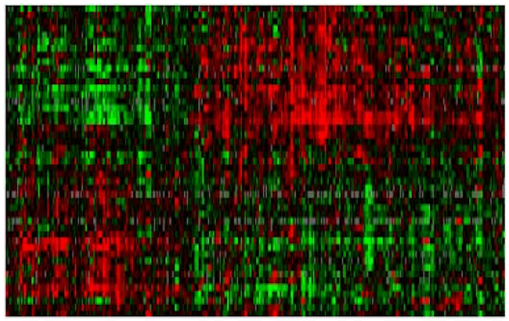
[1] Randolph et al (2018)

Kernel-Penalized Regression

Kernel-Penalized Regression

- Use the *duality* between the feature space ($\mathbb{R}^p$) and the observation space ($\mathbb{R}^n$) – formally, the duality diagram (Escoufier (1977), ...)

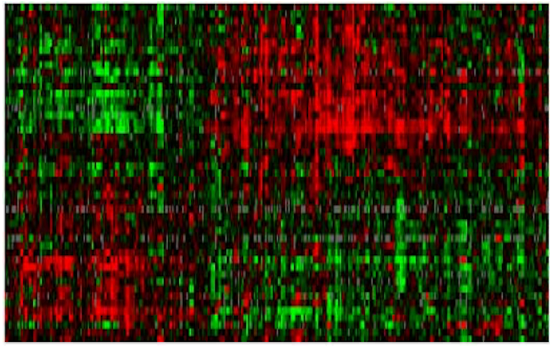
$$\mathbb{R}^p \underset{X}{\overset{X^\top}{\rightleftarrows}} \mathbb{R}^n$$

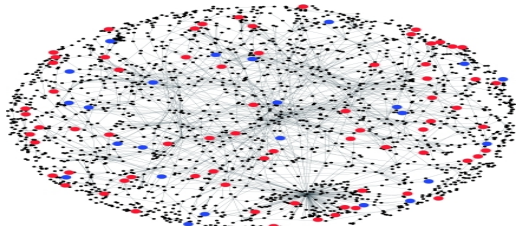


Kernel-Penalized Regression

- Use the *duality* between the feature space ($\mathbb{R}^p$) and the observation space ($\mathbb{R}^n$) – formally, the duality diagram (Escoufier (1977), ...)
- Can incorporate additional structure, e.g., network information

$$\mathbb{R}^p \underset{X}{\overset{X^\top}{\rightleftarrows}} \mathbb{R}^n$$

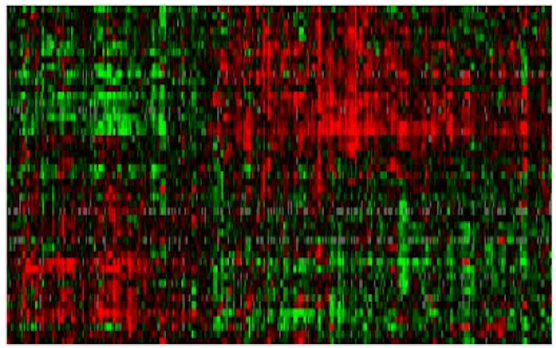


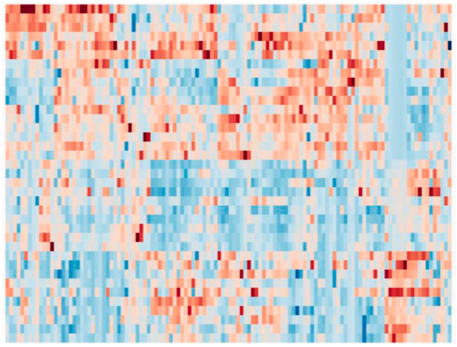


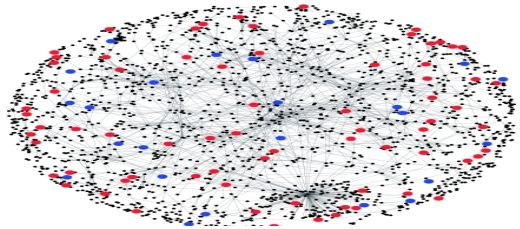
Kernel-Penalized Regression

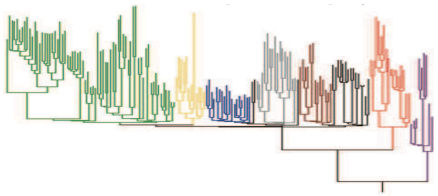
- Use the *duality* between the feature space ($\mathbb{R}^p$) and the observation space ($\mathbb{R}^n$) – formally, the duality diagram (Escoufier (1977), ...)
- Can incorporate additional structure, e.g., network information
- Can also incorporate multiple data omics data

$$\mathbb{R}^p \underset{X}{\overset{X^\top}{\leftrightarrows}} \mathbb{R}^n \underset{Z}{\overset{Z^\top}{\leftrightarrows}} \mathbb{R}^q$$









Example: Integrating Metabolomics Data

Integrating targeted (X_1) and unbiased (X_2) metabolomics profiling data for the same subjects

Example: Integrating Metabolomics Data

Integrating targeted (X_1) and unbiased (X_2) metabolomics profiling data for the same subjects

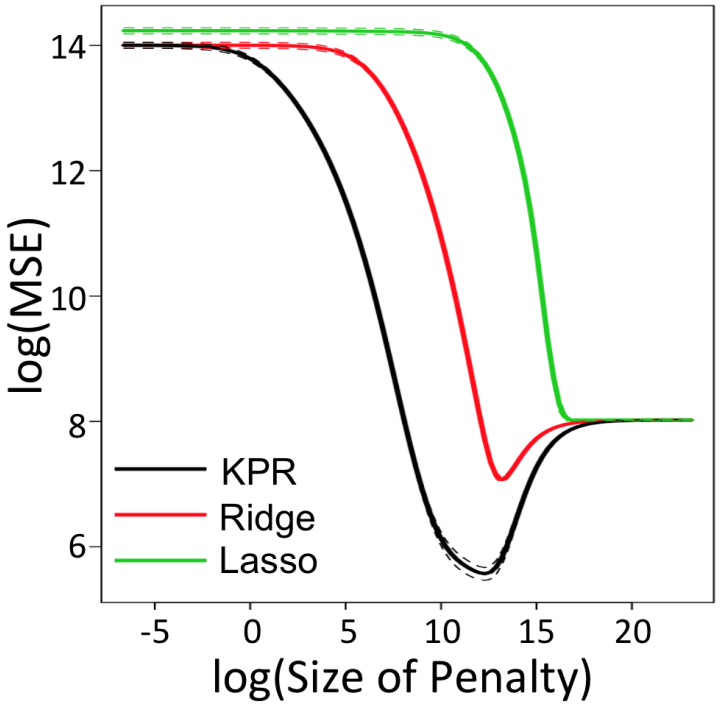
LOO CV-prediction error for the original data

	Lasso	Ridge	KPR
MSE	29.00	30.68	27.25

Example: Integrating Metabolomics Data

Integrating targeted (X_1) and unbiased (X_2) metabolomics profiling data for the same subjects

MSE for estimation of regression coefficients, based on our β^*



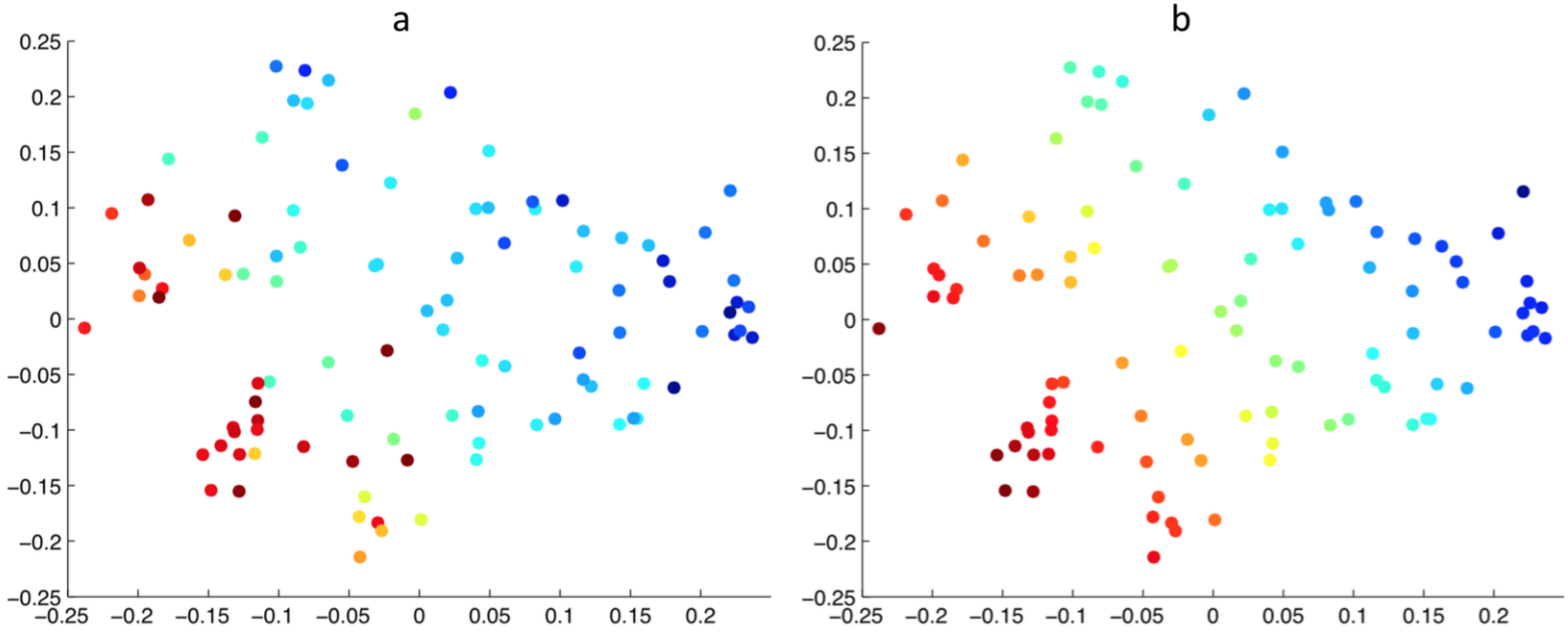
LOO CV-prediction error for the original data

	Lasso	Ridge	KPR
MSE	29.00	30.68	27.25

Another Example: Analysis of Microbiome Data

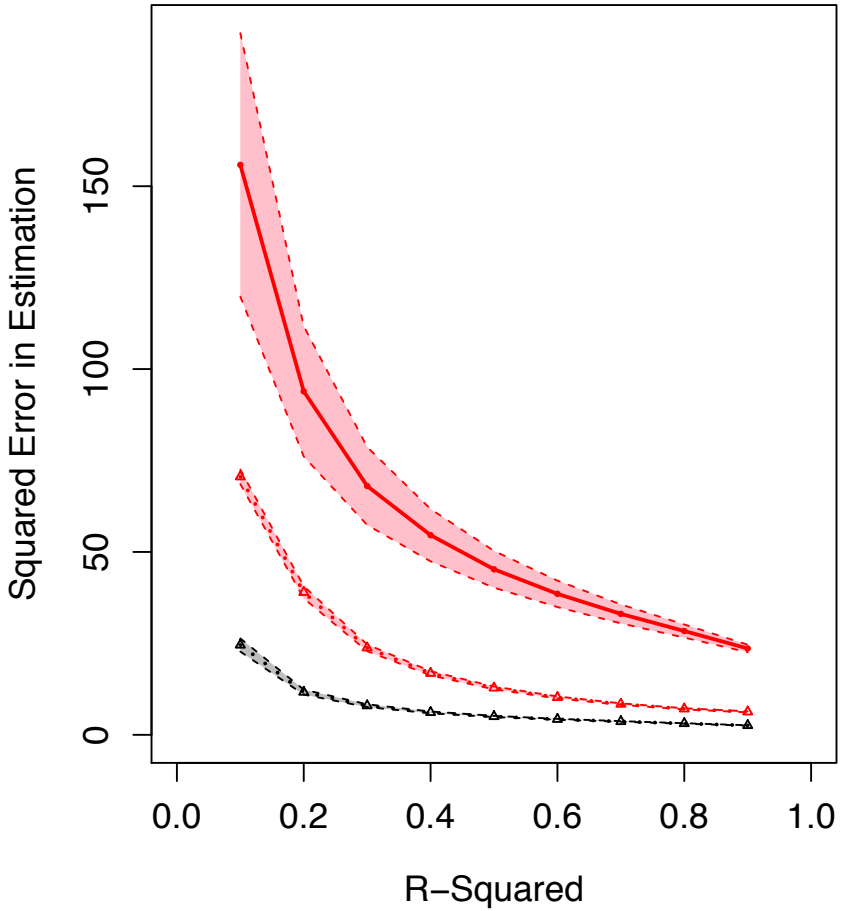
Simulation setup

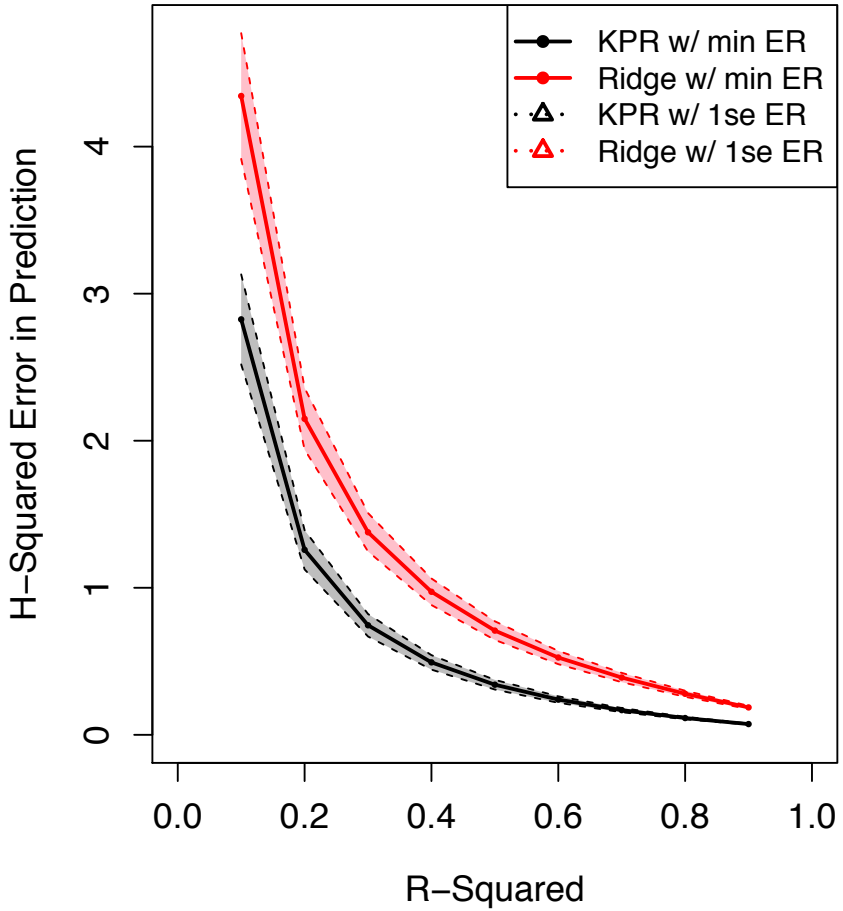
- *Simulate* the outcome based on real microbiome data:
  - we use the data from Yatsunenko et al (2012) consisting of $p = 495$ taxa for $n = 100$ subjects with $y = \log(age)$
  - the original study showed that a 2D-MDS based on the phylogenetic tree captures the pattern in response (left)
  - we generate y^* similarly in a phylogenetically-informed PCR (right)



Another Example: Analysis of Microbiome Data

Simulation results





Existing Approaches: Unsupervised Learning Methods

Existing Approaches: Unsupervised Learning Methods

Vertical integration

Existing Approaches: Unsupervised Learning Methods

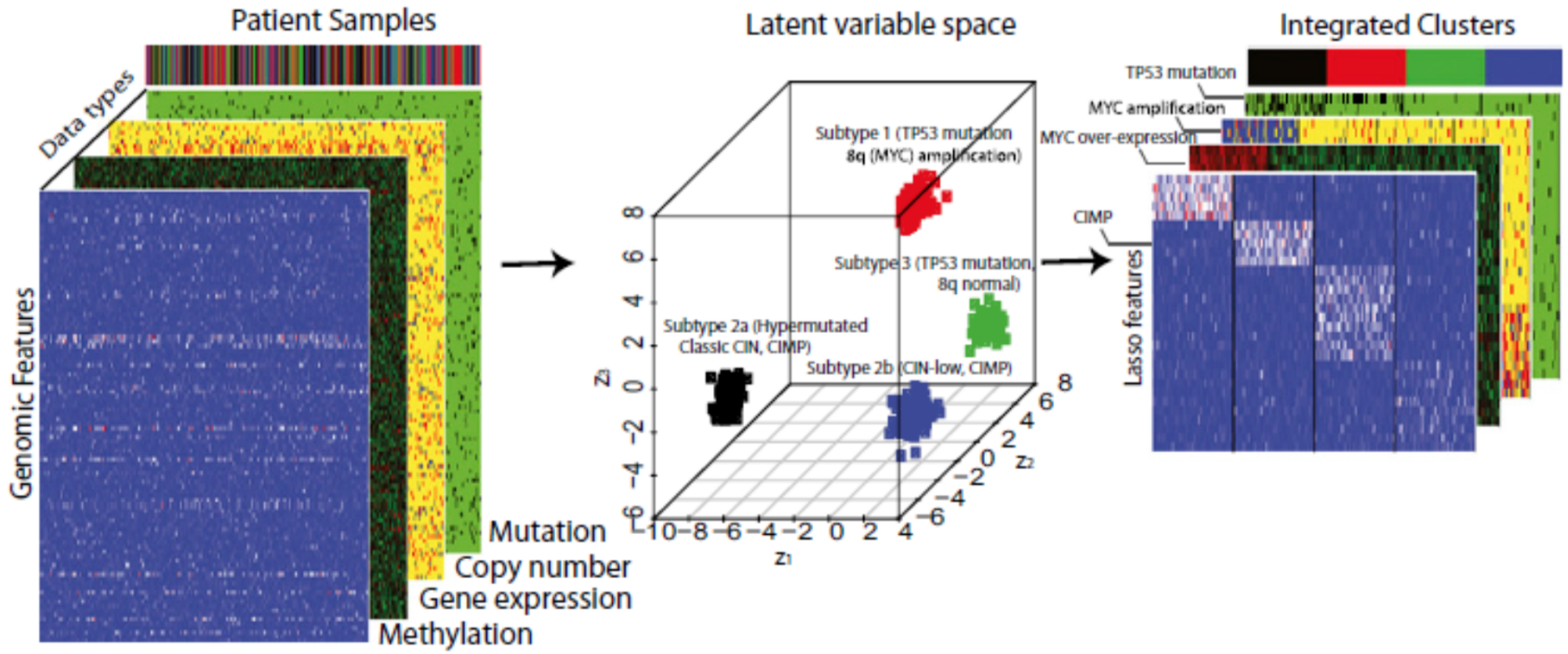
Vertical integration

- Integrative clustering

Existing Approaches: Unsupervised Learning Methods

Vertical integration

- Integrative clustering

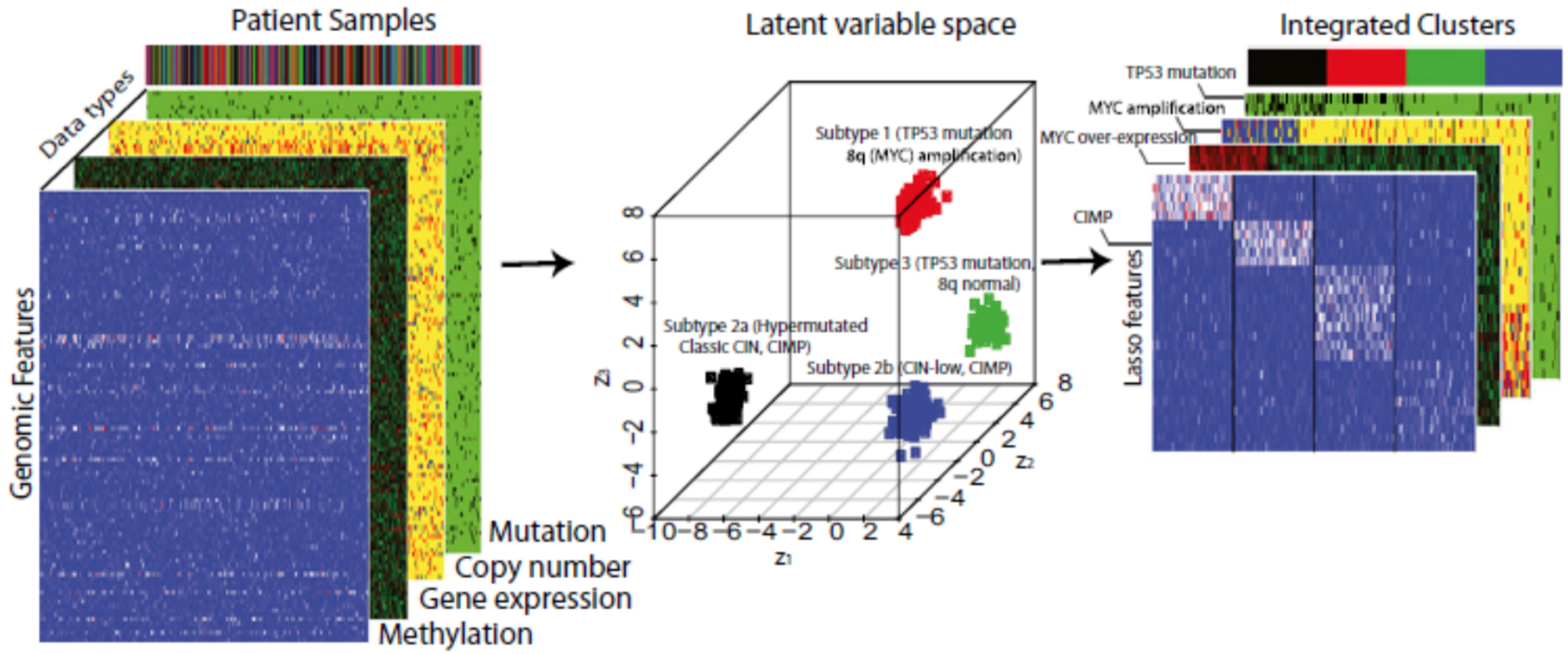


4246 | www.pnas.org/cgi/doi/10.1073/pnas.1208949110

Existing Approaches: Unsupervised Learning Methods

Vertical integration

- Integrative clustering



4246 | www.pnas.org/cgi/doi/10.1073/pnas.1208949110

- See Ronglai et al (2009, 2013); SungHwan et al (2015)

Existing Approaches: Unsupervised Learning Methods

Vertical integration

Existing Approaches: Unsupervised Learning Methods

Vertical integration

- Integrative dimension reduction

Existing Approaches: Unsupervised Learning Methods

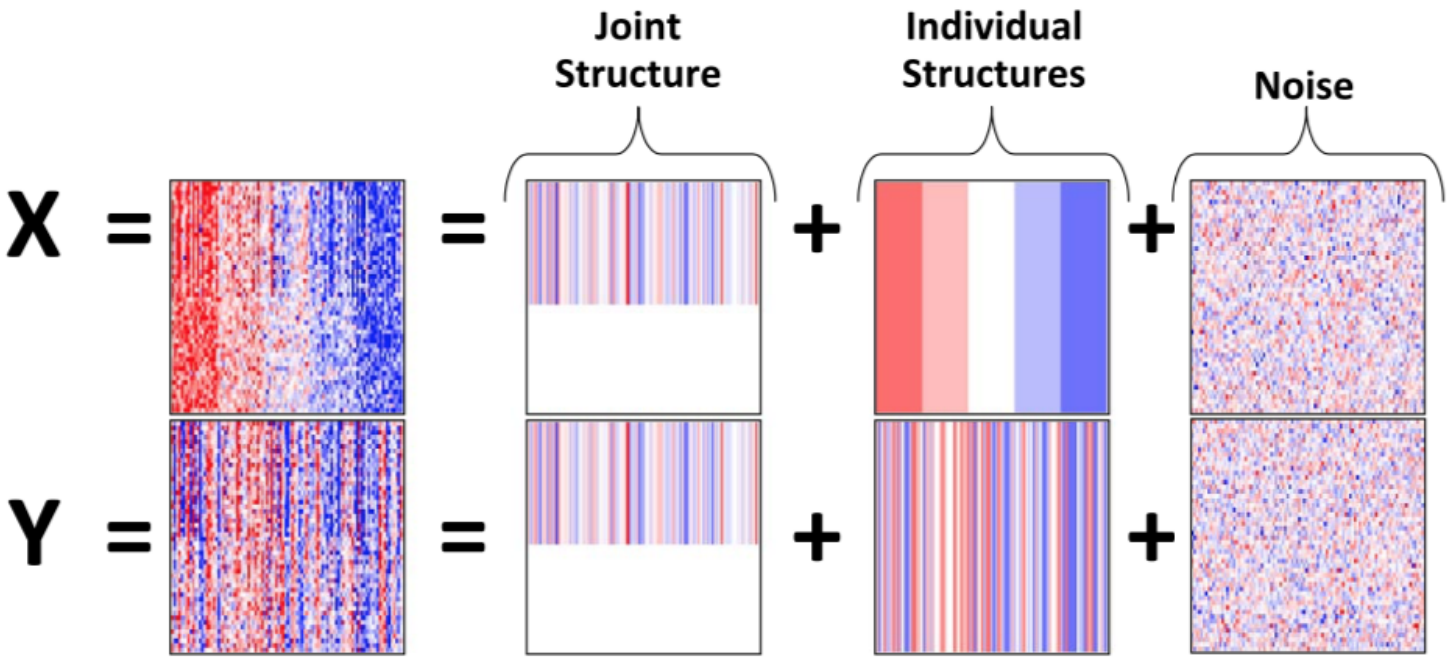
Vertical integration

- Integrative dimension reduction
  - Canonical Correlation Analysis (CCA), which looks for correlated omics measures — see, e.g. Witten et al (2009)

Existing Approaches: Unsupervised Learning Methods

Vertical integration

- Integrative dimension reduction
  - Canonical Correlation Analysis (CCA), which looks for correlated omics measures — see, e.g. Witten et al (2009)
  - Integrative Matrix Factorization (PCA, etc) — see Lock et al (2013); Argelaguet et al (2018)



Part III: Extensions

Starting Point

nature Vol 457 | 12 February 2009 | doi:10.1038/nature07762

LETTERS

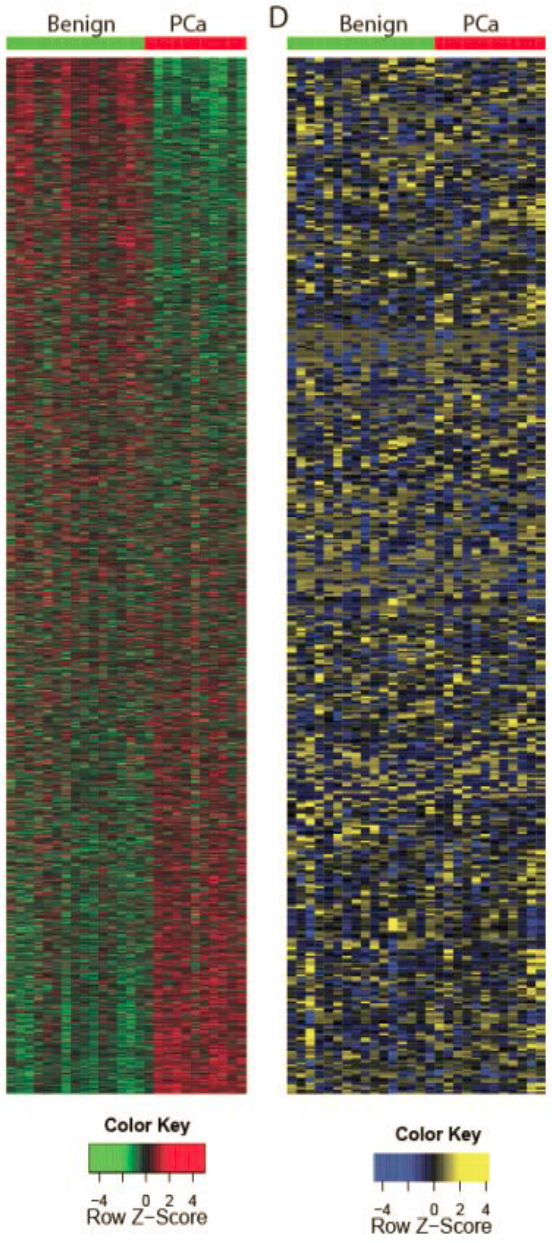
Metabolomic profiles delineate potential role for sarcosine in prostate cancer progression

Arun Sreekumar[1,2,3,4], Laila M. Poisson[5*], Thekkelnaycke M. Rajendiran[1,3*], Amjad P. Khan[1,3*], Qi Cao[1,3], Jindan Yu[1,3], Bharathi Laxman[1,3], Rohit Mehra[1,3], Robert J. Lonigro[1,4], Yong Li[1,3], Mukesh K. Nyati[4,6], Aarif Ahsan[6], Shanker Kalyana-Sundaram[1,3], Bo Han[1,3], Xuhong Cao[1,3], Jaeman Byun[7], Gilbert S. Omenn[2,7,8], Debashis Ghosh[4,5,11], Subramaniam Pennathur[2,4,7], Danny C. Alexander[12], Alvin Berger[12], Jeffrey R. Shuster[12], John T. Wei[4,9], Sooryanarayana Varambally[1,3,4], Christopher Beecher[1,2,3] & Arul M. Chinnaiyan[1,2,3,4,9,10]

Multiple, complex molecular events characterize cancer development and progression[1,2]. Deciphering the molecular networks that distinguish organ-confined disease from metastatic disease may lead to the identification of critical biomarkers for cancer invasion and disease aggressiveness. Although gene and protein expression have been extensively profiled in human tumours, little is known about the global metabolomic alterations that characterize neoplastic progression. Using a combination of high-throughput liquid-and-gas-chromatography-based mass spectrometry, we profiled more than 1,126 metabolites across 262 clinical samples related to prostate cancer (42 tissues and 110 each of urine and

were differential (Wilcoxon $P < 0.05$), with a false discovery rate (FDR) of 99%. Likewise, for urine, 36 out of 583 (6%) metabolites were differential (Wilcoxon $P < 0.05$), with an FDR of 67%. Thus, our initial focus was directed towards understanding the tissue metabolomic profiles as they showed more robust alterations.

Tissue samples were derived from benign adjacent prostate ($n = 16$), clinically localized prostate cancer ($n = 12$, PCA) and metastatic prostate cancer ($n = 14$) patients. Selection of metastatic tissue samples from different sites (see Supplementary Table 2) minimized characterization of analytes specific to cells of non-prostatic origin. In total, high-throughput profiling of the tissue quantitatively detected



Prostate Cancer...

- Prostate cancer (PCa) is the most common cancer in men
- About 221K new cases per year in the US
- About 28K deaths per year in the US – second leading cause of deaths in cancers (behind lung cancer)
- 5-year survival rate for localized PCa is nearly 100%

Prostate Cancer...(ctd.)

- PCa is driven by multiple factors & many genes implicated (androgen receptor, the TMPRSS2-ETS gene family fusion, BRCA1 and BRCA2)
- The prostate glands require androgen to work properly
- Androgen hormonal therapy is widely used in older patients (over 75 years) rather than radical prostatectomy or radiation therapy
- Castrate Resistant PCa (CRPC) does not respond to hormone (androgen) treatments or gets worse with hormone therapy
- poor survival prognostics for CRPC patients: mean survival time $\leq$ 2 yrs
- Precise molecular alterations driving CRPC not well-understood

Data from Sreekumar et al (2009)

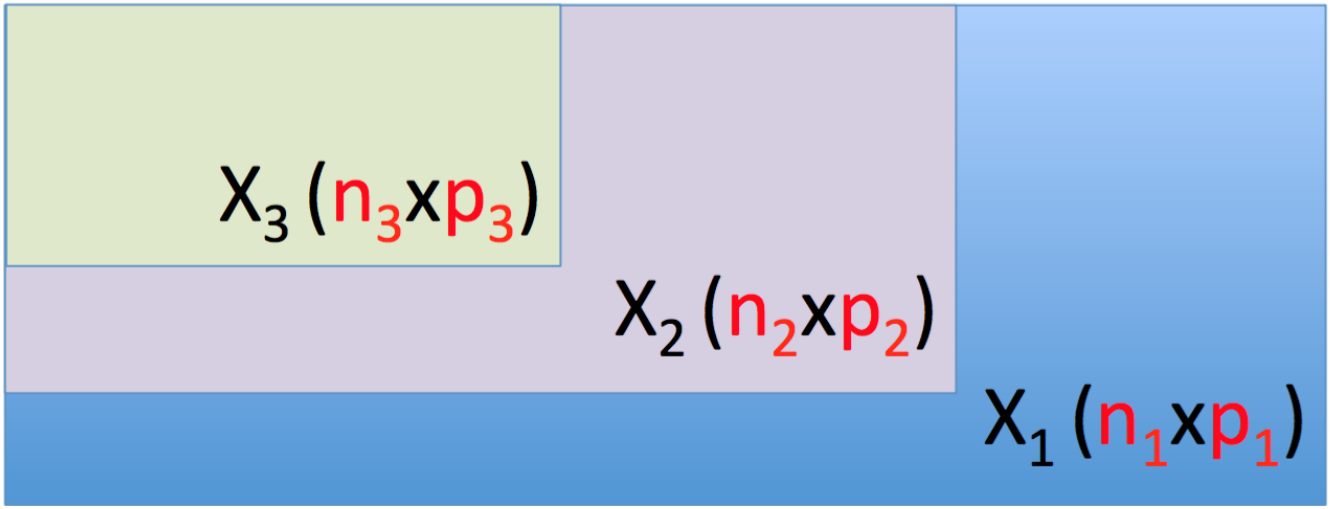
- Transcriptomic and metabolomic data for 12 PCa and 16 benign adjacent tissue samples
- Mostly matched samples, but few unmatched!
- Given the small sample size, need to
  - preserve all samples
  - reduce dimension

Omics Data Integration: Beyond Vertical and Horizontal

What if we have data on different platforms, but the samples don't match?

Omics Data Integration: Beyond Vertical and Horizontal

What if we have data on different platforms, but the samples don't match?

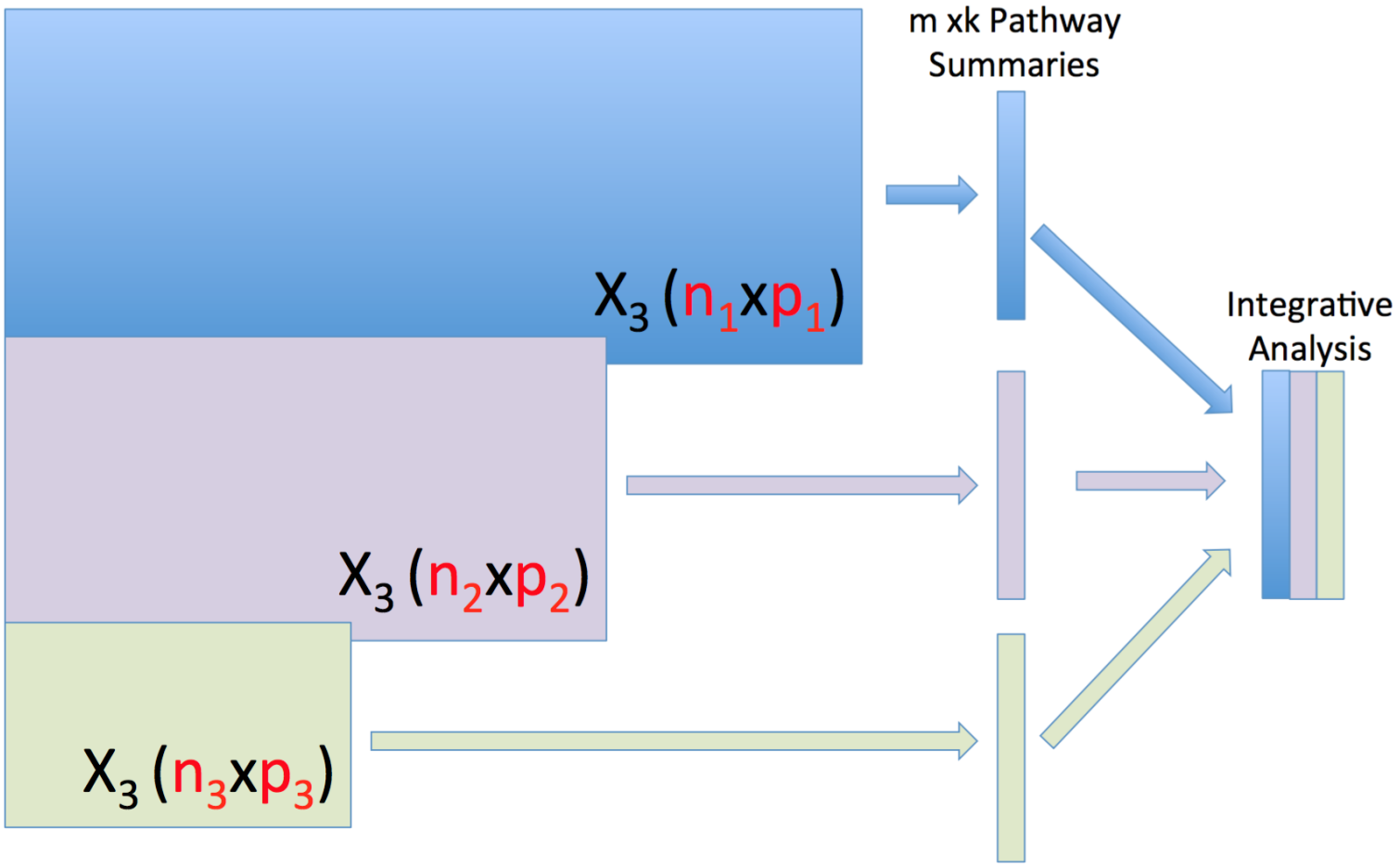


More on Omics Data Integration

Solution: Use pathways as the common dimension!

More on Omics Data Integration

Solution: Use pathways as the common dimension!

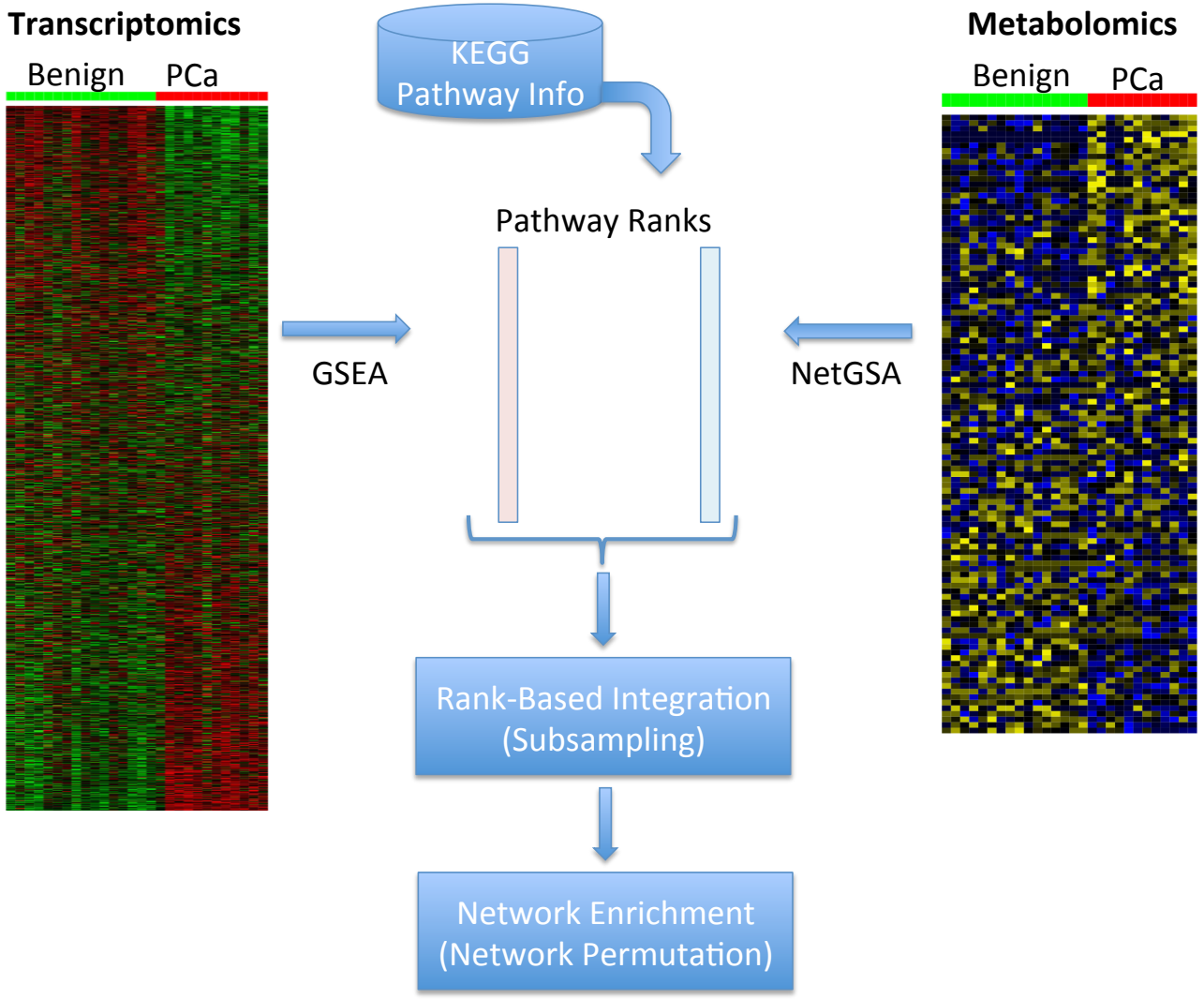


Rank-Based Integration

Metabolomics and Transcriptomics data from non-matching samples

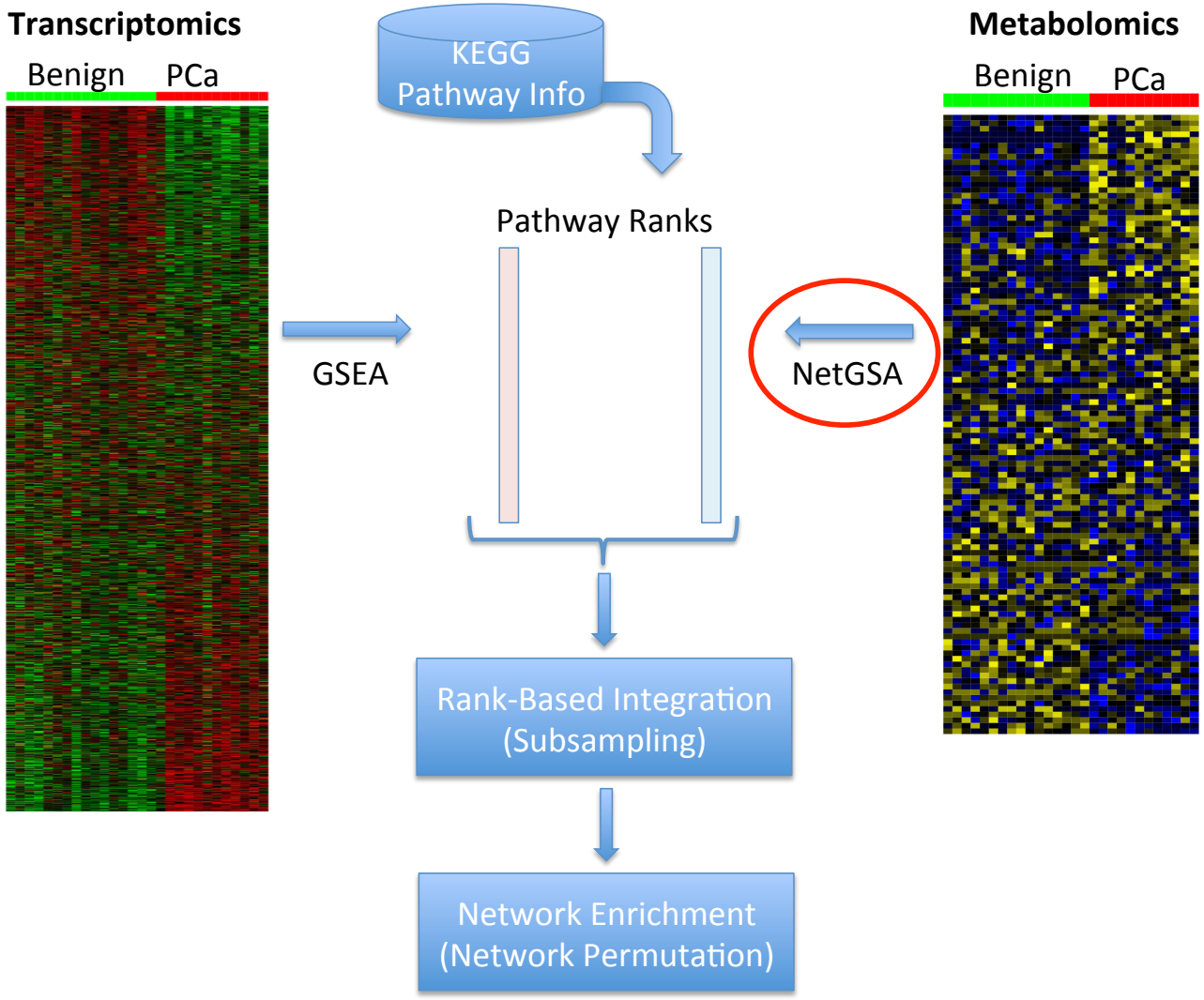
Rank-Based Integration

Metabolomics and Transcriptomics data from non-matching samples



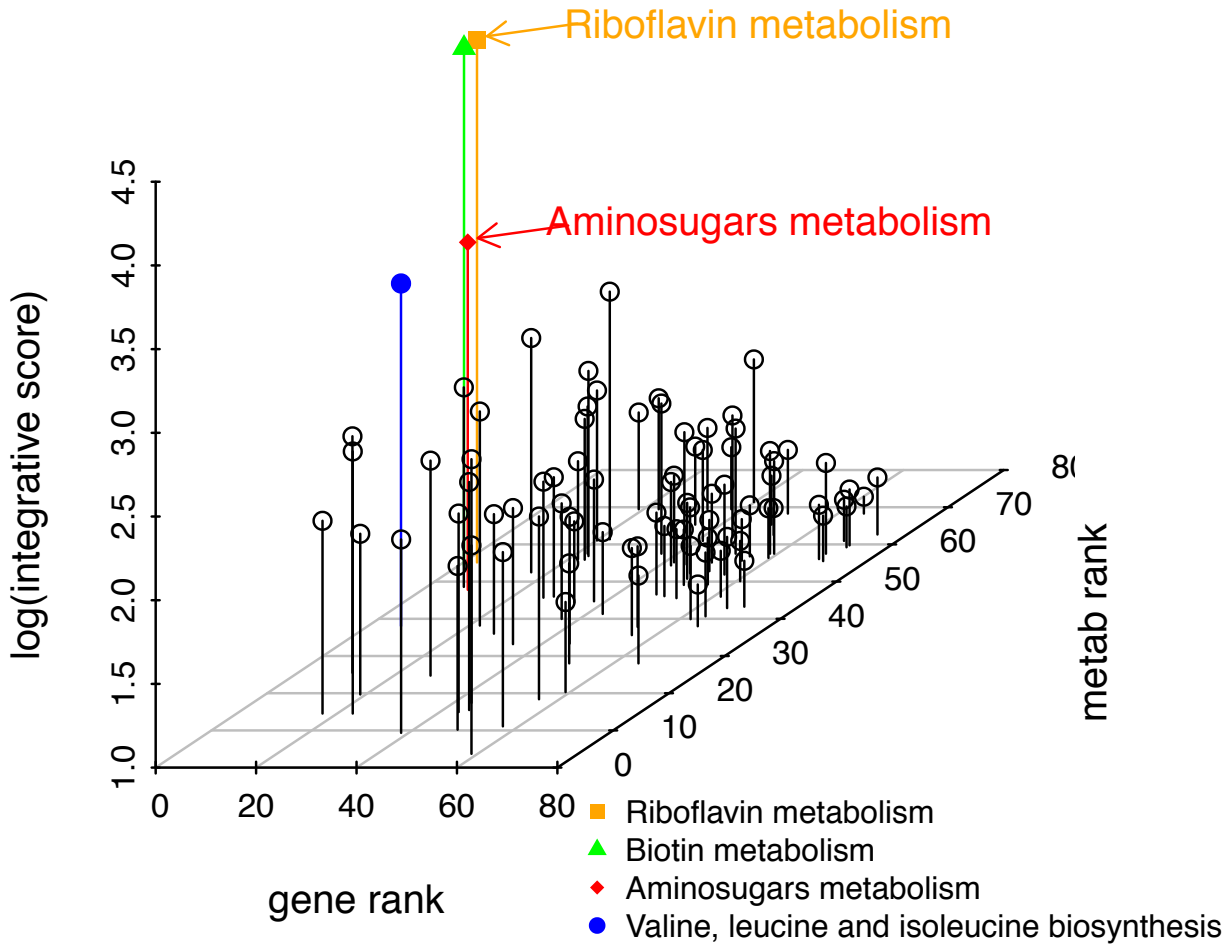
Rank-Based Integration

Metabolomics and Transcriptomics data from non-matching samples



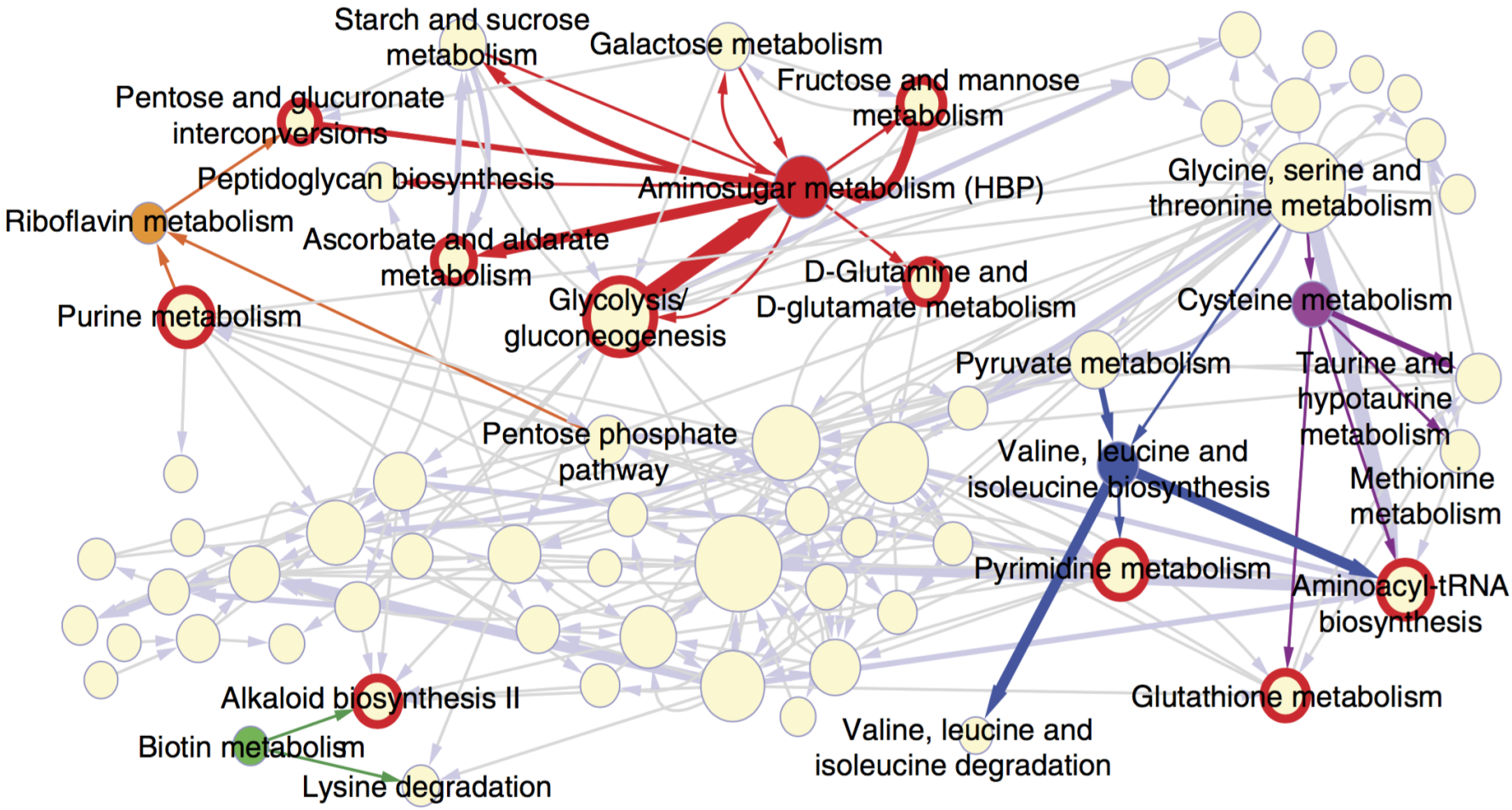
Step 1: Rank-Based Integration

Rankings vs Integrative Score



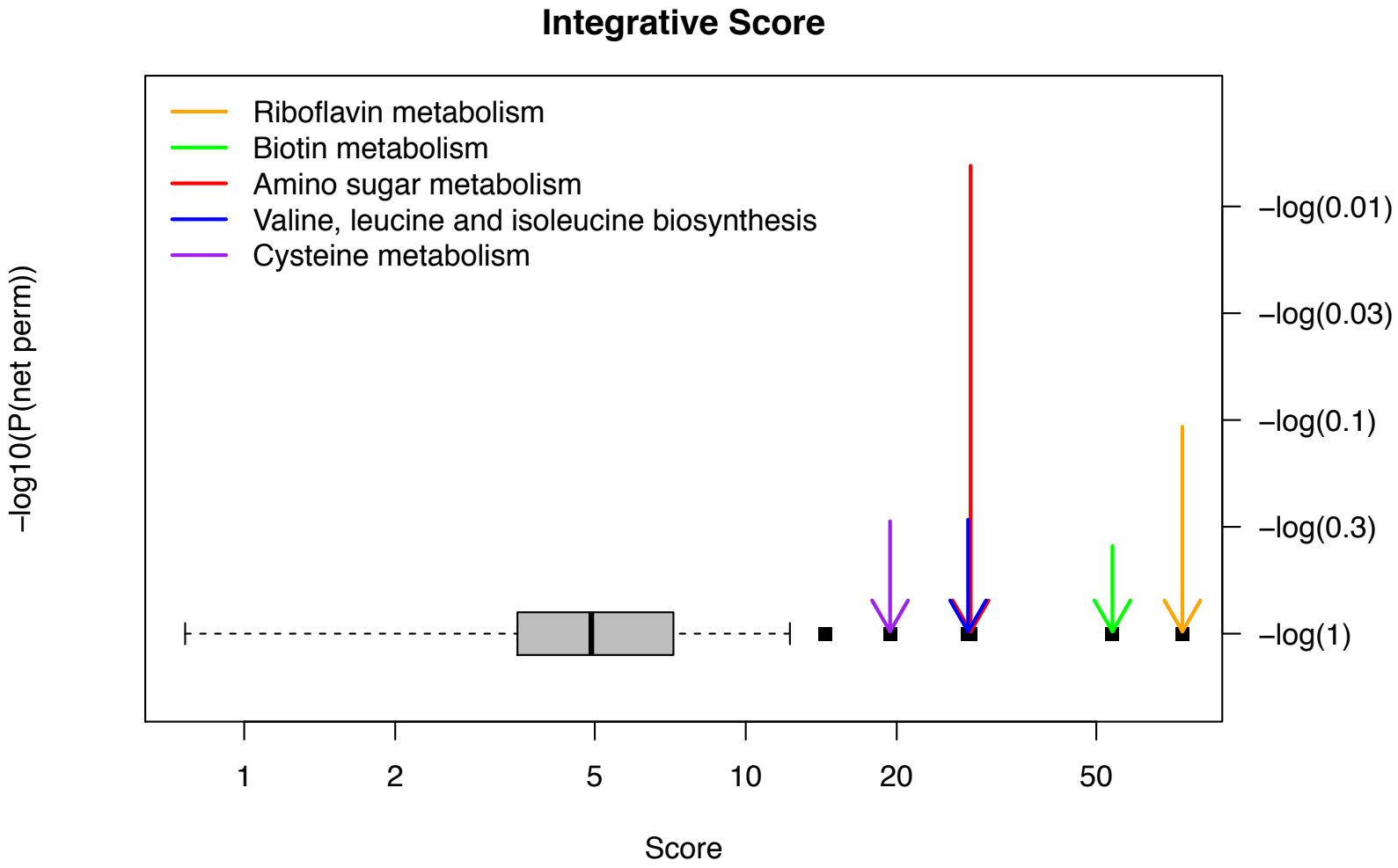
Step 2: Network Enrichment Analysis

Network permutation test to identify key pathways with active neighbors



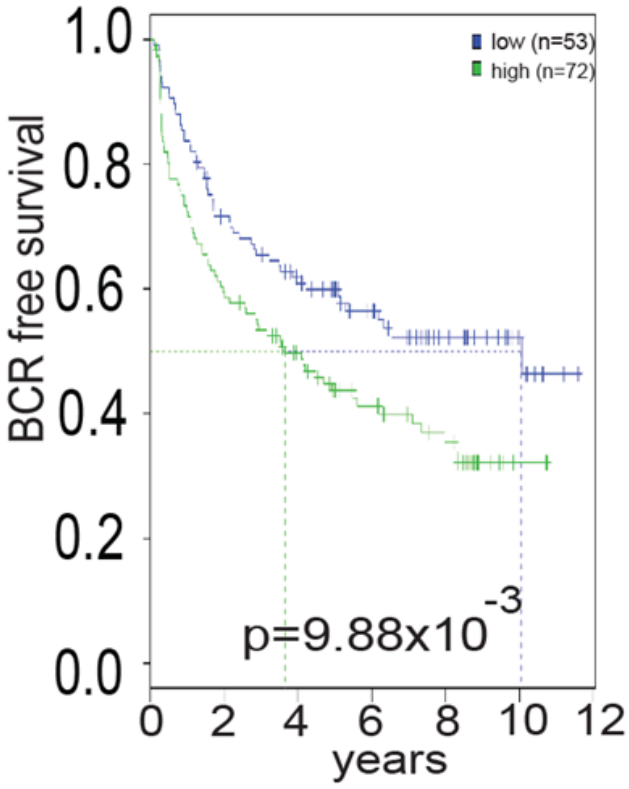
Putting Things Together

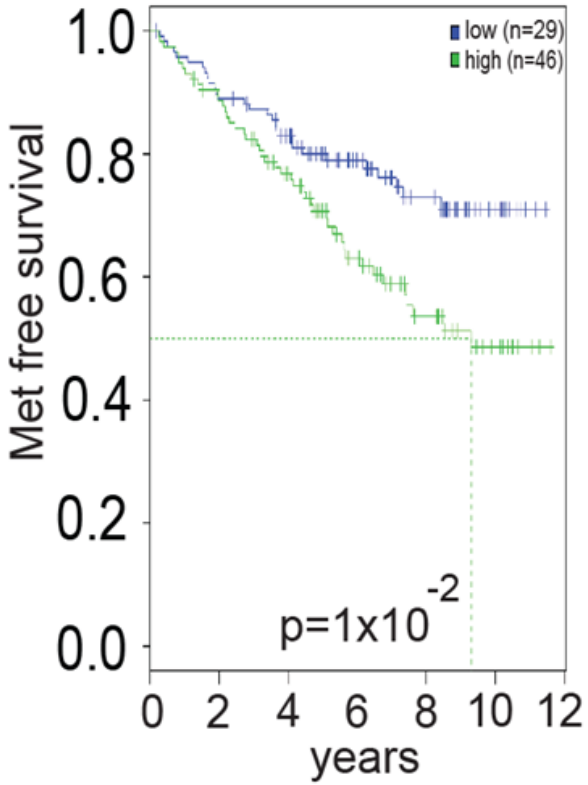
Rank-based integrative pathway scores vs. network enrichment p-values

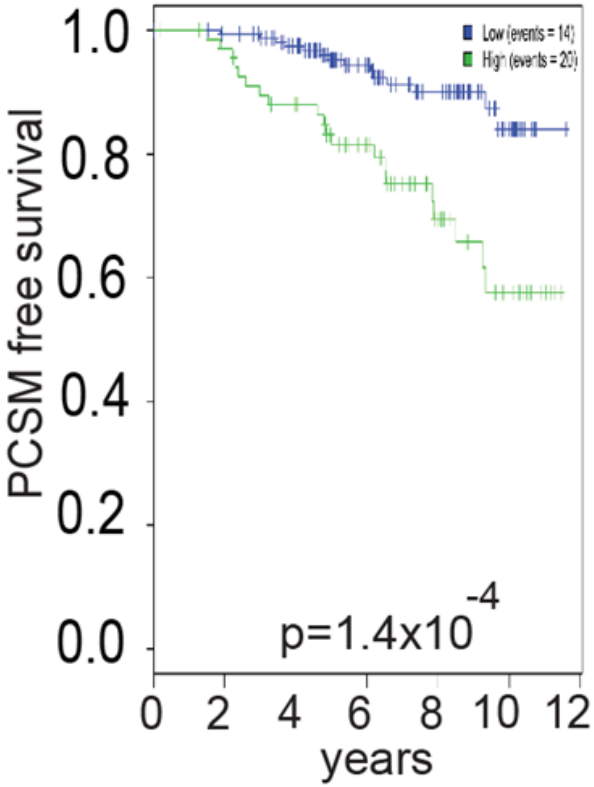


⇒ Aminosugar Metabolism, or Hexosamine Biosynthesis Pathway (HBP)

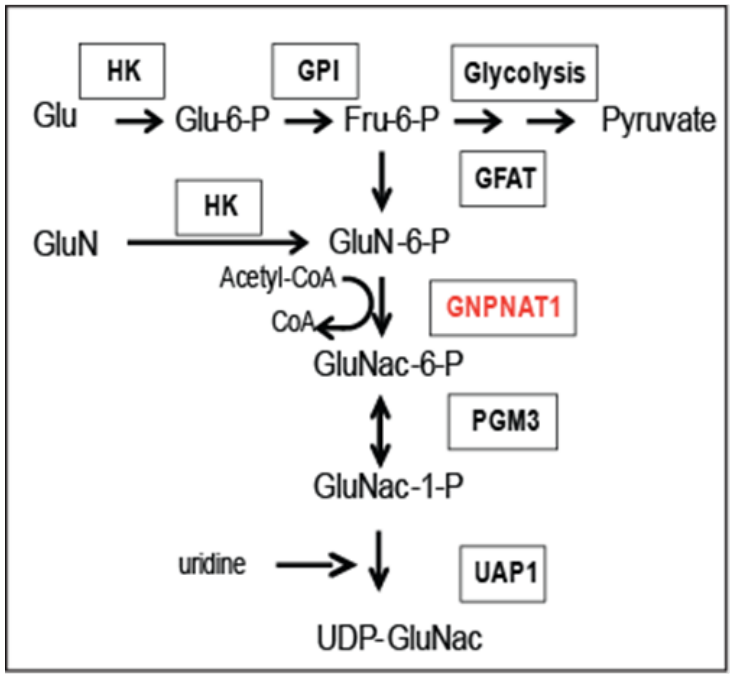
Clinical Relevance of HBP

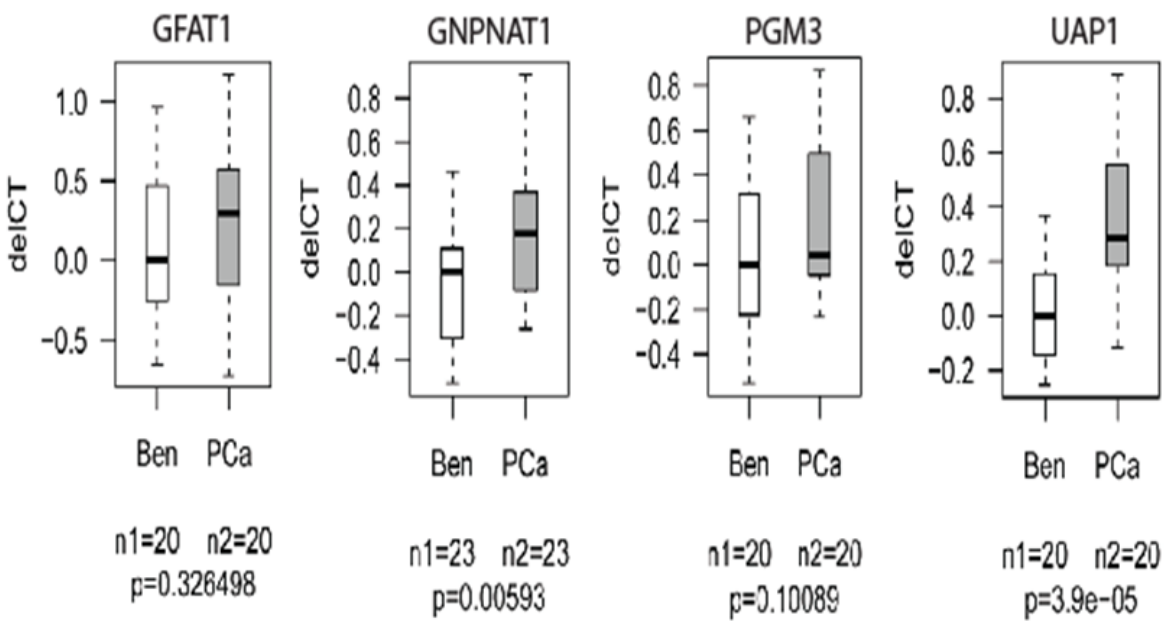




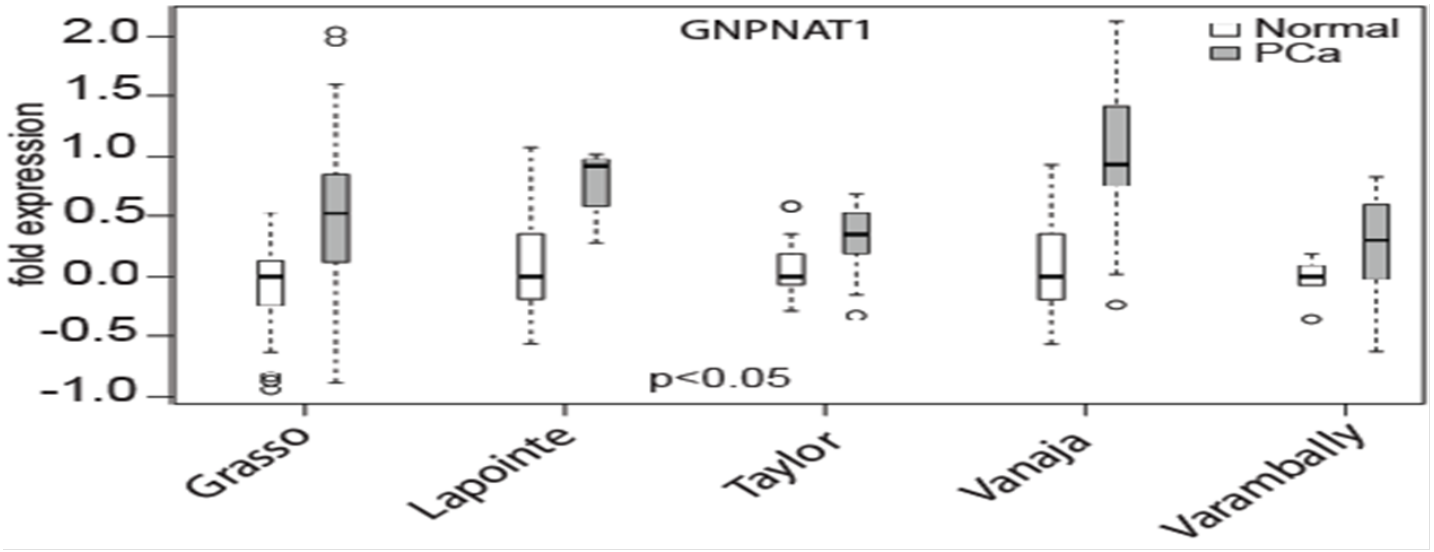


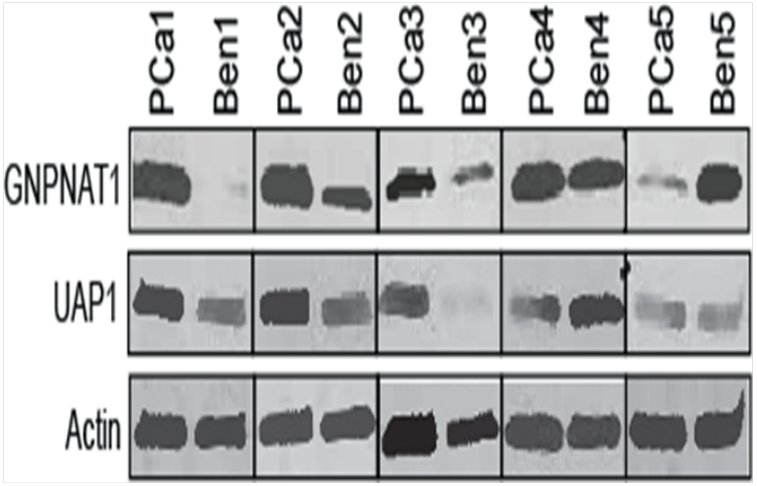
Expressions of HBP Genes in PCa

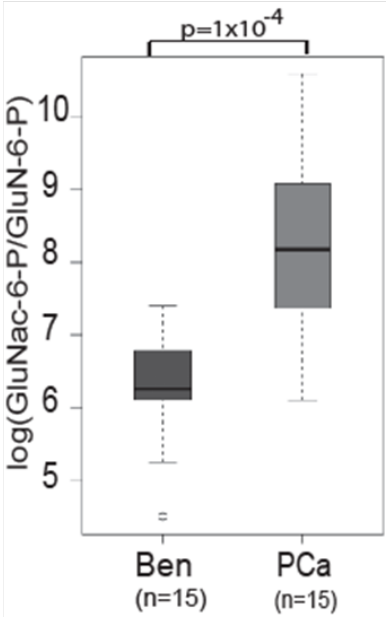




GNPNAT1 Expression in PCa







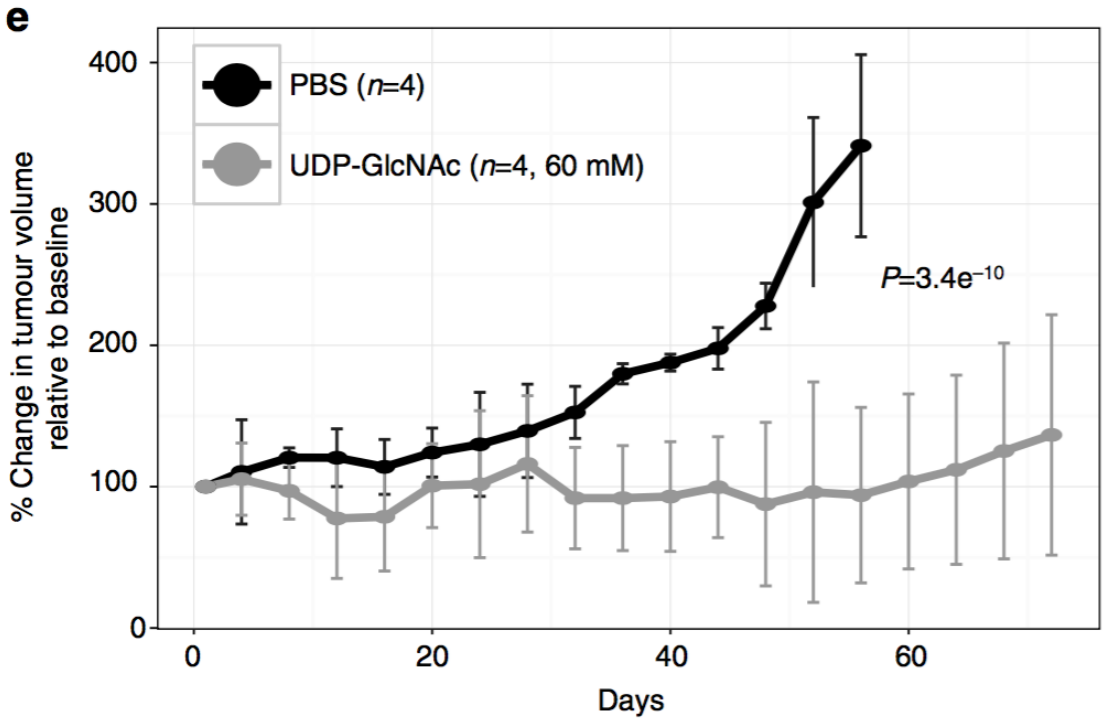
Therapeutic Potential

Therapeutic Potential

- HBP components elevated in localized PCa, but down-regulated in castrate resistant PCa (CRPC)
- Genetic loss of function experiments for GNPNAT1 in CRPC-like cells led to increased proliferation and aggressiveness, in vitro and in vivo

Therapeutic Potential

- Addition of HBP metabolite UDP-N-acetylglucosamine to CRPC-like cells reduced the expression of cell cycle genes and attenuated tumor cell proliferation, both in vitro and in vivo; also demonstrated additive efficacy when combined with enzalutamide in vitro



nature COMMUNICATIONS

ARTICLE

Received 21 Jan 2016 | Accepted 13 Apr 2016 | Published 19 May 2016

DOI: 10.1038/ncomms11612 OPEN

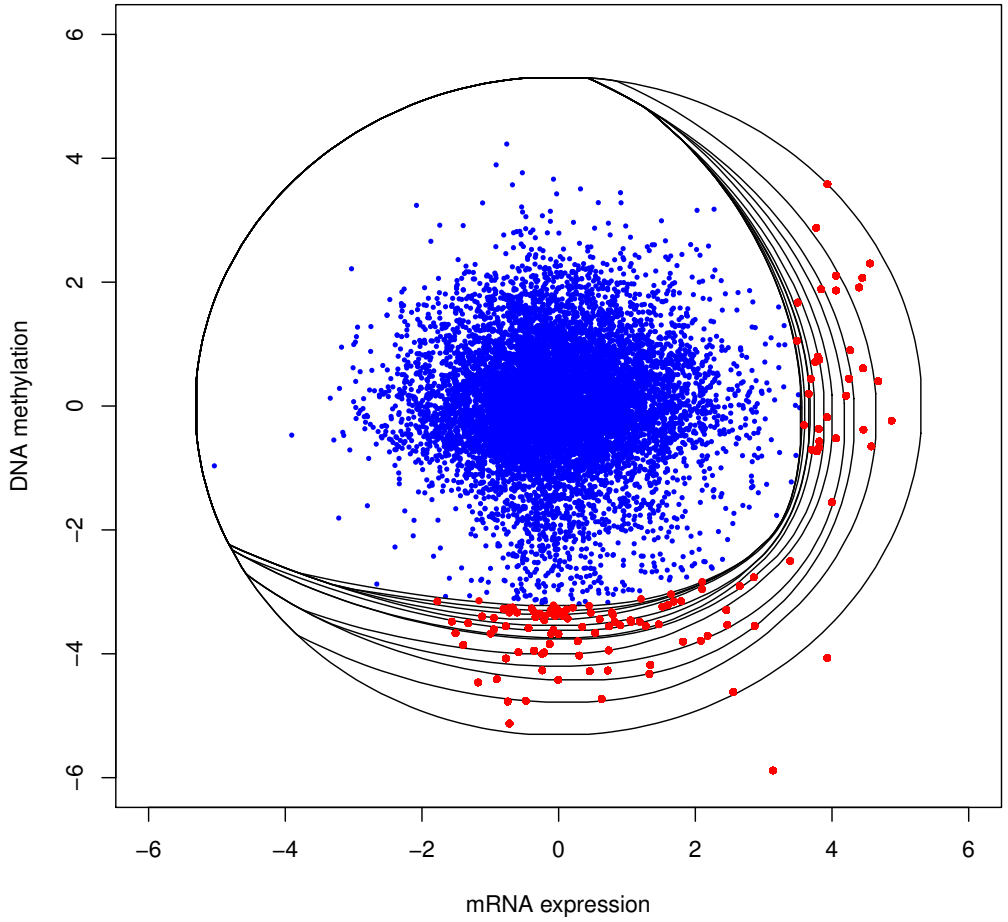
Inhibition of the hexosamine biosynthetic pathway promotes castration-resistant prostate cancer

Akash K. Kaushik[1,2,*], Ali Shojaie[3,*], Katrin Panzitt[1,†,*], Rajni Sonavane[1], Harene Venghatakrishnan[4,5], Mohan Manikkam[1], Alexander Zaslavsky[4,5], Vasanta Putluri[1], Vihas T. Vasu[6], Yiqing Zhang[1], Ayesha S. Khan[7], Stacy Lloyd[1], Adam T. Szafran[1], Subhamoy Dasgupta[1], David A. Bader[1], Fabio Stossi[1], Hangwen Li[4,5], Susmita Samanta[1], Xuhong Cao[4,5,8], Efrosini Tsouko[7], Shixia Huang[1,9], Daniel E. Frigo[7,10], Lawrence Chan[1,2], Dean P. Edwards[1,9], Benny A. Kaipparettu[11], Nicholas Mitsiades[1], Nancy L. Weigel[1], Michael Mancini[1], Sean E. McGuire[1], Rohit Mehra[4,8], Michael M. Ittmann[12], Arul M. Chinnaiyan[4,5,8,13], Nagireddy Putluri[1], Ganesh S. Palapattu[4,5], George Michailidis[14,†] & Arun Sreekumar[1,2,9]

Other Related Projects

Other Related Projects

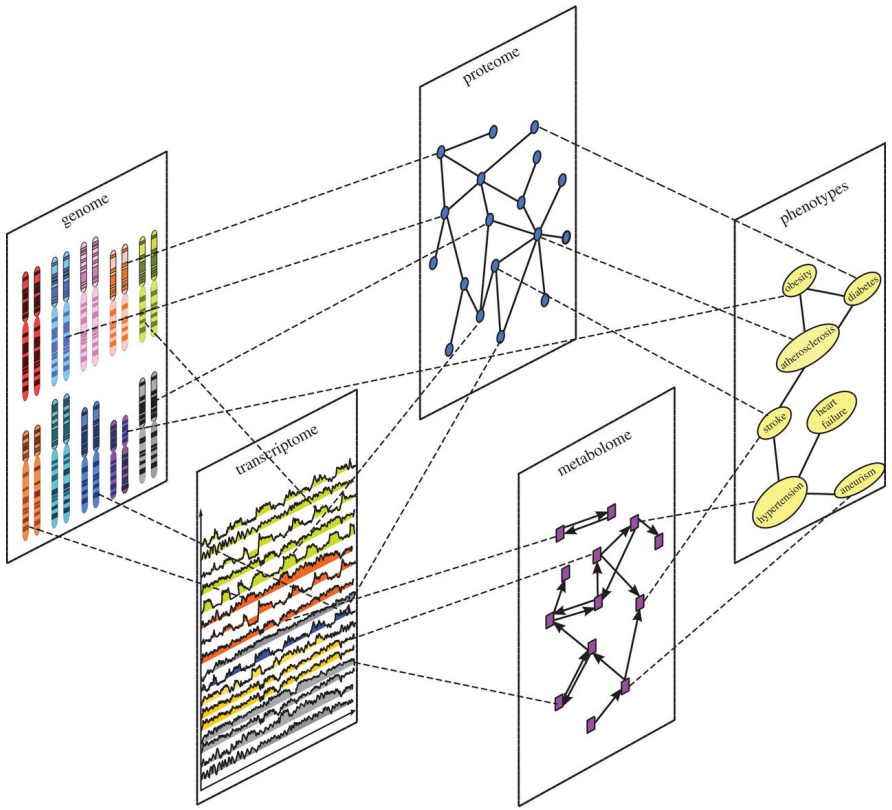
- FDR control for omics data integration (multivariate test statistics)[2]



[1] Alishahi, Ehyaei & **S.**, A generalized Benjamini-Hochberg procedure for multivariate hypothesis testing

Other Related Projects

- FDR control for omics data integration (multivariate test statistics)[2]
- Integrative multi-layer network analysis[3]

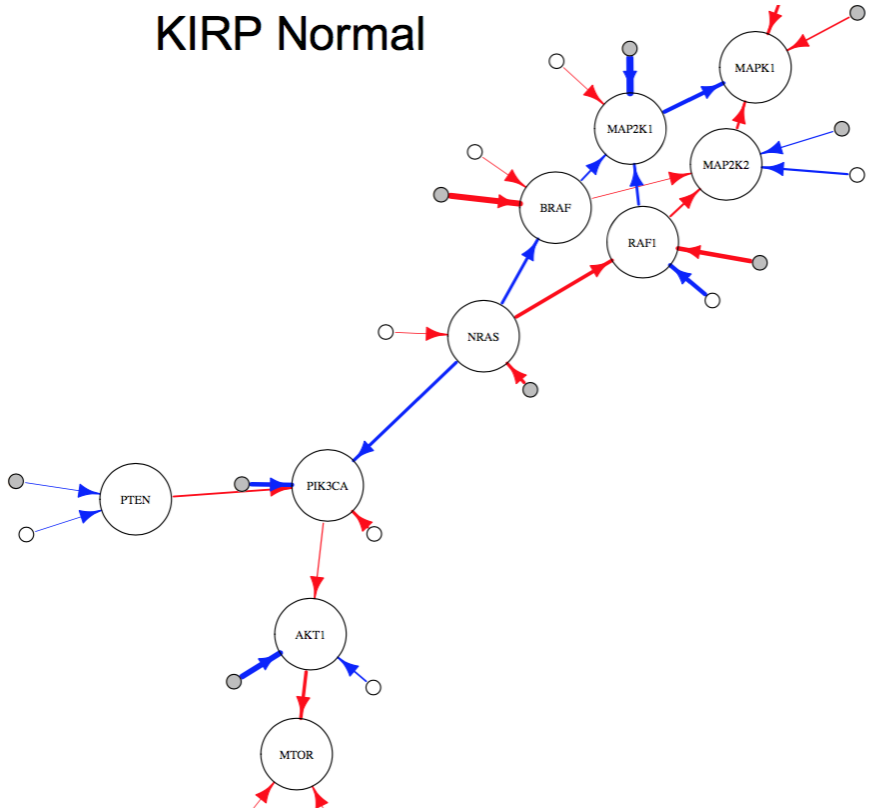


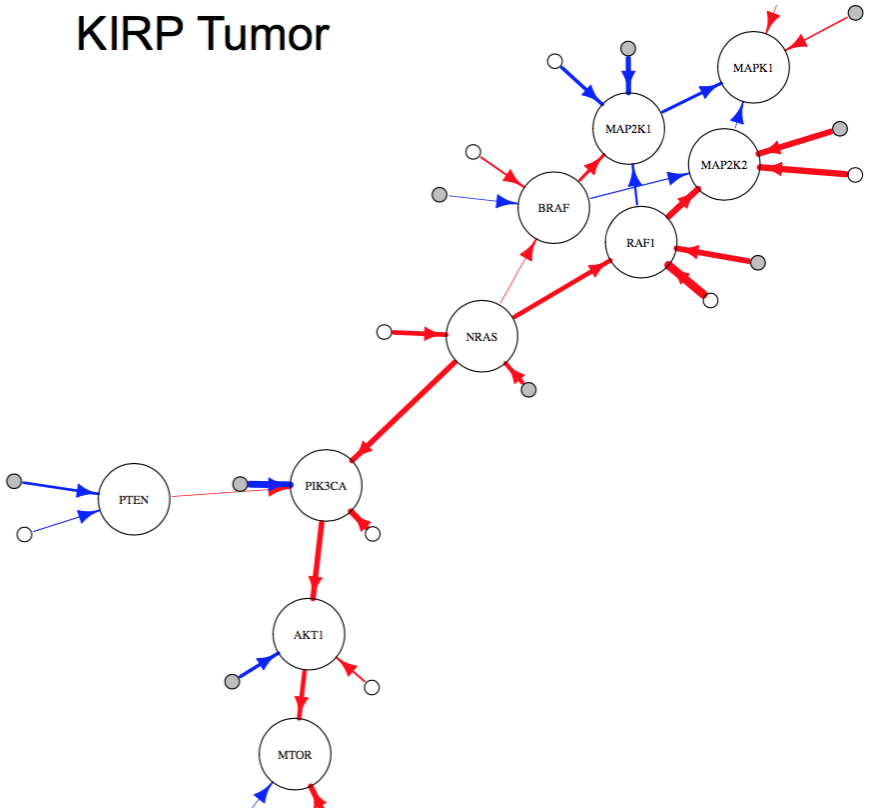
[1] Alishahi, Ehyaei & **S.**, A generalized Benjamini-Hochberg procedure for multivariate hypothesis testing

[2] Zhang et al (2018)

Other Related Projects

- FDR control for omics data integration (multivariate test statistics)[2]
- Integrative multi-layer network analysis[3]



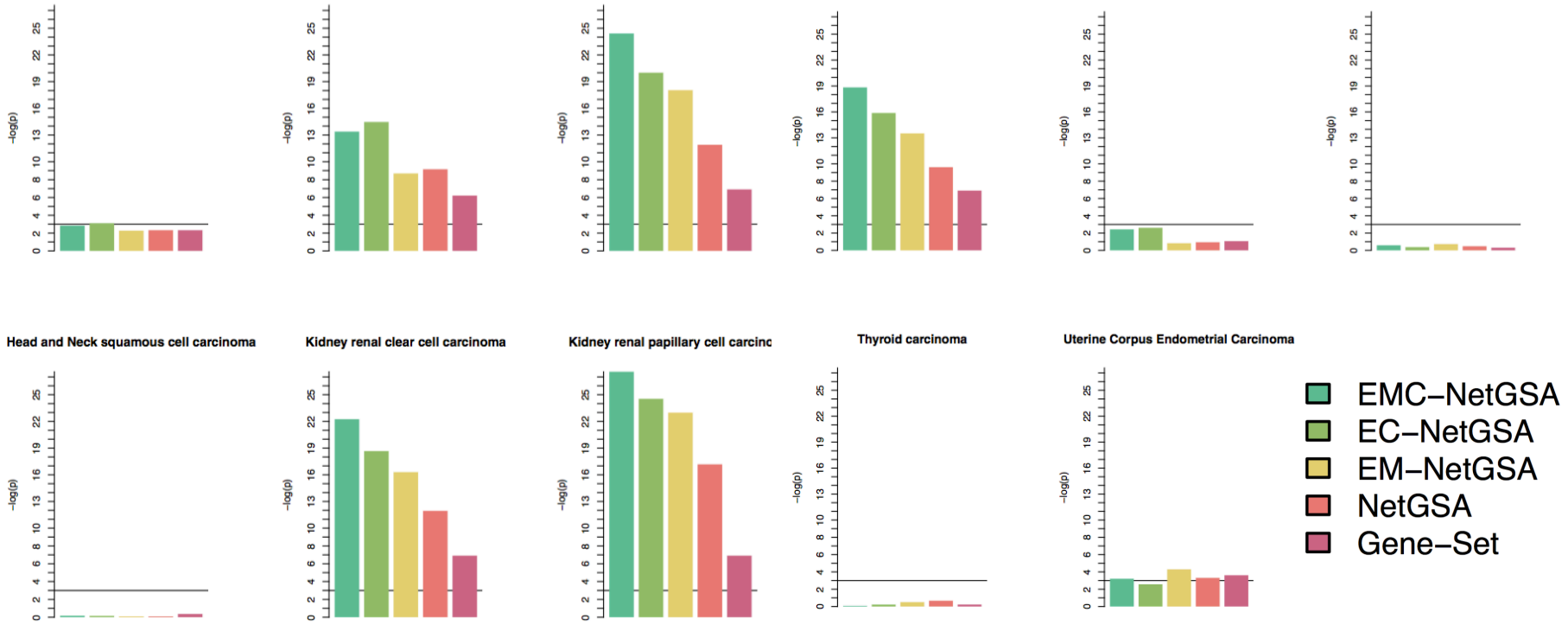


[1] Alishahi, Ehyaei & **S.**, A generalized Benjamini-Hochberg procedure for multivariate hypothesis testing

[2] Zhang et al (2018)

Other Related Projects

- FDR control for omics data integration (multivariate test statistics)[2]
- Integrative multi-layer network analysis[3]



[1] Alishahi, Ehyaei & **S.**, A generalized Benjamini-Hochberg procedure for multivariate hypothesis testing

[2] Zhang et al (2018)

What's Next?

What's Next?

Network-based integration of omics data over multiple subpopulations (horizontal and vertical!)