Contextualized Networks Reveal Heterogeneous Transcriptomic Regulation in Tumors at Sample-Specific Resolution

Caleb N. Ellington[*], Benjamin J. Lengerich[2,3], Thomas B.K. Watkins[4], Jiekun Yang[2,3], Hanxi Xiao[5], Manolis Kellis[2,3], and Eric P. Xing[1,6,7]

[1]Carnegie Mellon University
[2]Massachusetts Institute of Technology
[3]Broad Institute of MIT and Harvard
[4]Cancer Institute, University College London
[5]University of Pittsburgh
[6]Petuum, Inc
[7]Mohamed bin Zayed University of Artificial Intelligence

Abstract

Cancers are shaped by somatic mutations, microenvironment, and patient background, each altering both gene expression and regulation in complex ways, resulting in highly-variable cellular states and dynamics. Inferring gene regulatory networks (GRNs) from expression data can help characterize this regulation-driven heterogeneity, but network inference requires many statistical samples, traditionally limiting GRNs to cluster-level analyses that ignore intra-cluster heterogeneity. We propose to move beyond cluster-based analyses by using *contextualized* learning, a multi-task learning paradigm, to generate sample-specific GRNs from sample contexts. We unify three network classes (Correlation, Markov, Neighborhood) and estimate sample-specific GRNs for 7997 tumours across 25 tumor types, with each network contextualized by copy number and driver mutation profiles, tumor microenvironment, and patient demographics. Sample-specific GRNs provide a structured view of expression dynamics at sample-specific resolution, revealing co-expression modules in correlation networks (CNs), as well as cliques and independent regulatory elements in Markov Networks (MNs) and Neighborhood Regression Networks (NNs). Our generative modeling approach predicts GRNs for unseen tumor types based on a pan-cancer model of how somatic mutations affect transcriptomic regulation. Finally, sample-specific networks enable GRN-based precision oncology, explaining known biomarkers via network-mediated effects, leading to novel prognostic intra-disease and inter-disease subtypes.

Introduction

Tumors are heterogeneous, developing through clonal evolution that accumulates mutations, including cancer-driver single-nucleotide variants (SNVs) and somatic copy number alterations (SCNAs). In addition to tumor cell intrinsic changes, tumors develop in and are shaped by a microenvironment that includes immune cells, the extracellular matrix, blood vessels and surrounding cells. This extensive heterogeneity necessitates hetereogeneous treatments targeted to individual patients. To personalize treatment, precision oncology aims to discover prognostic biomarkers that stratify patients, and predictive biomarkers to identify effective therapies. While the identification of biomarkers using

[*]cellingt@cs.cmu.edu

NeurIPS 2023 Generative AI and Biology Workshop.

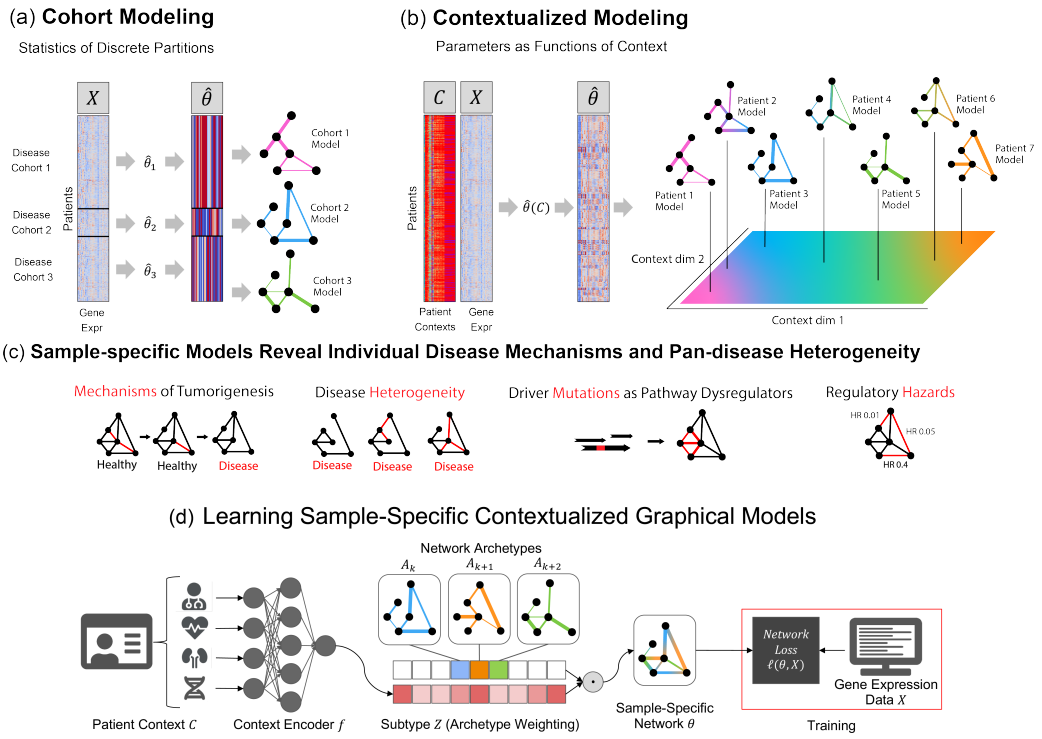


Figure 1: (a) Traditional modeling approaches assume each training (sub)population is homogeneous and samples are identically distributed. However, populations must be large enough to allow robust inference, presenting a tradeoff between personalization and power. (b) Contextualization assumes model parameters are a function of context, allowing powerful context-specific inference without *a priori* clustering of subpopulations or assuming population homogeneity. Contexts can be unique to each sample, permitting sample-specific model inference. (c) Sample-specific models reveal population heterogeneity, relate rare pathological mechanisms to more common ones, and provide new data views for prognosis and biomarker identification. (d) Graphical depiction of the deep learning framework. Sample context is used to predict weights on each of the model archetypes, which we call the subtype. The sample-specific network is estimated as the tensor dot product of archetypal networks and subtype weights. The network archetypes are learned simultaneously alongside the context encoder using backpropagation.

somatic DNA alterations or gene expression patterns has proved challenging [1], notable exceptions (e.g. HER2 amplification in breast cancer) motivate us to find a systematic way to search for differentiating regulatory factors that reflect cellular states and foreshadow cellular responses to treatments.

Modeling gene regulatory networks (GRNs) under a variety of clinical and molecular contexts can help to quantify this relationship between biomarkers, differential regulation, and tumor pathology. Single-cell and multiomic profiling have advanced the potential for studying highly context-specific regulatory relationships in GRNs, but computational methods of inferring GRNs continue to rely on partitioning samples into homogeneous sets of samples [2, 3, 4, 5]. As such, existing methods for high-resolution network inference either impose strong biological priors based on known transcription factor-gene regulation [6], or apply a sample-left-out approach that lacks statistical power [7, 8]. Partition-based modeling is insufficient to capture high-resolution or continuously rewiring GRNs, a problem for precision oncology because some types of cancer neither form discrete clusters [9] nor cluster by tissue of origin [10].

To infer tumor-specific GRNs that account for this heterogeneity, we propose to reframe GRN inference in a multi-task-learning paradigm that shares information between related molecular contexts. By recasting networks as the output of a learnable function, our approach shares statistical power between samples while also permitting fine-grained variation to capture the complexity of sample-

specific contexts such as tissue-of-origin, somatic mutation landscape, tumor microenvironment and clinical measurements. We formulate three types of GRNs (Markov, Neighborhood, and Correlation networks) under this paradigm, and estimate sample-specific GRNs which enable sample-specific analyses of latent regulatory processes. Applying this computational framework to over 7000 samples, we find that contextualized networks improve prediction of held-out expression data and reveal latent heterogeneity which has previously been obscured by partition-based methods of network inference.

Materials and Methods

Contextualized Networks We seek a context-specific density of model parameters $\mathbb{P}(\theta \mid C)$ where

$$\mathbb{P}(X|C) = \int_{\theta} d\theta \mathbb{P}_M(X|\theta)\mathbb{P}(\theta|C)$$

is maximized, where $\mathbb{P}_M(X|\theta)$ is the probability of gene expression X under network model M with parameters $\theta \in \mathbb{R}^{p\times p}$ and context C, which can contain both multivariate and real features. To overcome θ being a high-dimensional, structured latent variable, we assume that contextualized networks lie on a subspace measured by a latent variable ("subtype") $Z \in \mathbb{R}^K$ such that $C \perp (X, W) \mid Z$. Further, we constrain θ to be generated from network archetypes with latent mixing,

$$\mathbb{P}(\theta|C) = \delta\left(\theta - \sum_{k=1}^{K} f(C)_k A_k\right),$$

where A_k are archetypal networks, δ is the Dirac delta function, and $f : C \to Z$ is a deterministic context encoder. This architecture is shown in Fig. 1, and is learned without prior knowledge using end-to-end backpropagation. When estimated as CNs, network edges are the pairwise Pearson's correlation between nodes. When estimated as MNs, edges are the pairwise precision values representing conditional dependencies between nodes. When estimated as NNs, edges are directed interactions between nodes. Detailed estimation procedures are discussed in Appendix S2.

Our entire framework (Fig. 1) is implemented in PyTorch using the PyTorch Lightning framework within our open-source software ContextualizedML. The context encoder, network archetypes, and contextualized network models are learned simultaneously using end-to-end backpropagation of the network loss (defined in Section S2).

Data Our dataset is constructed from TCGA[2] and related studies, covering 7997 samples from 7648 patients with 6397 samples for training and validation and 1600 as testing. For context, we use clinical information, biopsy composition, SCNAs and SNVs (Appendix S3). Gene expression data was log-transformed and compressed to a set of cancer driver genes, then transformed using PCA into 50 metagenes. Networks were learned on the metagene expression data.

Results

Contextualization Recovers Latent Variation from Heterogeneous Observational Data Observational data is cheap and abundant in comparison to experimental data, but suffers from complex environmental confounders, contexts, and conditions. To model the latent processes of data generation using observational data, users must artificially control for on sample contexts and conditions to emulate a controlled experimental environment. However, controlling for all conditions and contexts simultaneously, especially on biomedical data with high-dimensional contexts, leads to conditions with as few as just one sample – too small to infer accurate context-specific models. Ignoring contextual effects (i.e. population modeling) is similarly ill-advised for heterogeneous data, leading to spurious results from models that are misspecified and inaccurate (e.g. Simpson's paradox).

Contextualization [11] addresses this by applying deep learning to the meta-relationship between contextual information and context-specific model parameters. Contextualization unifies previous approaches such as varying-coefficient modeling [12], cluster analysis, and cohort analysis by introducing two simple concepts: a context encoder which translates sample context into model parameters, and a sample-specific model which represents the latent context-specific mechanisms of data generation. By learning how models change in response to context, contextualization enables powerful control over high-dimensional and continuously varying contexts, discovering dynamic latent structure underlying data generation in heterogeneous populations.

[2] http://www.cancer.gov/tcga

Table 1: Fit of inferred networks to match held-out gene expression profiles. Mean-squared error (MSE) all three types of networks (Markov, Neighborhood, and Correlation) for gene expression predicted by the network. Reported values are mean ± std over 30 bootstrapped and randomly initialized runs. Error reduction is reported relative to the best baseline.

	Markov	Neighborhood	Correlation
Population	0.985 ± 0.006	0.984 ± 0.004	0.963 ± 0.000
Cluster-specific	0.365 ± 0.014	0.349 ± 0.012	0.683 ± 0.052
Disease-specific	0.368 ± 0.003	0.351 ± 0.003	0.673 ± 0.002
Contextualized	$\mathbf{0.322 \pm 0.014}$	$\mathbf{0.296 \pm 0.013}$	$\mathbf{0.529 \pm 0.019}$
Error Reduction	$14.6\% \pm 3.4\%$	$18.1\% \pm 3.3\%$	$20.2\% \pm 3.4\%$

Contextualization Enables Estimation of Sample-Specific Correlation, Markov, and Neighborhood Networks While GRNs are commonly interpereted as adjacency graphs [6, 2, 4], at their core existing methods for GRN inference can often be categorized as variants of three probabilistic models: CNs, MNs, and NNs. We unify these models via a linear reparameterizations equivalent to each models' unique constraints (See Appendix S2). Linear parameterization provides a differentiable objective for optimizing each model and the linear residual errors are proportional to the negative log likelihood of each network model under the data. Our unifying linearization of these models allows us to apply contextualization uniformly to each network class, and further enables us to benchmark and test the effects of common model personalization paradigms against contextualization in terms of model likelihood and modeling errors.

Contextualized Networks Improve Likelihood of Held-Out Expression Profiles We benchmark contextualized networks against common partition-based modeling regimes. Contextualization significantly improves the fit of networks models to gene expression data (Table 1). By accounting for contextual dependencies in model parameters, contextualized graphical models infer context-specific effects that can be overlooked by group-level modeling approaches (e.g. cluster-specific, disease-specific models), resulting more accurate representations of gene expression and regulation. Baselines are detailed in Section S2.1.

Contextualized Networks Share Power Between Cancer Types Contextualization relates transcriptional regulation to genomic variation through a context encoder (Fig. 1). During training, the encoder learns to modify the parameters of a network model in response to contextual signals. At test time, the encoder uses learned context signals to generalize between sparsely sampled contexts. Rare or undersampled diseases like KICH and GBM can benefit from contextual signals learned from well-sampled diseases in similar tissues, while disease-specific models cannot share information across disease types (Figure S1b). Furthermore, contextualization adapts models to unseen contexts at test time, responding to even extreme distribution shift (Fig. S1a).

Contextualized Networks Reveal Tissue-Specific Regulatory Modules Contextualization produces context-specific network models, resulting in patient-specific networks for all 7997 patients in our TCGA dataset. Organizing patients according to their network models reveals that tissue type is a primary driver, but not the sole factor in determining gene-gene interactions (Fig. S2). In particular, diseased networks differ drastically from healthy networks, while gene and PCA-derived metagene expression profiles are still largely tissue-derived. Additionally, intra-disease (Fig. S4a) and inter-disease (Fig. 2a) subtypes are visible even at pan-cancer resolution (Fig. S2), making obvious common tumorigenesis mechanisms that underly noisy gene expression dynamics.

Contextualized Networks Reveal Regulatory Modules Conserved Across Tissues in Cancer Cancers of the GI tract have a shared set of clinical subtypes according to state-of-the-art TCGA subtypes [13]. Contextualized networks reveal that tumors of the GI tract display a continuum of GRN dysregulation, and discover disparate types of GRN dysregulation within patients assigned to TCGA's GI.CIN subtype (Figure 2). Re-assigning patients based on GRN-derived subtypes improves prognosis (Fig. 2b) and reveals biomarkers of these dysregulation subtypes (Fig. 2a).

Contextualized Networks Reveal Novel Prognostic Subtypes For each of the 25 tumor types, we cluster patients by their contextualized network parameters to identify network-based tumor subtypes. We compare the subtypes' prognostic ability against state-of-the-art TCGA subtypes. On average,

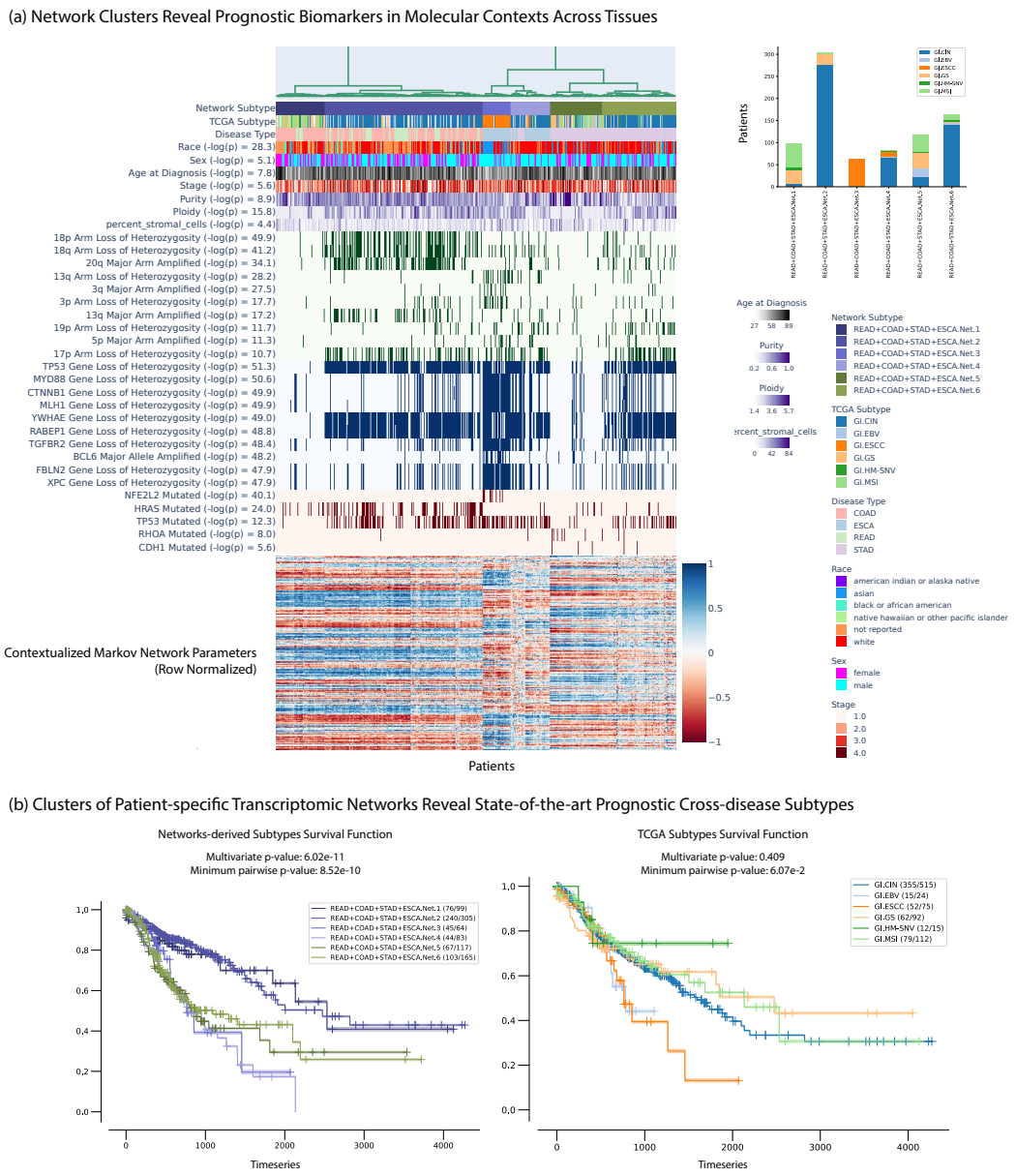


Figure 2: Exploration of network subtypes for tumors of the GI tract, including READ, COAD, STAD, and ESCA, looking at correlated molecular and clinical contexts.

GRN-based subtypes outperform TCGA subtypes and expression-derived subtypes by several orders of magnitude (Table S1). Detailed subtype comparisons are available in Appendices S1 and S5.

Discussion

In this study, we propose contextualized GRNs as cohesive sample-specific representations of latent tumor states underlying disease progression and survival. Our models reveal new insights about cancer heterogeneity by relating transcriptomic, genetic, immune, and clinical factors to through tumor regulatory network topology. More broadly, contextualized modeling seeks to estimate context-specific models beyond context-specific sampling constraints. By sharing information among samples while also allowing sample-specific variation, our framework models complex and dynamic distributions despite physical and technical barriers that would prohibit sample-specific inference.

References

[1] Rotem Ben-Hamo, Adi Jacob Berger, Nancy Gavert, Mendy Miller, Guy Pines, Roni Oren, Eli Pikarsky, Cyril H. Benes, Tzahi Neuman, Yaara Zwang, Sol Efroni, Gad Getz, and Ravid Straussman. Predicting and affecting response to cancer therapy based on pathway-level biomarkers. 11(1):3296.

[2] Pau Badia-i Mompel, Lorna Wessels, Sophia Müller-Dott, Rémi Trimbour, Ricardo O. Ramirez Flores, Ricard Argelaguet, and Julio Saez-Rodriguez. Gene regulatory network inference in the era of single-cell multi-omics. pages 1–16. Publisher: Nature Publishing Group.

[3] Matthew Stone, Sunnie Grace McCalla, Alireza Fotuhi Siahpirani, Viswesh Periyasamy, Junha Shin, and Sushmita Roy. Identifying strengths and weaknesses of methods for computational network inference from single cell RNA-seq data. Publication Title: bioRxiv.

[4] Aditya Pratapa, Amogh P. Jalihal, Jeffrey N. Law, Aditya Bharadwaj, and T. M. Murali. Benchmarking algorithms for gene regulatory network inference from single-cell transcriptomic data. 17(2):147–154.

[5] Dawn Thompson, Aviv Regev, and Sushmita Roy. Comparative analysis of gene regulatory networks: from network reconstruction to evolution. 31:399–428.

[6] Sara Aibar, Carmen Bravo González-Blas, Thomas Moerman, Vân Anh Huynh-Thu, Hana Imrichova, Gert Hulselmans, Florian Rambow, Jean-Christophe Marine, Pierre Geurts, Jan Aerts, Joost van den Oord, Zeynep Kalender Atak, Jasper Wouters, and Stein Aerts. SCENIC: single-cell regulatory network inference and clustering. 14(11):1083–1086. Number: 11 Publisher: Nature Publishing Group.

[7] Marieke Lydia Kuijjer, Matthew George Tung, Guocheng Yuan, John Quackenbush, and Kimberly Glass. Estimating sample-specific regulatory networks. 14:226–240.

[8] Olga Lazareva, Zakaria Louadi, Johannes Kersting, Jan Baumbach, David B. Blumenthal, and Markus List. DysRegNet: Patient-specific and confounder-aware dysregulated network inference. Pages: 2022.04.29.490015 Section: New Results.

[9] Oana Ursu, James T. Neal, Emily Shea, Pratiksha I. Thakore, Livnat Jerby-Arnon, Lan Nguyen, Danielle Dionne, Celeste Diaz, Julia Bauman, Mariam Mounir Mosaad, Christian Fagre, April Lo, Maria McSharry, Andrew O. Giacomelli, Seav Huong Ly, Orit Rozenblatt-Rosen, William C. Hahn, Andrew J. Aguirre, Alice H. Berger, Aviv Regev, and Jesse S. Boehm. Massively parallel phenotyping of coding variants in cancer with perturb-seq. 40(6):896–905. Number: 6 Publisher: Nature Publishing Group.

[10] Katherine A. Hoadley, Christina Yau, Toshinori Hinoue, Denise M. Wolf, Alexander J. Lazar, Esther Drill, Ronglai Shen, Alison M. Taylor, Andrew D. Cherniack, Vésteinn Thorsson, Rehan Akbani, Reanne Bowlby, Christopher K. Wong, Maciej Wiznerowicz, Francisco Sanchez-Vega, A. Gordon Robertson, Barbara G. Schneider, Michael S. Lawrence, Houtan Noushmehr, Tathiane M. Malta, Joshua M. Stuart, Christopher C. Benz, and Peter W. Laird. Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. 173(2):291–304.e6.

[11] Benjamin Lengerich, Caleb N. Ellington, Andrea Rubbi, Manolis Kellis, and Eric P. Xing. Contextualized machine learning.

[12] Trevor Hastie and Robert Tibshirani. Varying-coefficient models. 55(4):757–796. Publisher: [Royal Statistical Society, Wiley].

[13] Yang Liu, Nilay S. Sethi, Toshinori Hinoue, Barbara G. Schneider, Andrew D. Cherniack, Francisco Sanchez-Vega, Jose A. Seoane, Farshad Farshidfar, Reanne Bowlby, Mirazul Islam, Jaegil Kim, Walid Chatila, Rehan Akbani, Rupa S. Kanchi, Charles S. Rabkin, Joseph E. Willis, Kenneth K. Wang, Shannon J. McCall, Lopa Mishra, Akinyemi I. Ojesina, Susan Bullman, Chandra Sekhar Pedamallu, Alexander J. Lazar, Ryo Sakai, Vésteinn Thorsson, Adam J. Bass, and Peter W. Laird. Comparative molecular analysis of gastrointestinal adenocarcinomas. 33(4):721–735.e8.

[14] Nicolai Meinshausen and Peter Bühlmann. High-dimensional graphs and variable selection with the lasso. 34(3):1436–1462. Publisher: Institute of Mathematical Statistics.

[15] Maruan Al-Shedivat, Avinava Dubey, and Eric Xing. Contextual explanation networks. 21(194):1–44.

[16] Jie Peng, Pei Wang, Nengfeng Zhou, and Ji Zhu. Partial correlation estimation by joint sparse regression models. 104(486):735–746.

[17] Zeya Wang, Ahmed O. Kaseb, Hesham M. Amin, Manal M. Hassan, Wenyi Wang, and Jeffrey S. Morris. Bayesian edge regression in undirected graphical models to characterize interpatient heterogeneity in cancer. 117(538):533–546.

[18] Travis I. Zack, Stephen E. Schumacher, Scott L. Carter, Andre D. Cherniack, Gordon Saksena, Barbara Tabak, Michael S. Lawrence, Cheng-Zhong Zhsng, Jeremiah Wala, Craig H. Mermel, Carrie Sougnez, Stacey B. Gabriel, Bryan Hernandez, Hui Shen, Peter W. Laird, Gad Getz, Matthew Meyerson, and Rameen Beroukhim. Pan-cancer patterns of somatic copy number alteration. 45(10):1134–1140.

[19] Claudia Calabrese, Natalie R. Davidson, Deniz Demircioğlu, Nuno A. Fonseca, Yao He, André Kahles, Kjong-Van Lehmann, Fenglin Liu, Yuichi Shiraishi, Cameron M. Soulette, Lara Urban, Liliana Greger, Siliang Li, Dongbing Liu, Marc D. Perry, Qian Xiang, Fan Zhang, Junjun Zhang, Peter Bailey, Serap Erkek, Katherine A. Hoadley, Yong Hou, Matthew R. Huska, Helena Kilpinen, Jan O. Korbel, Maximillian G. Marin, Julia Markowski, Tannistha Nandi, Qiang Pan-Hammarström, Chandra Sekhar Pedamallu, Reiner Siebert, Stefan G. Stark, Hong Su, Patrick Tan, Sebastian M. Waszak, Christina Yung, Shida Zhu, Philip Awadalla, Chad J. Creighton, Matthew Meyerson, B. F. Francis Ouellette, Kui Wu, Huanming Yang, Alvis Brazma, Angela N. Brooks, Jonathan Göke, Gunnar Rätsch, Roland F. Schwarz, Oliver Stegle, and Zemin Zhang. Genomic basis for RNA alterations in cancer. 578(7793):129–136. Number: 7793 Publisher: Nature Publishing Group.

[20] Peter Van Loo, Silje H. Nordgard, Ole Christian Lingjærde, Hege G. Russnes, Inga H. Rye, Wei Sun, Victor J. Weigman, Peter Marynen, Anders Zetterberg, Bjørn Naume, Charles M. Perou, Anne-Lise Børresen-Dale, and Vessela N. Kristensen. Allele-specific copy number analysis of tumors. 107(39):16910–16915. Publisher: Proceedings of the National Academy of Sciences.

[21] Ruben M. Drews, Barbara Hernando, Maxime Tarabichi, Kerstin Haase, Tom Lesluyes, Philip S. Smith, Lena Morrill Gavarró, Dominique-Laurent Couturier, Lydia Liu, Michael Schneider, James D. Brenton, Peter Van Loo, Geoff Macintyre, and Florian Markowetz. A pan-cancer compendium of chromosomal instability. 606(7916):976–983. Number: 7916 Publisher: Nature Publishing Group.

[22] Collin Tokheim and Rachel Karchin. CHASMplus reveals the scope of somatic missense mutations driving human cancers. 9(1):9–23.e8.

[23] Christopher D. Steele, Ammal Abbasi, S. M. Ashiqul Islam, Amy L. Bowes, Azhar Khandekar, Kerstin Haase, Shadi Hames-Fathi, Dolapo Ajayi, Annelien Verfaillie, Pawan Dhami, Alex McLatchie, Matt Lechner, Nicholas Light, Adam Shlien, David Malkin, Andrew Feber, Paula Proszek, Tom Lesluyes, Fredrik Mertens, Adrienne M. Flanagan, Maxime Tarabichi, Peter Van Loo, Ludmil B. Alexandrov, and Nischalan Pillay. Signatures of copy number alterations in human cancer. 606(7916):984–991. Number: 7916 Publisher: Nature Publishing Group.

[24] Zbyslaw Sondka, Sally Bamford, Charlotte G. Cole, Sari A. Ward, Ian Dunham, and Simon A. Forbes. The COSMIC cancer gene census: describing genetic dysfunction across all human cancers. 18(11):696–705. Number: 11 Publisher: Nature Publishing Group.

[25] Arthur Liberzon, Chet Birger, Helga Thorvaldsdóttir, Mahmoud Ghandi, Jill P. Mesirov, and Pablo Tamayo. The molecular signatures database (MSigDB) hallmark gene set collection. 1(6):417–425.

[26] Eiru Kim, Veronica Gheorge, and Traver Hart. Dynamic rewiring of biological activity across genotype and lineage revealed by context-dependent functional interactions. Publication Title: bioRxiv.

[27] Shahin Mohammadi, Jose Davila-Velderrain, and Manolis Kellis. Reconstruction of cell-type-specific interactomes at single-cell resolution. 9(6):559–568.e4.

[28] Mladen Kolar, Le Song, Amr Ahmed, and Eric P Xing. Estimating time-varying networks. ISBN: 0812.5087 Publication Title: arXiv [stat.ML].

[29] Benjamin J Lengerich, Maruan Al-Shedivat, Amir Alavi, Jennifer Williams, Sami Labbaki, and Eric P Xing. Discriminative subtyping of lung cancers from histopathology images via contextual deep learning. Publisher: Cold Spring Harbor Laboratory Press.

Supplemental Information

S1 Continued Results

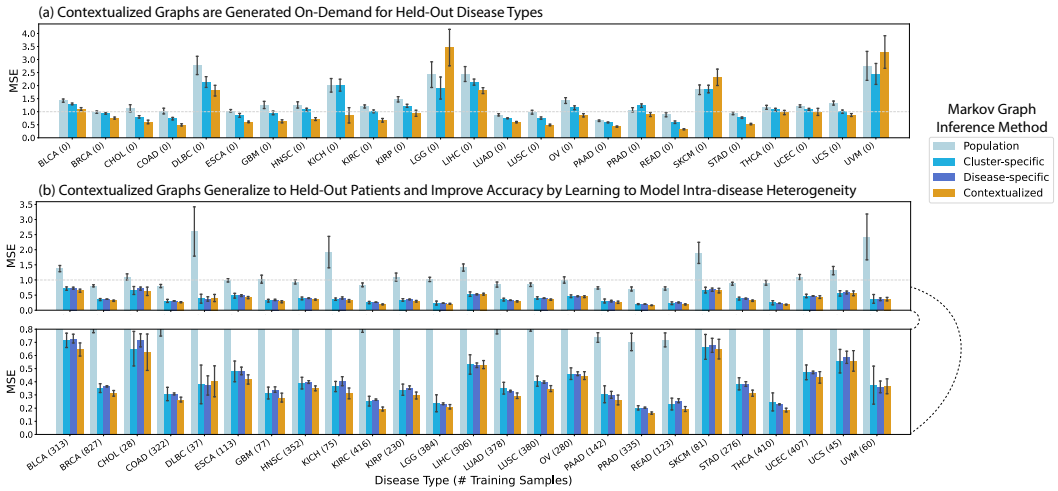


Figure S1: Performance of Contextualized Markov Networks. **(a)** Disease-fold cross-validation, in which each of the 25 disease types are held out from training and evaluated only at testing time. Disease-specific network inference cannot be applied in this regime. **(b)** Testing on held-out patients. Results are from 30 bootstrapped runs for each hold-out disease type and the hold-out patient set. Bar height is the MSE of the bootstrap-averaged network, error bars are the standard deviation of network MSEs over all bootstraps.

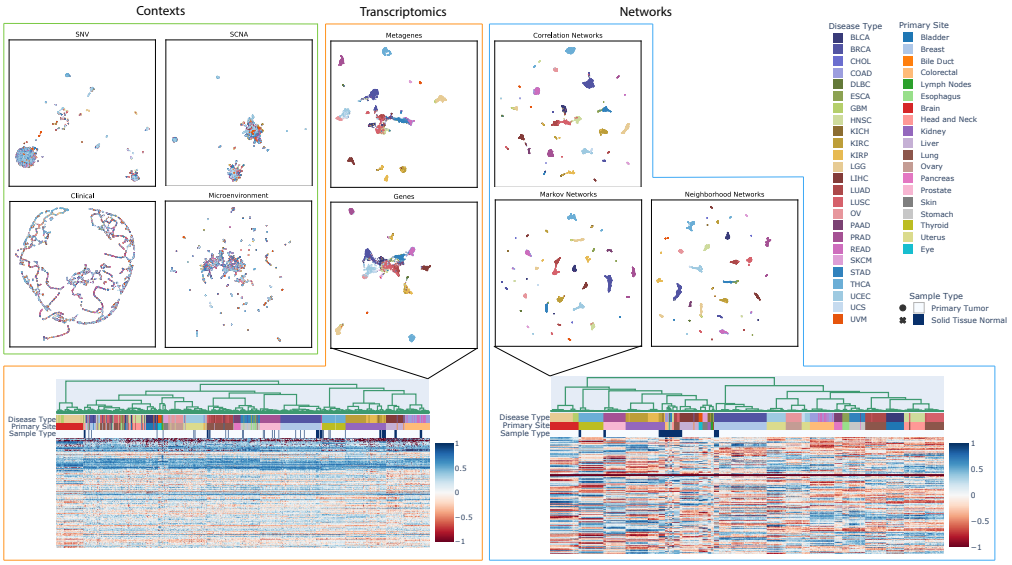
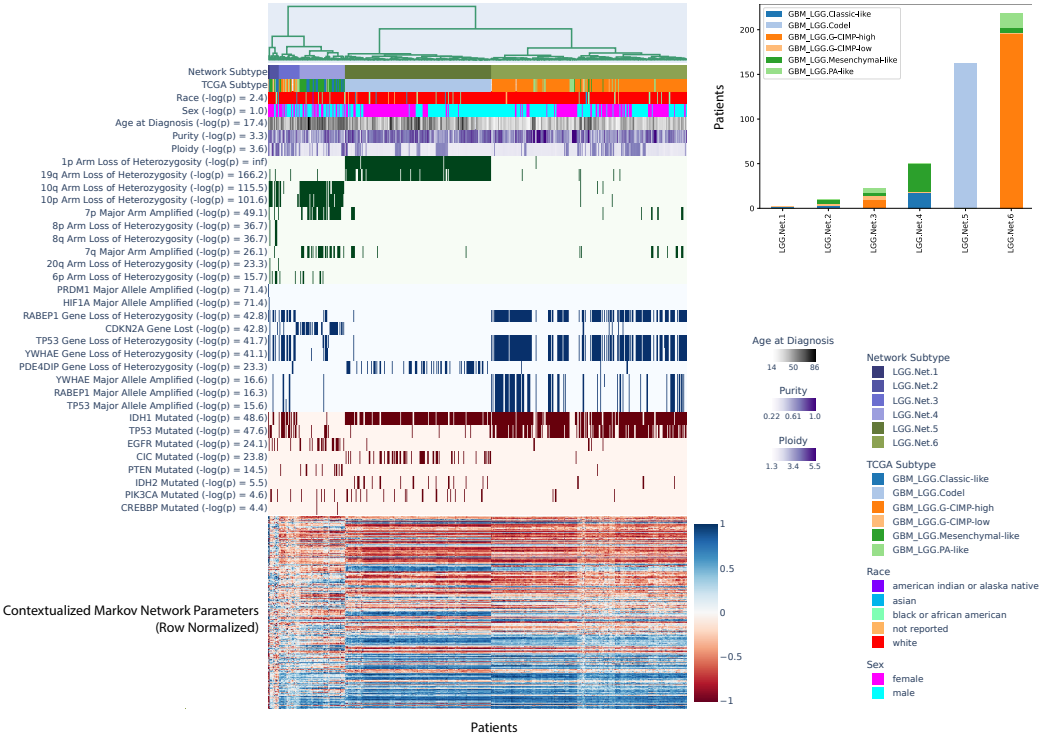


Figure S2: Embeddings reveal the organization of different disease views. Context views alone cannot capture tumor disease types. Transcriptomic views recapitulate disease types, aligning by cell type and tissue of origin. Contextualized networks reveal new separations and similarities, revealing disease subtypes and cross-disease relationships.

Table S1: Average p-values of survival tests for disease subtyping. Multivariate log-rank tests the stratification of survival time across all subtypes. Minimum pairwise tests the maximum survival stratification between all pairs of prognostic subtypes.

	Expression	TCGA	Networks
Multivariate log-rank test (-log(p))	8.53	9.65	11.24
Minimum Pairwise log-rank test (-log(p))	8.27	9.55	11.71

(a) Network Clusters Reveal Prognostic Biomarkers in Molecular Contexts



(b) Clusters of Patient-specific Transcriptomic Networks Reveal State-of-the-art Prognostic Subtypes

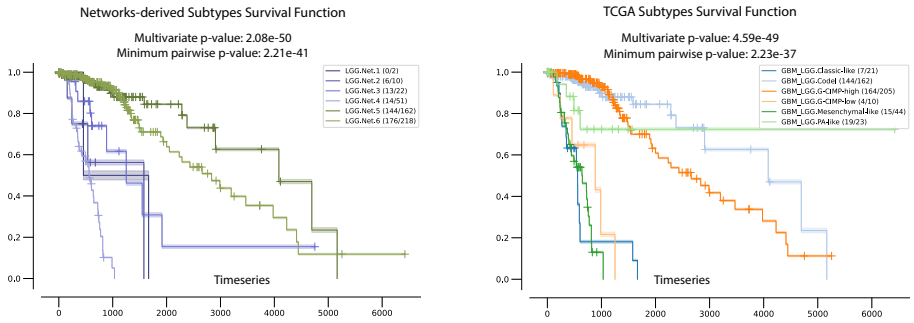
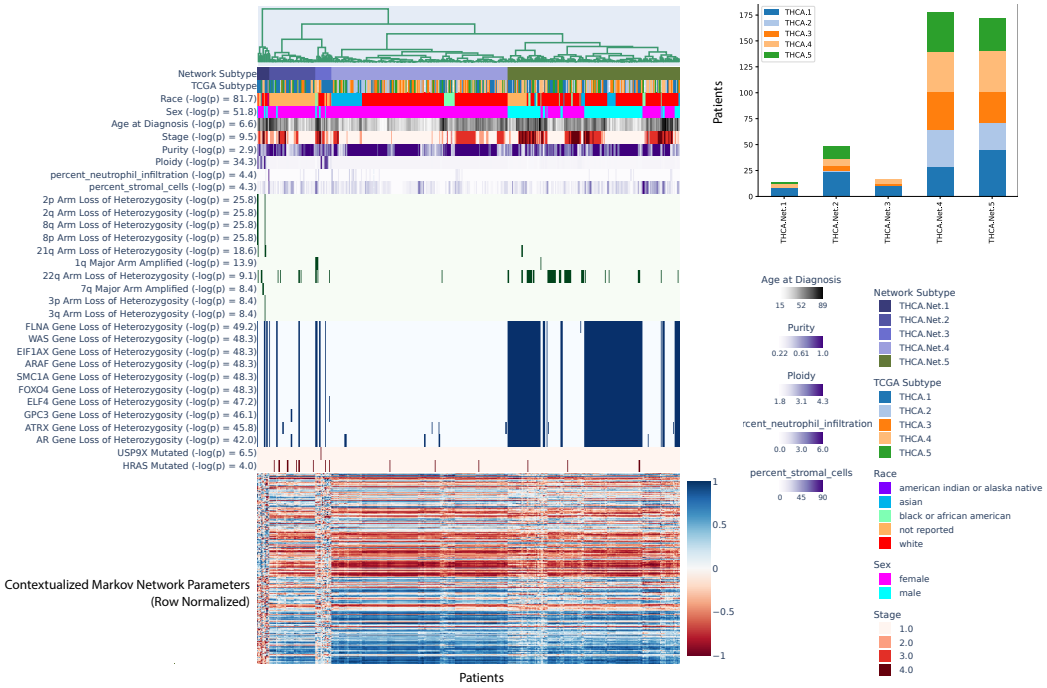


Figure S3: Exploration of network subtypes for LGG, looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.

(a) Network Clusters Reveal Prognostic Biomarkers in Molecular Contexts



(b) Clusters of Patient-specific Transcriptomic Networks Reveal State-of-the-art Prognostic Subtypes

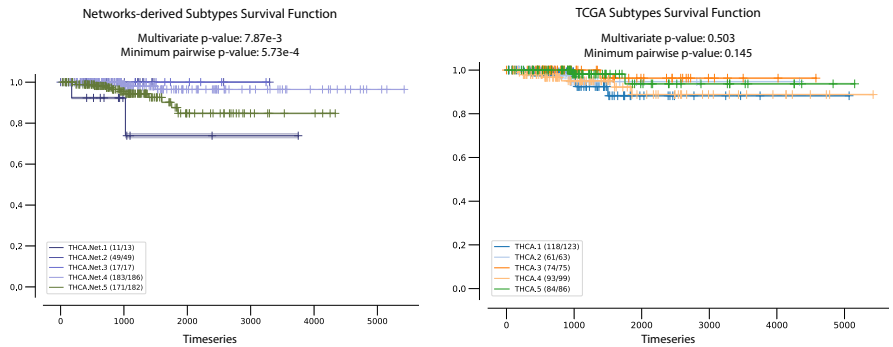


Figure S4: Exploration of network subtypes for THCA, looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.

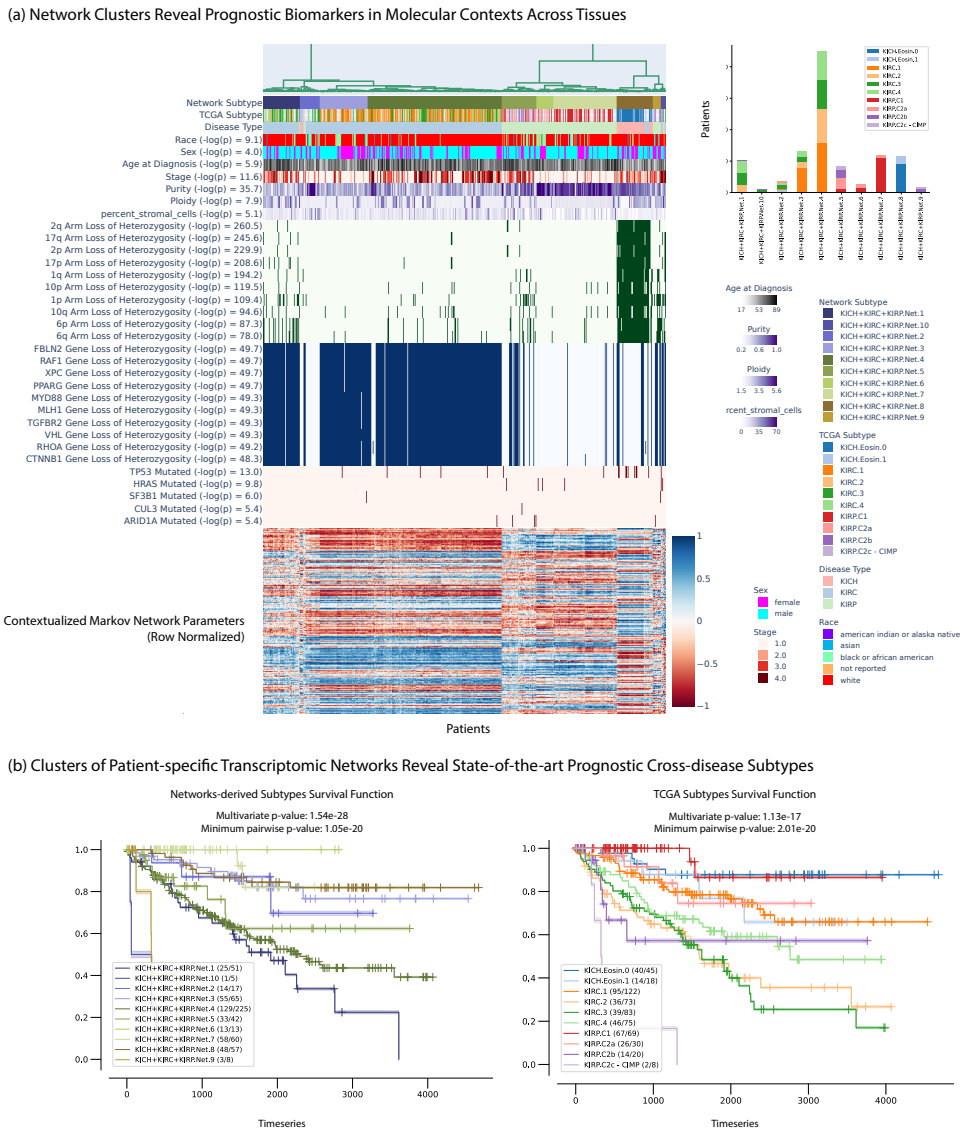


Figure S5: Exploration of cross-disease network subtypes for cancers of the kidneys, including KICH, KIRC, and KIRP, looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.

S2 Method Details

We seek network density $\mathbb{P}(\theta \mid C)$ such that

$$\mathbb{P}(X \mid C) = \int_{\theta} d\theta \mathbb{P}_M(X \mid \theta)\mathbb{P}(\theta \mid C)$$

is maximized, where $\mathbb{P}_N(X \mid \theta)$ is the probability of gene expression $X \in \mathbb{R}^p$ under network model class M with parameters $\theta \in \mathbb{R}^{p \times p}$ and context C, which can contain both multivariate and real features. To overcome θ being a high-dimensional, structured latent variable, we assume that all contextualized networks lie on a subspace spanned by a set of K network archetypes $\mathcal{A} \coloneqq \text{span}(\{A_k \in \mathbb{R}^{p \times p} : A_1, ..., A_K\})$, i.e. $\theta \in \mathcal{A}$. Further, the space spanned by $\mathcal{A}$ is parameterized by a latent variable ("subtype") $Z \in \mathbb{R}^K$ such that Z is a deterministic function of context $Z = f(C)$ and the context-specific network model θ (and subsequently the gene expression observations X) are independent of context given Z, i.e. $C \perp (X, \theta) \mid Z$. In this way, we constrain θ as a convex combination of network archetypes via latent mixing.

$$\begin{aligned}
\mathbb{P}(X \mid C) &= \int_{\theta, Z} d\theta dZ \mathbb{P}_M(X \mid \theta)\mathbb{P}(\theta|Z)\mathbb{P}(Z \mid C) \\
&= \int_{\theta, Z} d\theta dZ \mathbb{P}_M(X \mid \theta)\delta(\theta - \sum_{k=1}^{K} Z_k A_k)\delta(Z - f(C)) \\
&= \mathbb{P}_M(X|\phi(C; f, \mathcal{A})) \\
\phi(C; f, \mathcal{A}) &= \sum_{k=1}^{K} Z_k A_k = \sum_{k=1}^{K} f(C)_k A_k
\end{aligned}$$

Where the context encoder $\phi(C; f, \mathcal{A})$ is parameterized by a learnable context-to-subtype mapping f and the set of archetypes $\mathcal{A}$. This architecture is shown in Figure 1d, and is learned end-to-end with backpropagation. While the archetypal networks A_k could use prior knowledge for initialization or regularization, no prior knowledge is required. In all experiments reported here, we do not use any prior knowledge of network structure or parameters.

This framework unites three different perspectives of GRNs: (1) CNs, in which network edges are the pairwise Pearson's correlation between nodes, (2) MNs, in which edges are the pairwise precision values representing conditional dependencies between nodes, and (3) NNs, in which edges represent directed linear relationships between nodes. The key challenge for each network class is to define a differentiable loss function ℓ_M that is proportional to the negative log probability of our contextualized network model.

$$\begin{aligned}
\widehat{f}, \widehat{\mathcal{A}} &= \underset{f, \mathcal{A}}{\text{argmin}} - \sum_{n=1}^{N} \log\left(\mathbb{P}_M(X_n|\phi(C_n; f, \mathcal{A}))\right) \\
&= \underset{f, \mathcal{A}}{\text{argmin}} \sum_{n=1}^{N} \ell_M\left(\phi(C_n; f, \mathcal{A}), X_n\right)
\end{aligned}$$

The loss objective can be used in the end-to-end optimization, solving for the context encoder and the network archetypes simultaneously, and subsequently inferring the context-specific parameters θ. Below, we derive a unifying linear parameterization of each network loss and discuss implementation details.

Contextualized Neighborhood Regression

We first apply contextualization to the graph variable selection algorithm proposed by Meinhausen and Buhlmann [14]. The direct relationship of this model to lasso regression links contextualized neighborhood regression to original works on contextualized linear models [15], making it a convenient stepping stone toward the graphical models in the sequel. The model is a Gaussian graphical model where $X \sim N(0, \Sigma)$ and Σ has sparse off-diagonal entries. The algorithm, neighborhood regression, recovers edges between nodes with non-zero partial correlations by solving the lasso regression for every feature X_i, given every other feature X_{-i}, where regression maximizes $P(X_i|X_{-i})$ via the loss

$$\widehat{\theta}_i = \underset{\theta}{\text{argmin}} \|X_i - X_{-i}\theta\|_2^2 + \lambda\|\theta\|_1$$

resulting in edges with source X_j for every $j \neq i$ and sink X_i and strength θ_{ij}, or no edge if $\theta_{ij} = 0$. Equivalently, we parameterize the neighborhood selection objective using the square matrix of network edge parameters $\theta \in \mathbb{R}^{p \times p}$.

$$\widehat{\theta} = \underset{\theta}{\operatorname{argmin}} \|X - X\theta\|_F^2 + \lambda \sum_i \|\theta_i\|_1 \text{ s.t. } \operatorname{diag}(\theta) = [0]$$

Where the contextualized neighborhood network objective replaces θ for each sample with a context-specific $\theta_n = \phi(C_n; f, \mathcal{A})$. Finally, we define a function ϕ' to mask the diagonal of θ, presenting the loss function ℓ_{NN} for contextualized neighborhood regression networks

$$\ell_{NN}(\phi(C; f, \mathcal{A}), X) = \|X - X\phi'(C; f, \mathcal{A})\|_2^2 + \lambda \sum_i \|\phi'(C; f, \mathcal{A})_i\|_1$$

$$\phi'(C; f, \mathcal{A}) = (1 - \mathbb{I}) \otimes \phi(C; f, \mathcal{A})$$

where $\otimes$ is the hadamard product.

Contextualized Markov Networks

Linear regression and Gaussian graphical models are intrinsically related, allowing us to extend work on contextualized linear models to various graphical representations of the Gaussian graphical model. To estimate sample-specific precision matrices representing the conditional dependency structure of an undirected graphical model or Markov network, we assume the data is drawn from $X \sim N(0, \Omega^{-1})$ where $\Omega = \Sigma^{-1}$ and estimate pairwise partial correlation coefficients. Using an equivalence defined by Peng et al. [16], the partial correlation coefficient is defined by regression as

$$\rho_{ij} = \operatorname{sign}(\beta_{ij})\sqrt{\beta_{ij}\beta_{ji}} = -\frac{\omega_{ij}}{\omega_{ii}\omega_{jj}} \tag{1}$$

Where the precision matrix Ω has elements ω_{ij} and β is the ordinary least squares solution to multivariate linear regression $\beta_i = \operatorname{argmin}_\beta \|X_i - X_{-i}\beta\|_2^2$. Critically, the precision matrix directly encodes conditional independence between features in X, and thus precision encodes the Markov network.

$$\omega_{ij} = 0 \iff X_i \perp X_j \mid X_{-\{i,j\}}$$

Following [17], we assume constant diagonal precision $w_{ii} = w_{jj} \forall i, j$ and therefore achieve proportionality between the regression and the precision matrix.

$$\omega_{ij} \propto -\operatorname{sign}(\widehat{\beta}_{ij})\sqrt{\widehat{\beta}_{ij}\widehat{\beta}_{ji}}$$

Assuming unit diagonal precisions $\omega_{ii} = 1$, the proportionality becomes exact equivalence. Further, proportionality induces symmetry in the regression, i.e. $\beta_{ij} = \beta_{ji}$. We encode this in the objective by requiring our estimate for θ to be a symmetrically augmented matrix based on γ, i.e. $\beta = \gamma + \gamma^T$

$$\widehat{\gamma} = \underset{\gamma}{\operatorname{argmin}} \|X - X(\gamma + \gamma^T)\|_F^2 \text{ s.t. } \operatorname{diag}(\gamma) = [0]$$

If Ω is sparse, we can apply lasso regularization to the multivariate regression objective [14]. Given the similarity between this differential Markov network objective and the neighborhood regression objective, we follow the exact contextualization procedure from above to contextualize γ and arrive at a loss function ℓ_{MN}

$$\ell_{MN}(\phi(C; f, \mathcal{A}), X) = \|X - X(\phi'(C; f, \mathcal{A}) + \phi'(C; f, \mathcal{A})^T)\|_2^2 + \lambda \sum_i \|\phi'(C; f, \mathcal{A})_i\|_1$$

where ϕ' is defined identically for masking the diagonal. The resulting precision matrix estimate is $\widehat{\Omega} = -(\phi'(C; \widehat{f}, \widehat{\mathcal{A}}) + \phi'(C; \widehat{f}, \widehat{\mathcal{A}})^T)$. In practice we do not threshold the estimated precision to non-zero values, instead using the exact precision matrix to represent the Markov network, retaining information about dependency strength as well as dependency structure in the network.

Contextualized Correlation Networks

Correlation networks are simple to estimate and often state-of-the-art for gene regulatory network inference [4]; contextualized correlation expand this utility to the granularity of sample-specific

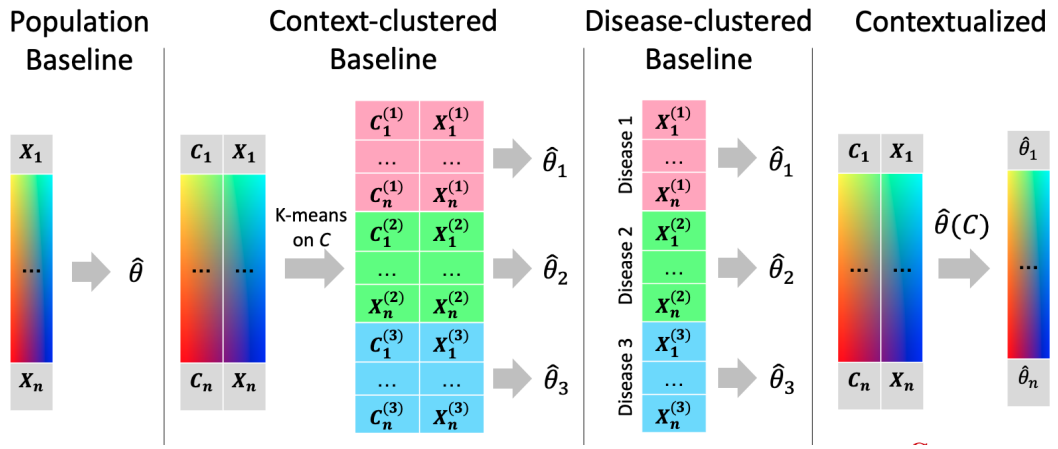


Figure S6: Modeling regimes for personalized inference.

network inferences. To estimate sample-specific correlation networks, we assume the data was drawn from $X \sim N(0, \Sigma)$ and use the well known univariable regression view of Pearson's marginal correlation coefficient:

$$\rho_{ij}^2 = \frac{\sigma_{ij}}{\sigma_{ii}\sigma_{jj}} = \beta_{ij}\beta_{ji} \tag{2}$$

where the covariance matrix Σ has elements σ_{ij}, and $\beta_{ij} = \text{argmin}_\beta (X_j - X_i\beta)^2$. This form converts correlation into two separable univariate least-squares regressions that maximize the marginal conditional probabilities $P(X_i|X_j)$ and $P(X_j|X_i)$. Contextualizing this differentiable objective, we get the contextualized correlation network loss

$$\ell_{CN}(\phi(C; f, \mathcal{A}), X) = \|X - X \otimes \phi(C; f, \mathcal{A})\|_F^2$$

where the context-specific correlation matrix is reconstructed as $\widehat{\rho}^2 = \phi(C; \widehat{f}, \widehat{\mathcal{A}}) \otimes \phi(C; \widehat{f}, \widehat{\mathcal{A}})^T$.

S2.1 Baselines

We compare contextualized modeling with several traditional approaches for context-controlled and context-agnostic inference, including population modeling, cluster modeling, and cohort modeling (Fig. S6). A population model assumes that the entire sample population is identically distributed. As a result, population modeling infers a single model representing all observations. In reality, sample populations often contains two or more uniquely distributed subpopulations. If we expect that there are several subpopulations with many observations each, and that these subpopulations can be stratified by context, it may be appropriate to cluster the data by context to identify these subpopulations and then infer a model for each context-clustered subpopulation. This assumes that all context features are equally important and therefore does not tolerate noise features well. Alternatively, when subpopulation groupings are known to be determined by a few important features, cohort modeling is more appropriate. Sample cohorts can be identified based on prior knowledge about important context features (e.g. disease type).

The baseline modeling regimes enjoy the benefits of traditional inference methods (i.e. identifiability, convergence) by relying on the assumption there are a discrete number of subpopulations underlying the observed data that are each defined by a latent model, and each of these subpopulations is well-sampled. This assumption is rarely, if ever, satisfied in a real-world setting. We develop contextualized modeling as a synthesis between traditional statistical inference and modern deep learning to enable model-based analysis of heterogeneous real data. Contextualized modeling assumes a functional dependency between models, but unlike prior methods makes no assumption about the form or complexity of this dependency. As such, contextualized models permit context-informed inference even when contexts are sparsely sampled and high dimensional.

Our entire framework (Fig. 1) is implemented in PyTorch using the PyTorch Lightning framework within our open-source software ContextualizedML. The context encoder, network archetypes, and

contextualized network models are learned simultaneously using end-to-end backpropagation of the network loss (defined in Section S2).

S2.2 Context Encoder & Training

The context encoder is implemented as a multi-layer perceptron with 3 hidden layers, each 100 neurons wide with ReLU activations. The context data views (S3.2) are concatenated sample-wise to create a single context feature vector encompassing all views for each patient. We use a batch-size of 10 and our learning rate is chosen automatically using PyTorch Lightning's `auto-lr-find` with an initial state of 1e-3. Model weights are initialized as `Uniform[-0.01, 0.01]`. We split our dataset into 80% training-validation and 20% testing. We create 30 bootstraps of the training-validation set and finally split into 80% training and 20% validation, resulting in a 64-16-20 split for train-validation-test where the train and validation sets are bootstrapped to evaluate model variance. We use early-stopping with a patience of 5 to end training when the minimum validation loss has not been improved for 5 epochs. We retain only the model with the minimum validation loss for each bootstrap. In Table 1, we evaluate these bootstraps individually to get error means and variances. Following this, we apply each of our bootstrapped models to the non-bootstrapped training-validation set and average the outputs of each model to obtain a single graph for each patient in this set, which we evaluate in-depth in Figures S3, S4, 2, and S5, and for all disease types in Appendix S5.

The context encoder is a highly flexible component of our framework and a driving force for future work. This attribute can be used to enforce assumptions about the relationships between contexts and models, between context features, and about the archetype space. For instance, by using a neural additive model instead of a multi-layer perceptron, we provide context-feature-specific archetype weights for interpretability. Similarly, we can augment our context encoder with a convolutional base and include imaging modalities in our context views. At the context encoder head, we currently use an unconstrained output, but applying a softmax activation would require all of the sample-specific models to lie within a polytope defined by the archetypal networks.

S3 Data

The TCGA data used is public and available for download via the Genomic Data Commons Data Portal.

S3.1 Data sources

The Cancer Genome Atlas[3] (TCGA) is a publicly-available pan-cancer datasource containing genomic, transcriptomic, and clinical profiling of tumors from dozens of landmark studies. We queried TCGA for samples with bulk RNA-sequencing and merged this dataset with two follow-up studies on an overlapping set of patients.

Somatic copy number alterations (SCNAs) SCNAs affect a larger fraction of the genome than do any other type of somatic genetic alteration [18] and are a major driver of expression variation in cancer [19]. We used copy number profiles derived from TCGA samples using ASCAT [20] from a pan-cancer study of the role of allele-specific SCNAs in cancer [21].

Driver single-nucleotide mutations (SNVs) SNVs can be classified into "driver" mutations thought to provide selective growth advantage and "passenger" mutations thought to have little role in promoting cancer development. We incorporated driver SNVs from the TCGA-derived CHASMplus dataset [22]

S3.2 Context data views

Clinical information This data view incorporates sample tissue-of-origin, race, age at diagnosis, gender, year of birth, and days to collection provided by TCGA.

[3] `www.cancer.gov/tcga`

Biopsy Composition This data view contains the sample's percent tumor cells, percent normal cells, percent tumor nuclei, percent monocyte infiltration, percent lymphocyte infiltration, and percent neutrophil infiltration provided by TCGA. We also incorporate expression-derived estimates of the fraction of a sample consisting of tumour cells from [21].

Copy Number Alterations From ASCAT [20], we gather whole genome doubling events as well as gain and loss events for bp-specific regions of hg19 based on data from [23]. We transform these gain and loss events into both arm-level and gene-level events, where arm-level events affect 85% of an entire arm in the same event, while genes-level events affect a single gene. We transform these into number of major and number of minor chromosome arms, and the number of major and minor alleles for the set of 295 genes that overlap between COSMIC [24] and MSigDB [25]. For both gene and arm-level events, we create a separate indicator for loss of heterozygosity on each gene.

Driver Mutations From CHASMplus [22] we gather the mutations on all COSMIC [24] oncogenes/tumor suppresor genes and binarize the presence or absence of a mutation in each gene.

S3.3 Transcriptomic data views

Transcriptomics We take the set of known oncogenes/tumor suppressor genes annotated in COSMIC [24] and included in TCGA gene expression panels. We then calculate the variance of each gene in each tumor type and take a weighted sum of these variances according to the total number of samples in each tumor type. We select the top 100 genes by this metric of "intra-disease variance".

Baselines We are not aware of any other scalable meta-learning, deep learning, or varying-coefficient methods to produce context-informed correlation, Markov, and Bayesian networks under a universal framework. As such, our baselines apply the network estimators in S2 under several well-known and general paradigms for improving model personalization, broadly relating to cluster analysis. Our population baseline provides no personalization, learning a single model for the entire population of training samples. Our context-clustered baseline takes an unsupervised approach to personalization by first doing a k-means clustering with k=25 on the aggregated context views (S3) and then inferring cluster-specific networks. Our disease-clustered baseline uses a personalization oracle, grouping samples by tumor type and then inferring disease-specific networks.

S4 Related work

State-of-the-art gene regulatory network estimators are limited to population, cohort, and cluster-based approaches [26, 27, 3]. Other proposals to estimate networks as the difference between a population model and a sample-left-out model lack statistical power [7]. Kolar et. al achieve sample-specific network estimation without sacrificing statistical power by using an approach similar to classic varying-coefficient models that weighs samples by their distance over context [28]. However, this approach inherently assumes smoothness of the parameters over a context, which does not align with our understanding of the non-linear, switch-like changes in biological systems that lead to disease. Contextual estimation networks (CENs) remove this smoothness assumption by inferring the relationship between context and model parameters with a neural network, but the CEN framework is only proposed as an adaptive learning approach for linear models [15]. Context-varying linear models have previously been applied to multi-omic cancer data, where context-varying coefficients inform how epigenetic markers have patient-specific effects on clinical outcomes [29]. Linear models do not inform us of the differential gene-gene interactions that explain changes in cellular behavior. To understand regulatory and metabolic variation at per-sample resolution, we require network models with context-varying structures and parameters.

S5 Extra Results

Table S2: Multivariate log-rank test comparison across different subtyping methods in terms of -log(p-value). Only samples shared between all datasets are used. – indicates no samples are shared, or subtypes do not exist for TCGA.

Disease Type	CoCA Subtypes	Expression Subtypes	Network Subtypes
BLCA	0.411	0.151	0.713
BRCA	1.484	0.616	1.558
CHOL	–	–	–
COAD	0.016	0.014	1.219
DLBC	–	–	–
ESCA	0.044	0.884	0.049
GBM	1.644	0.958	0.101
HNSC	1.209	0.312	3.465
KICH	0.715	13.802	0.211
KIRC	5.042	6.109	13.741
KIRP	14.538	10.582	15.205
LGG	48.338	33.438	49.681
LIHC	0.009	0.427	0.827
LUAD	0.687	1.172	0.507
LUSC	0.123	0.105	0.249
OV	0.704	0.684	0.05
PAAD	0.439	1.104	1.494
PRAD	–	–	–
READ	0.221	0.203	0.117
SKCM	–	–	–
STAD	0.044	1.117	0.575
THCA	0.298	0.164	2.104
UCEC	6.937	3.343	7.07
UCS	0.319	0.023	0.048
UVM	4.838	2.589	3.565

Table S3: Minimum pairwise log-rank test comparison across different subtyping methods in terms of -log(p-value). Only samples shared between all datasets are used. – indicates no samples are shared, or subtypes do not exist for TCGA.

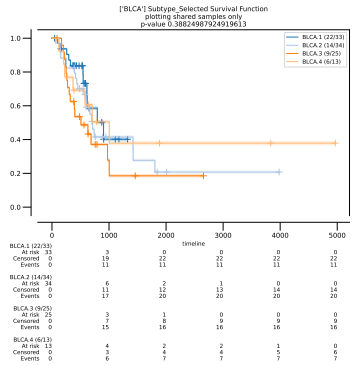
Disease Type	CoCA Subtypes	Expression Subtypes	Network Subtypes
BLCA	1.059	0.589	0.949
BRCA	2.056	1.13	2.542
CHOL	–	–	–
COAD	0.162	0.159	2.026
DLBC	–	–	–
ESCA	0.323	1.598	0.284
GBM	1.631	1.509	0.63
HNSC	1.855	0.853	3.307
KICH	0.715	13.802	0.211
KIRC	5.61	5.247	14.82
KIRP	19.696	9.241	18.661
LGG	36.533	25.894	40.656
LIHC	0.105	0.619	1.263
LUAD	1.67	2.29	1.198
LUSC	0.614	0.417	0.625
OV	1.414	1.325	0.287
PAAD	0.937	1.465	2.18
PRAD	–	–	–
READ	0.431	0.723	0.474
SKCM	–	–	–
STAD	0.469	1.777	1.428
THCA	0.837	0.831	3.242
UCEC	5.555	3.319	7.42
UCS	0.319	0.023	0.048
UVM	5.076	2.61	4.536

Table S4: Multivariate log-rank test comparison across different subtyping methods in terms of -log(p-value). All samples are used, even if they do not have a given TCGA subtype or were removed from our dataset during preprocessing.

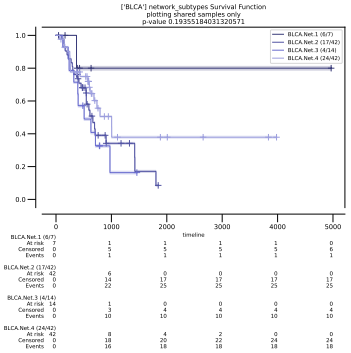
Disease Type	CoCA Subtypes	Expression Subtypes	Network Subtypes
BLCA	0.312	1.649	0.817
BRCA	5.262	0.616	1.558
CHOL	–	0.402	0.306
COAD	0.013	0.01	0.851
DLBC	–	0.148	0.326
ESCA	0.034	0.843	0.192
GBM	1.644	0.362	0.123
HNSC	1.516	0.807	1.781
KICH	0.713	13.802	0.211
KIRC	5.567	7.761	3.989
KIRP	14.004	6.221	11.002
LGG	48.338	33.891	50.065
LIHC	0.077	0.315	2.466
LUAD	0.874	1.41	1.556
LUSC	0.075	0.223	0.035
OV	0.704	0.76	0.076
PAAD	0.429	3.004	2.144
PRAD	0.554	0.23	0.084
READ	0.221	0.075	0.703
SKCM	0.086	0.473	0.129
STAD	0.128	1.015	0.519
THCA	0.3	0.182	2.197
UCEC	7.128	3.356	6.944
UCS	0.319	0.023	0.048
UVM	4.838	2.589	3.565

Table S5: Minimum pairwise log-rank test comparison across different subtyping methods in terms of -log(p-value). All samples are used, even if they do not have a given TCGA subtype or were removed from our dataset during preprocessing.

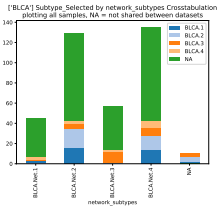
Disease Type	CoCA Subtypes	Expression Subtypes	Network Subtypes
BLCA	0.872	2.736	1.554
BRCA	6.94	1.13	2.542
CHOL	–	0.402	0.306
COAD	0.171	0.201	1.608
DLBC	–	0.148	0.326
ESCA	0.331	1.419	0.546
GBM	1.631	1.141	0.582
HNSC	2.597	1.382	2.675
KICH	0.713	13.802	0.211
KIRC	6.239	6.073	3.193
KIRP	21.38	6.375	11.705
LGG	36.533	26.156	41.199
LIHC	0.194	0.67	3.025
LUAD	1.854	3.042	2.702
LUSC	0.487	0.6	0.22
OV	1.414	1.452	0.345
PAAD	0.914	3.215	2.372
PRAD	1.718	0.54	0.084
READ	0.431	0.43	0.866
SKCM	1.135	0.473	0.129
STAD	0.737	1.739	1.296
THCA	0.812	0.843	3.426
UCEC	5.629	2.904	7.273
UCS	0.319	0.023	0.048
UVM	5.076	2.61	4.536



(a) Known subtypes survival function



(b) Network subtypes survival function



(c) Network subtypes crosstabulation with known subtype

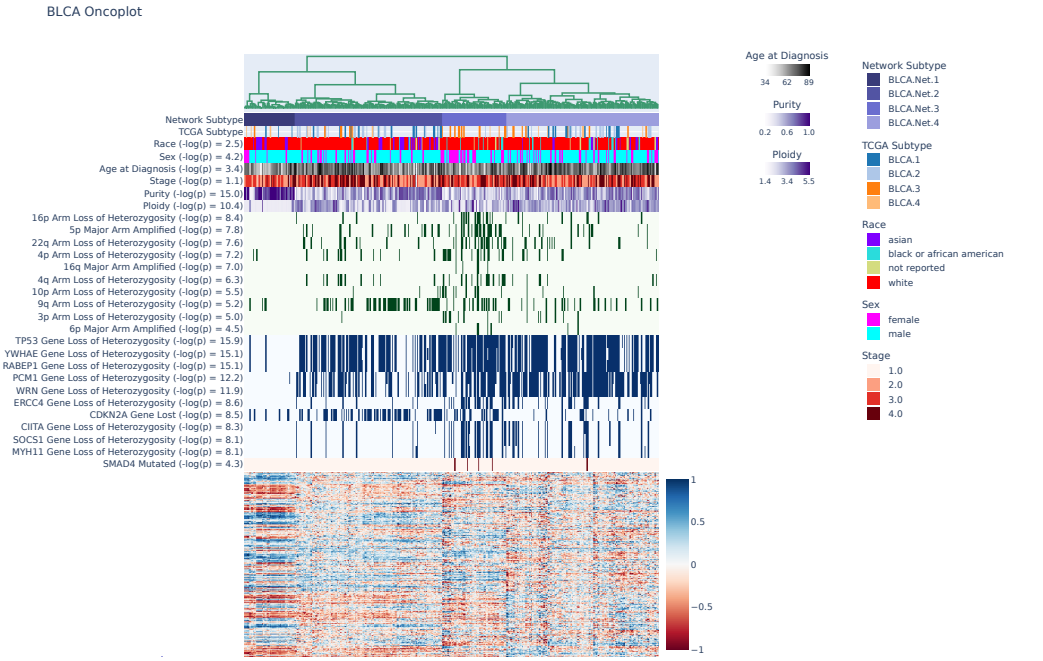
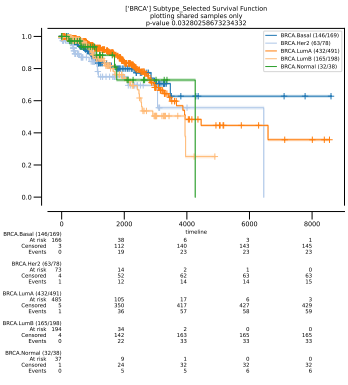
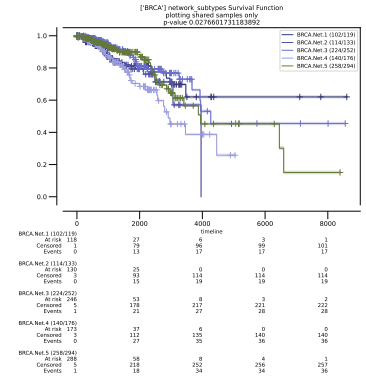


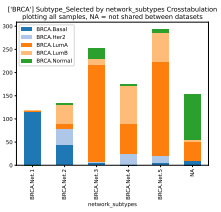
Figure S7: Exploration of network subtypes for Bladder Urothelial Carcinoma (BLCA), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function



(b) Network subtypes survival function



(c) Network subtypes crosstabulation with known subtype

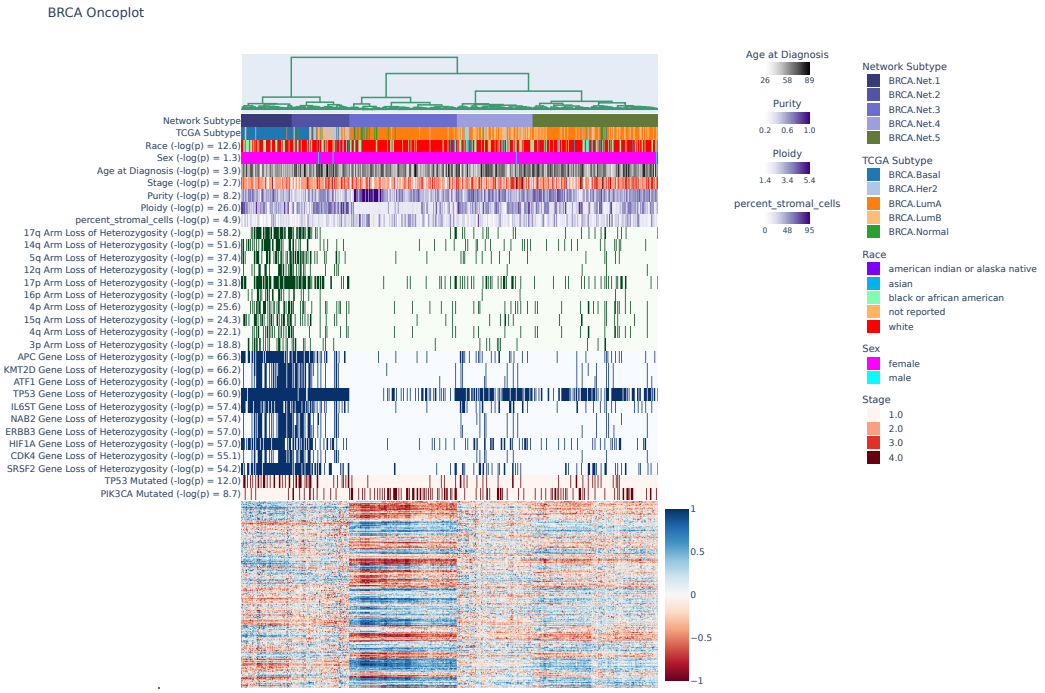
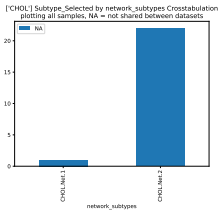


Figure S8: Exploration of network subtypes for Breast invasive carcinoma (BRCA), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Network subtypes crosstabulation with known subtype

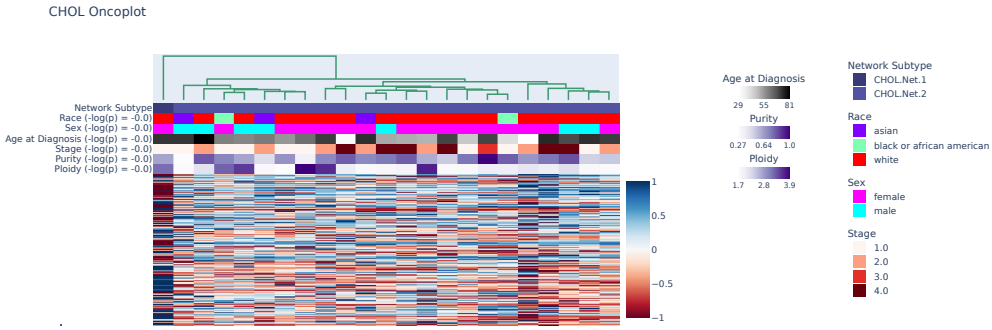
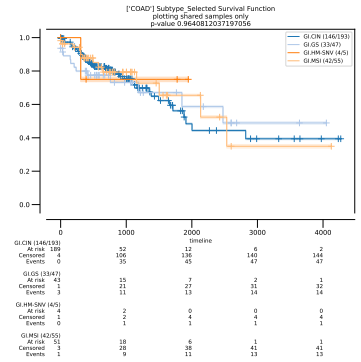
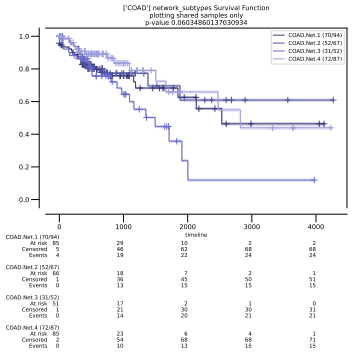


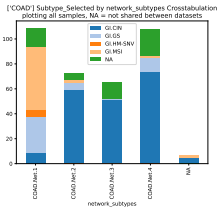
Figure S9: Exploration of network subtypes for Cholangiocarcinoma (CHOL), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function



(b) Network subtypes survival function



(c) Network subtypes crosstabulation with known subtype

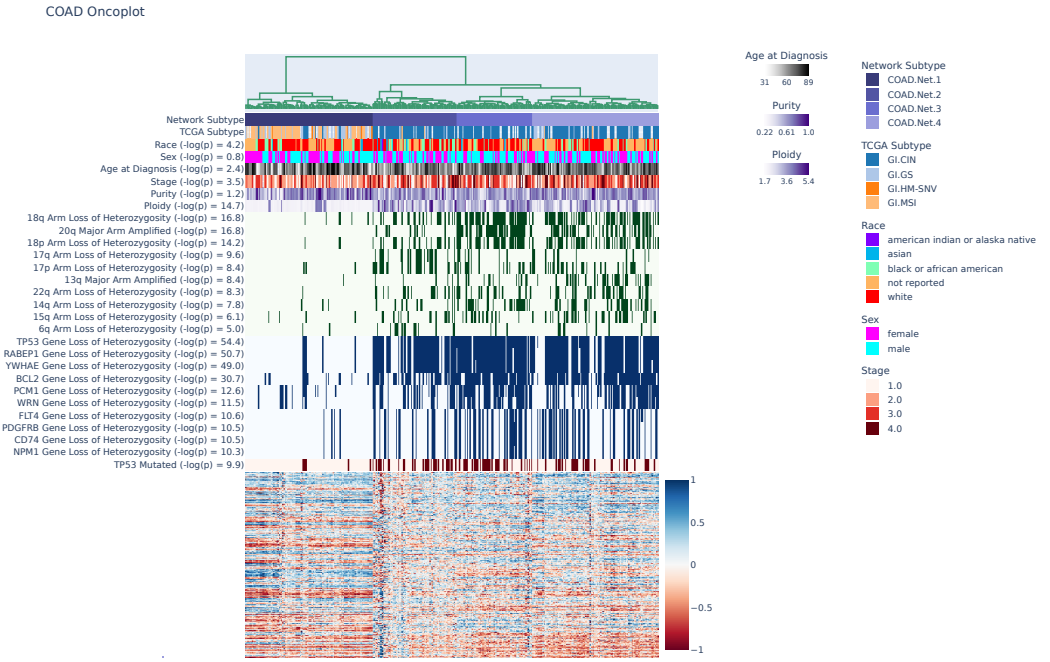
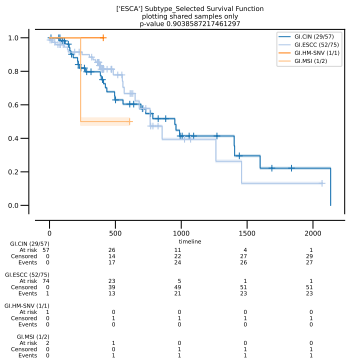
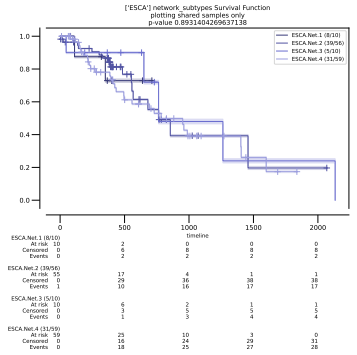


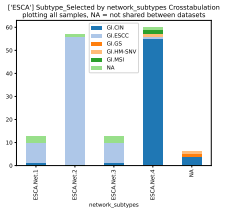
Figure S10: Exploration of network subtypes for Colon adenocarcinoma (COAD), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function



(b) Network subtypes survival function



(c) Network subtypes crosstabulation with known subtype

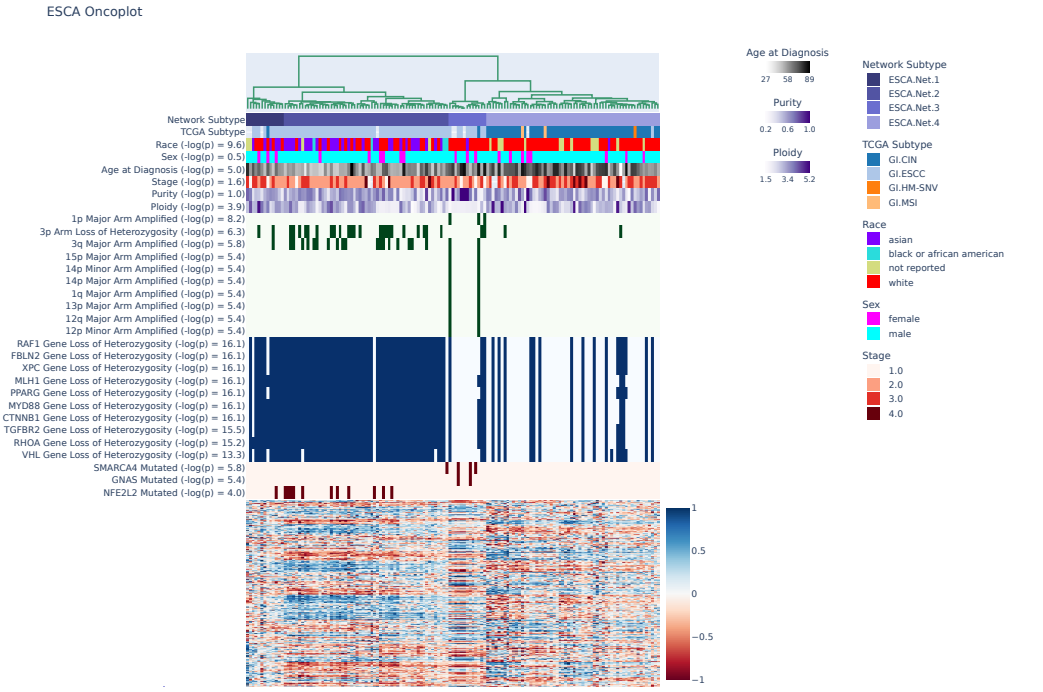
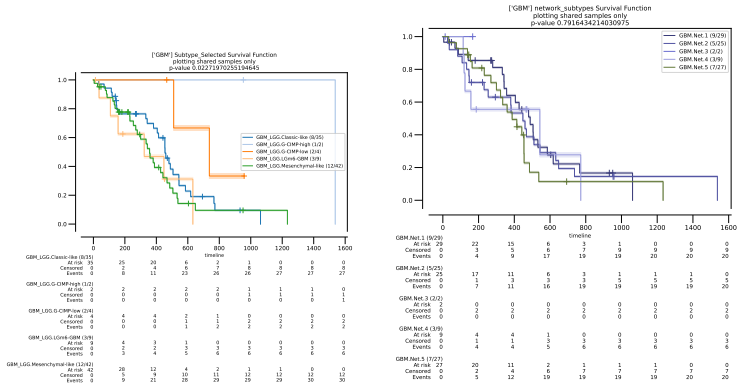
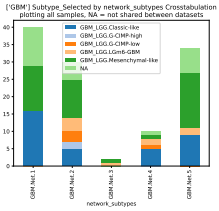


Figure S11: Exploration of network subtypes for Esophageal carcinoma (ESCA), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function

(b) Network subtypes survival function



(c) Network subtypes crosstabulation with known subtype

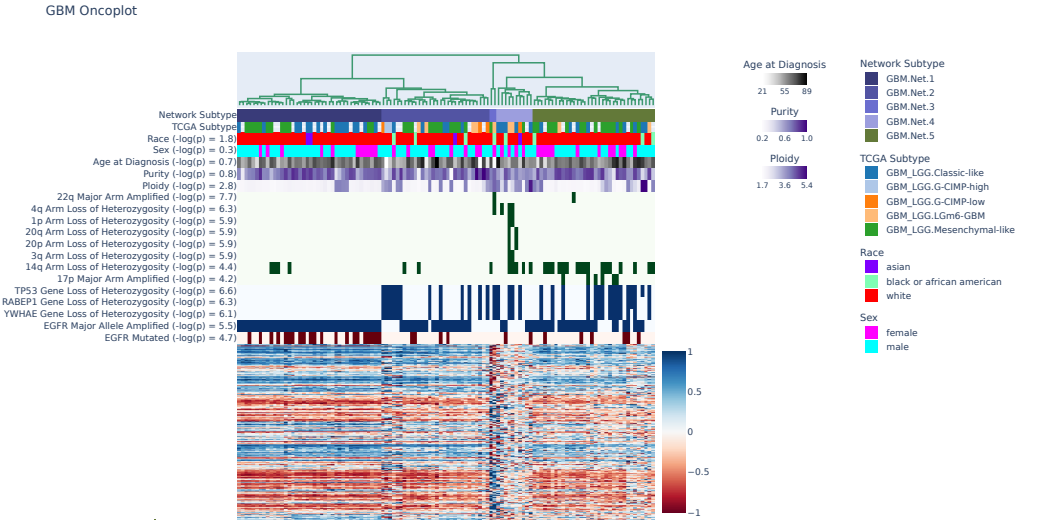
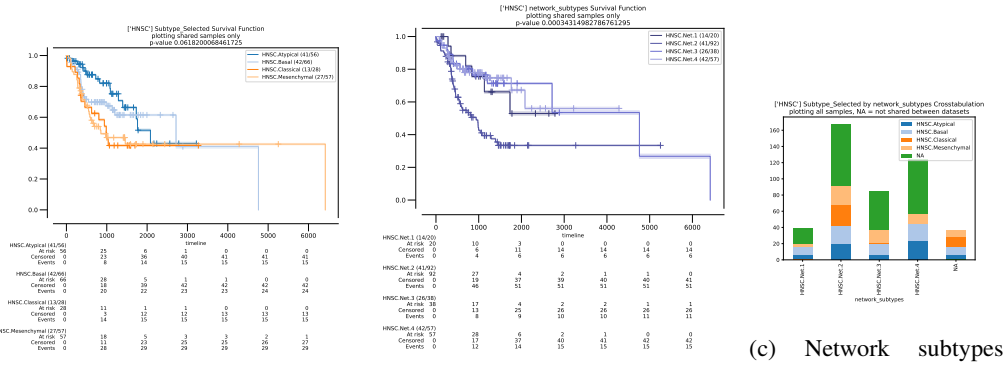


Figure S12: Exploration of network subtypes for Glioblastoma multiforme (GBM), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function (b) Network subtypes survival function (c) Network subtypes crosstabulation with known subtype

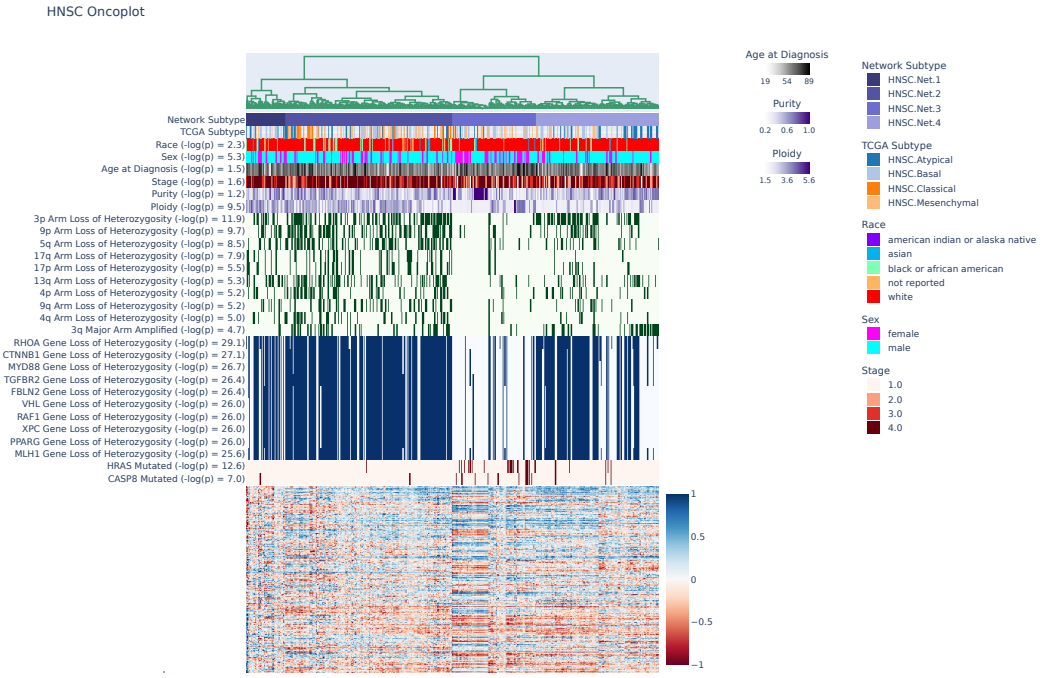
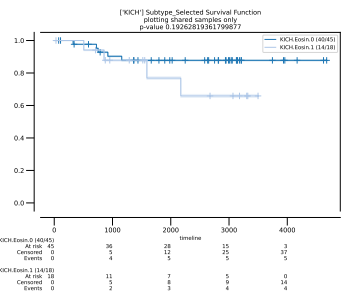
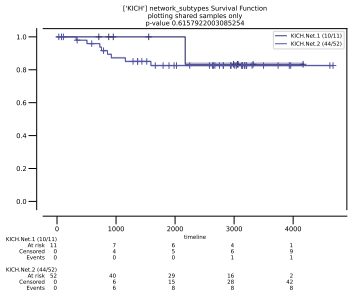


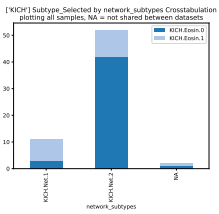
Figure S13: Exploration of network subtypes for Head and Neck squamous cell carcinoma (HNSC), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function



(b) Network subtypes survival function



(c) Network subtypes crosstabulation with known subtype

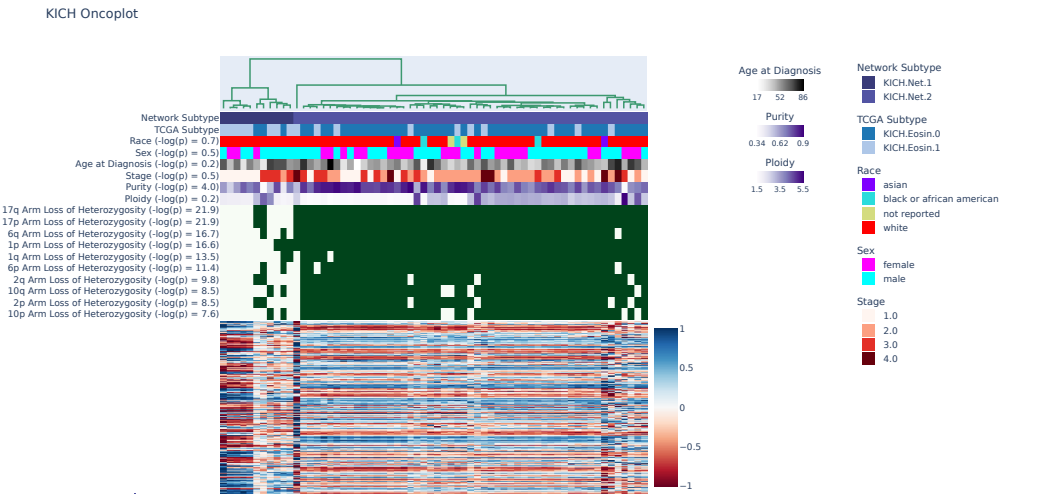
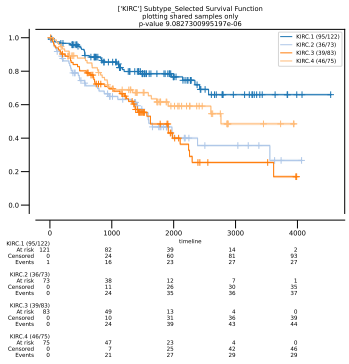
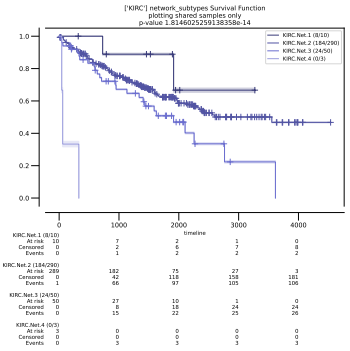


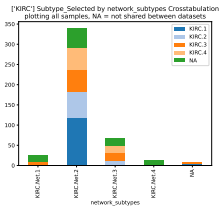
Figure S14: Exploration of network subtypes for Kidney Chromophobe (KICH), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function



(b) Network subtypes survival function



(c) Network subtypes crosstabulation with known subtype

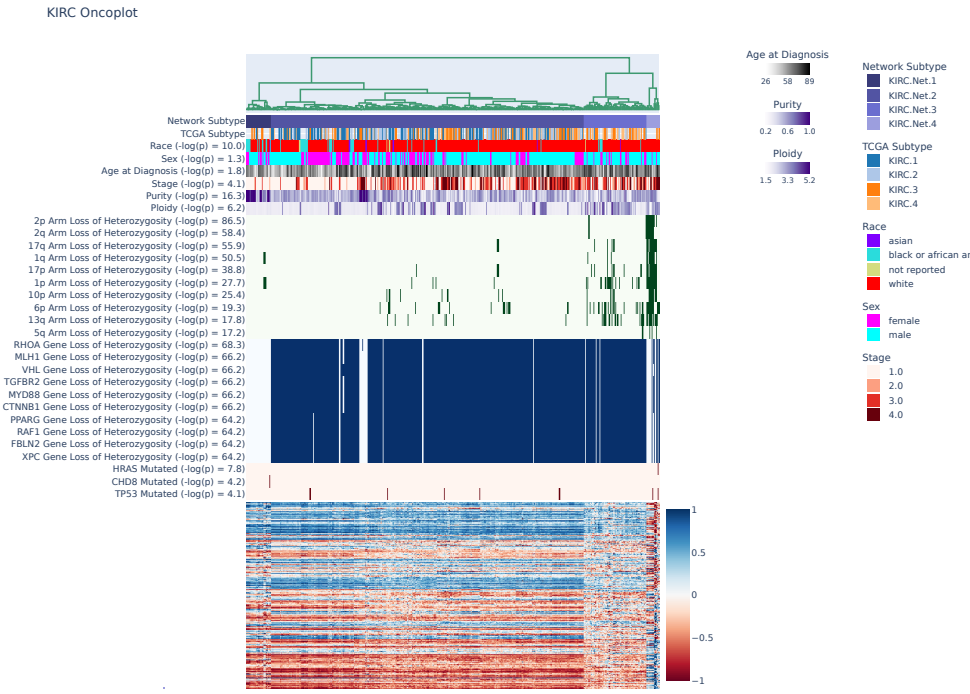
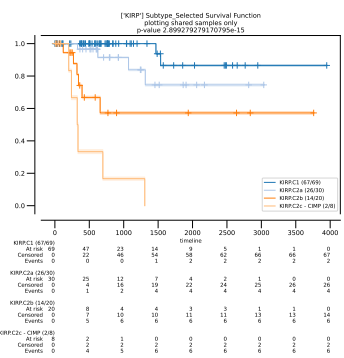
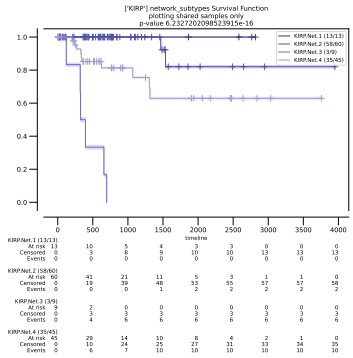


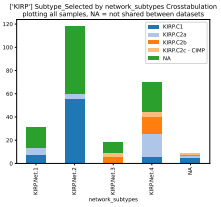
Figure S15: Exploration of network subtypes for Kidney renal clear cell carcinoma (KIRC), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function



(b) Network subtypes survival function



(c) Network subtypes crosstabulation with known subtype

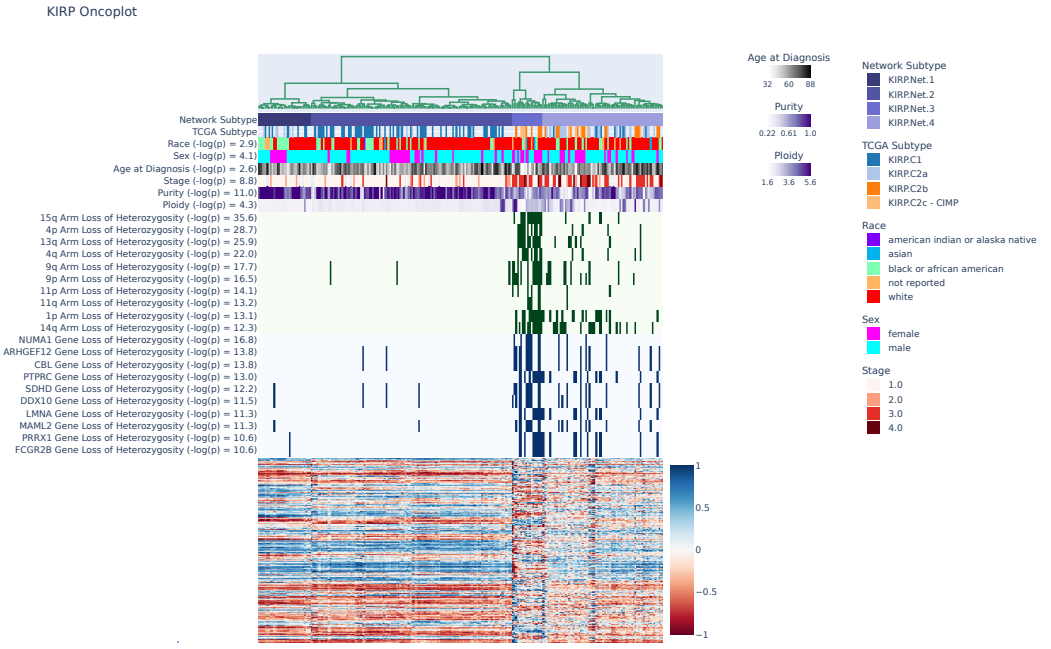
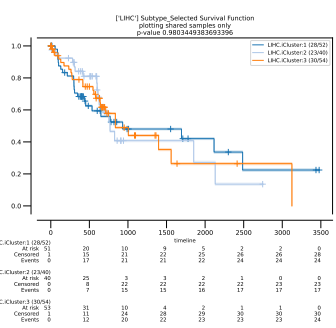
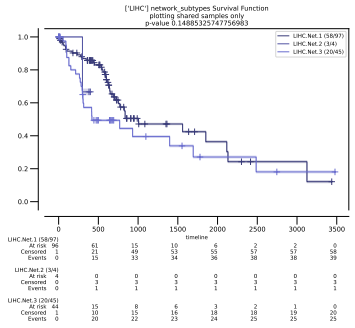


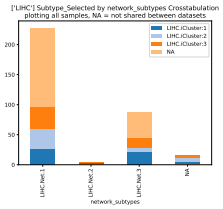
Figure S16: Exploration of network subtypes for Kidney renal papillary cell carcinoma (KIRP), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function



(b) Network subtypes survival function



(c) Network subtypes crosstabulation with known subtype

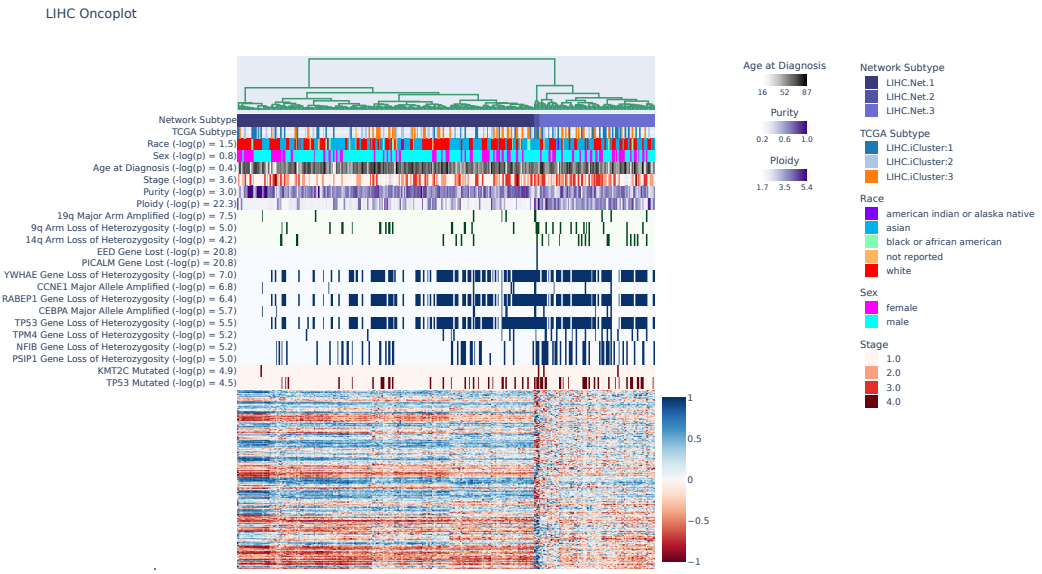
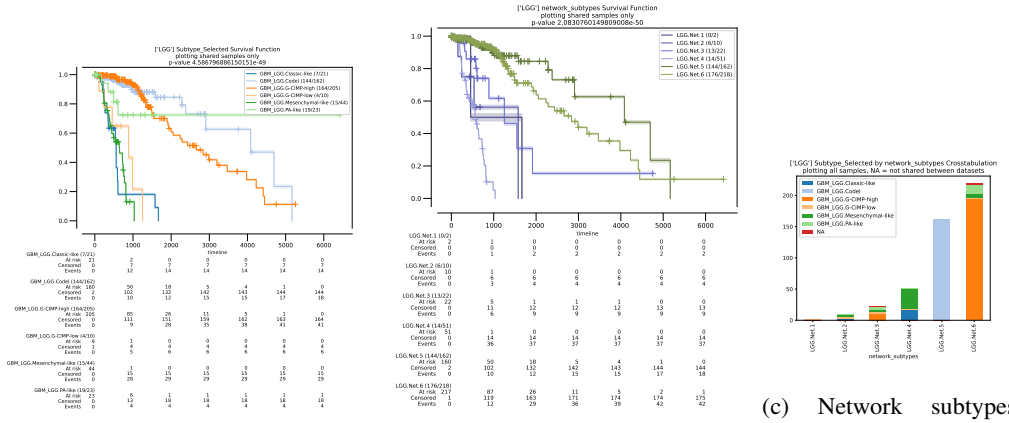


Figure S17: Exploration of network subtypes for Liver hepatocellular carcinoma (LIHC), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function (b) Network subtypes survival function (c) Network subtypes crosstabulation with known subtype

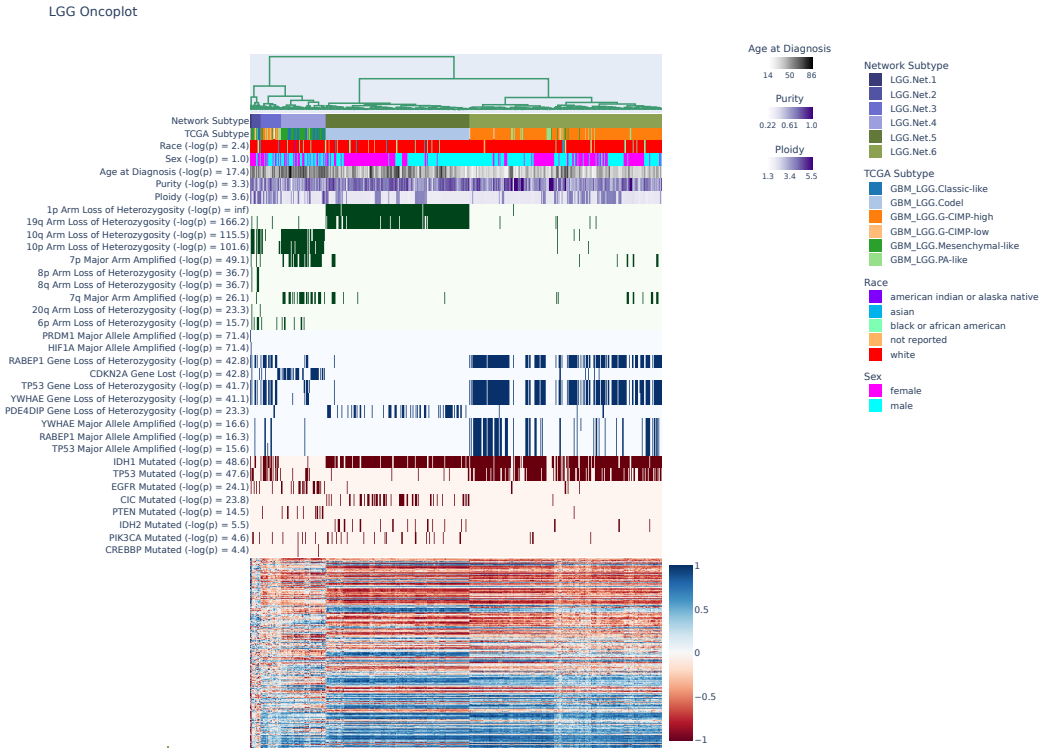
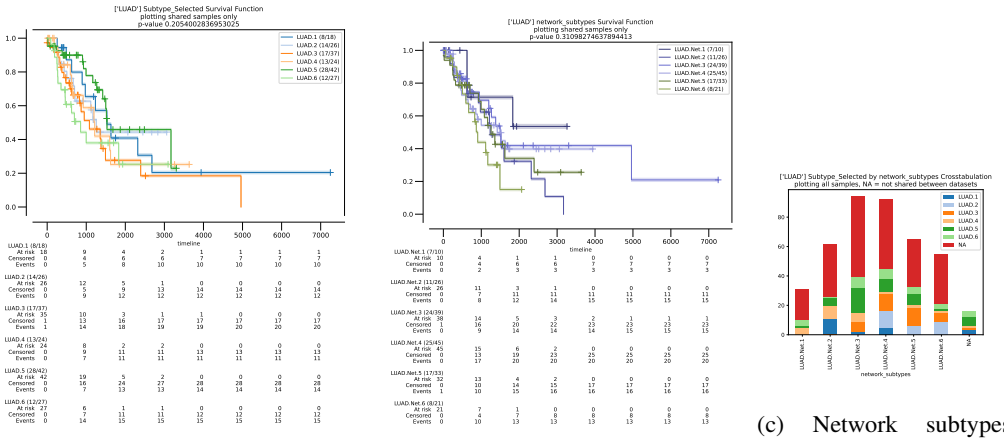


Figure S18: Exploration of network subtypes for Brain Lower Grade Glioma (LGG), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function (b) Network subtypes survival function (c) Network subtypes crosstabulation with known subtype

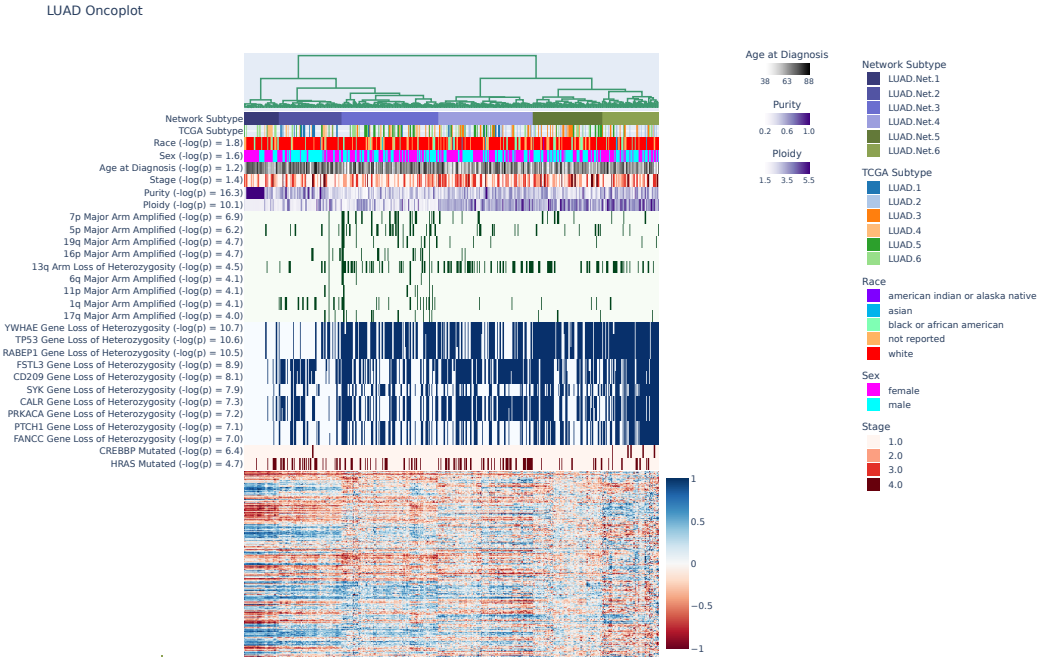
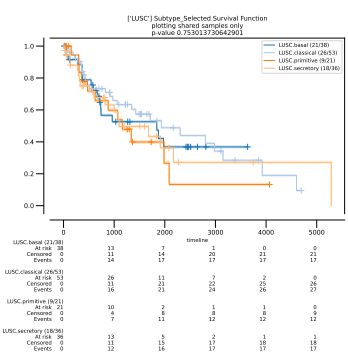
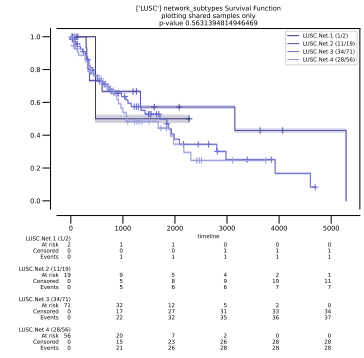


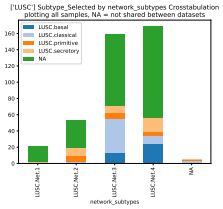
Figure S19: Exploration of network subtypes for Lung adenocarcinoma (LUAD), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function



(b) Network subtypes survival function



(c) Network subtypes crosstabulation with known subtype

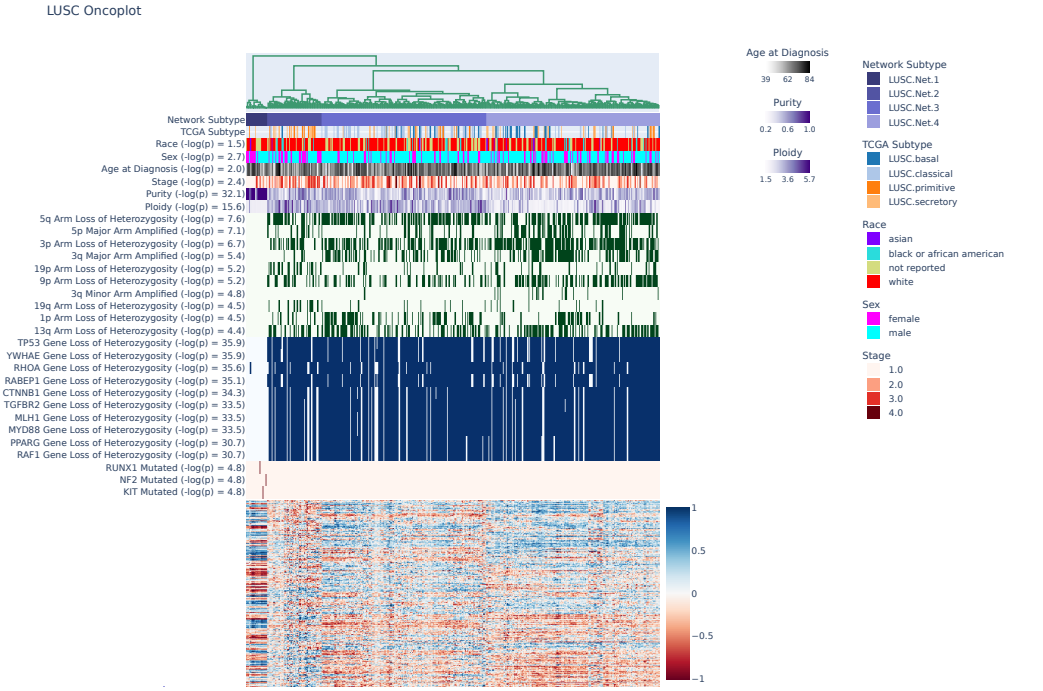
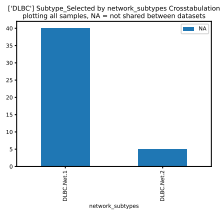


Figure S20: Exploration of network subtypes for Lung squamous cell carcinoma (LUSC), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Network subtypes crosstabulation with known subtype

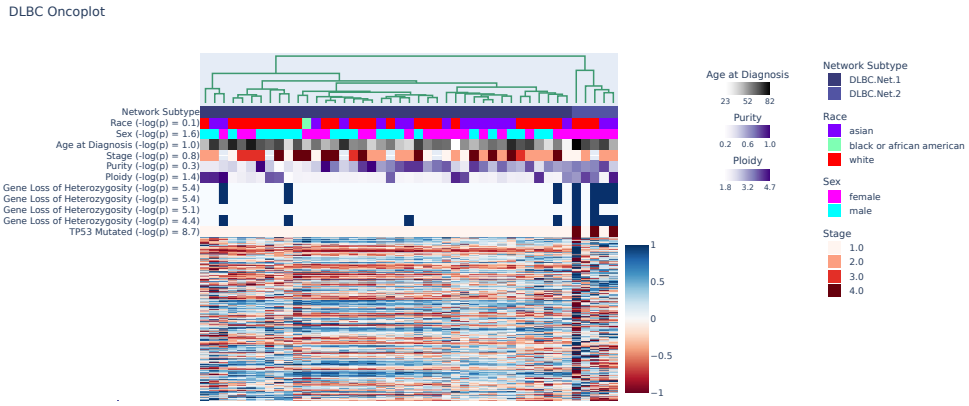
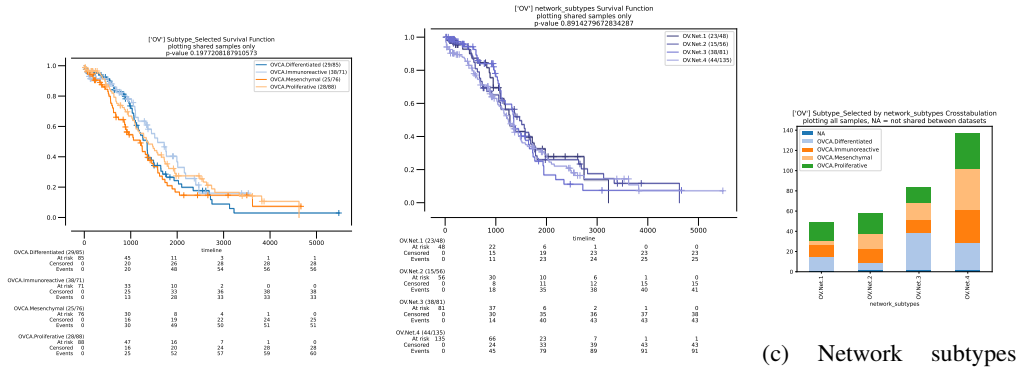


Figure S21: Exploration of network subtypes for Lymphoid Neoplasm Diffuse Large B-cell Lymphoma (DLBC), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function (b) Network subtypes survival function (c) Network subtypes crosstabulation with known subtype

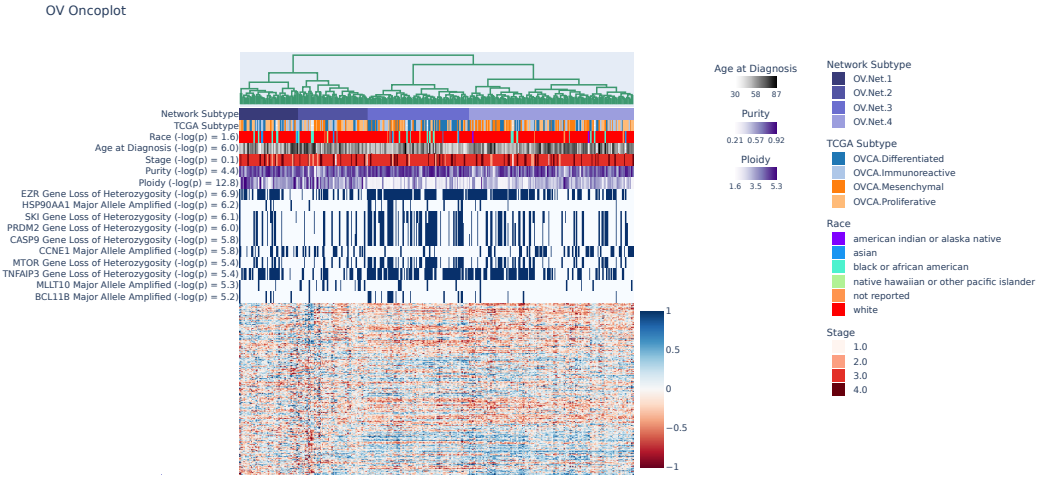
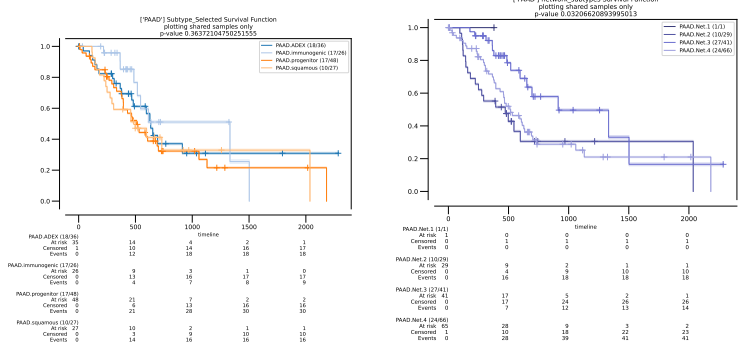
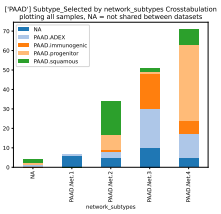


Figure S22: Exploration of network subtypes for Ovarian serous cystadenocarcinoma (OV), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function

(b) Network subtypes survival function



(c) Network subtypes crosstabulation with known subtype

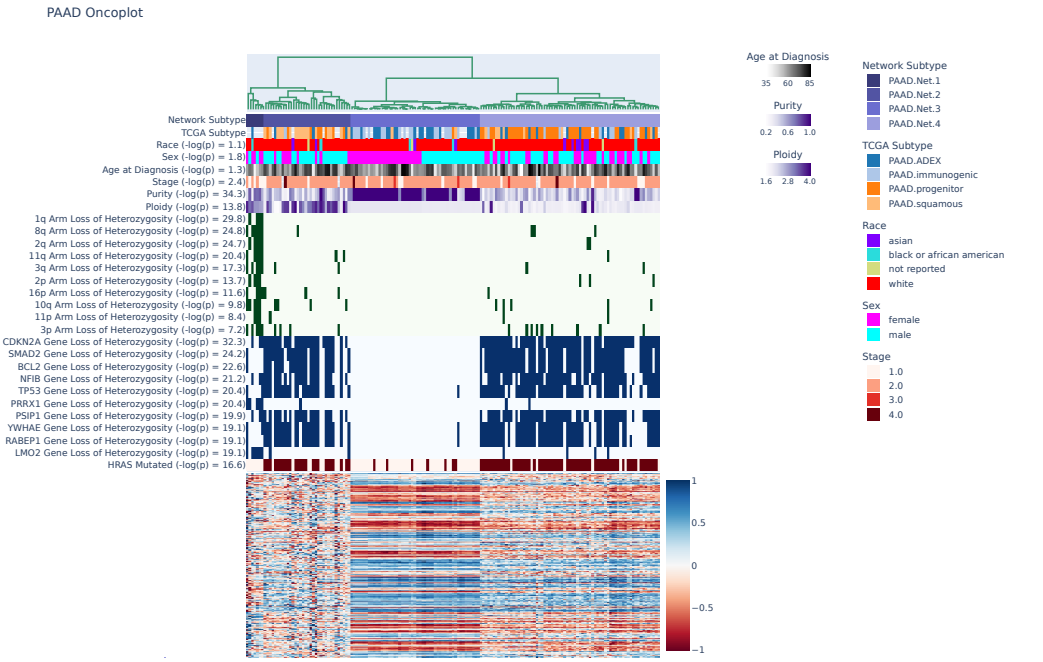
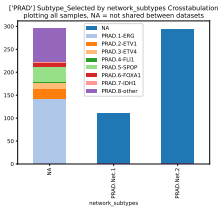


Figure S23: Exploration of network subtypes for Pancreatic adenocarcinoma (PAAD), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Network subtypes crosstabulation with known subtype

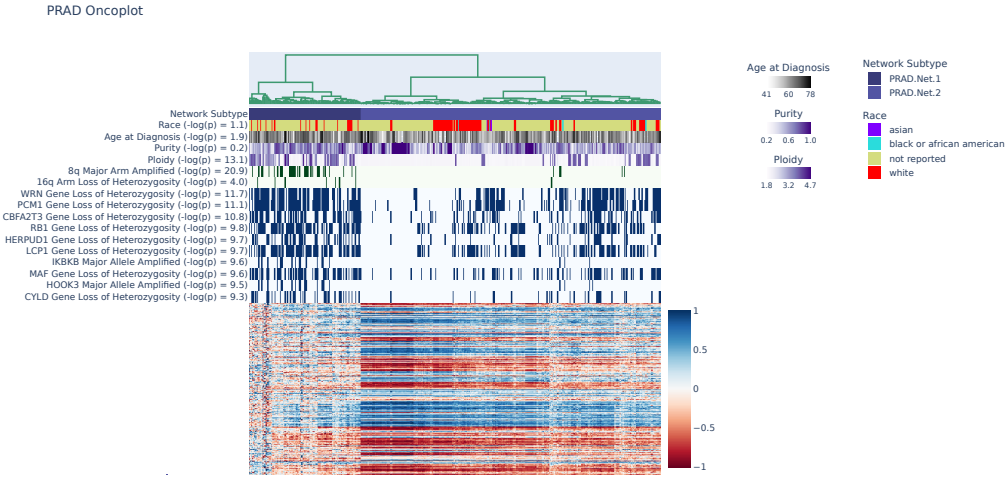
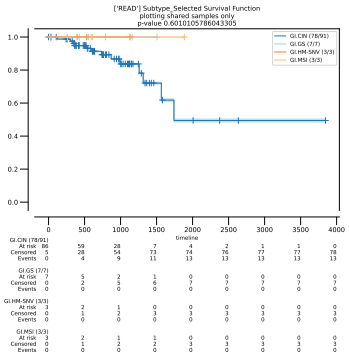
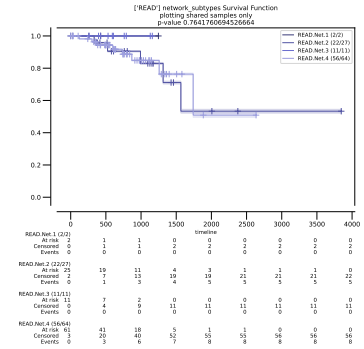


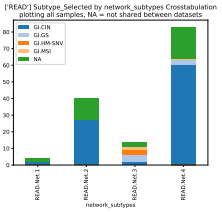
Figure S24: Exploration of network subtypes for Prostate adenocarcinoma (PRAD), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function



(b) Network subtypes survival function



(c) Network subtypes crosstabulation with known subtype

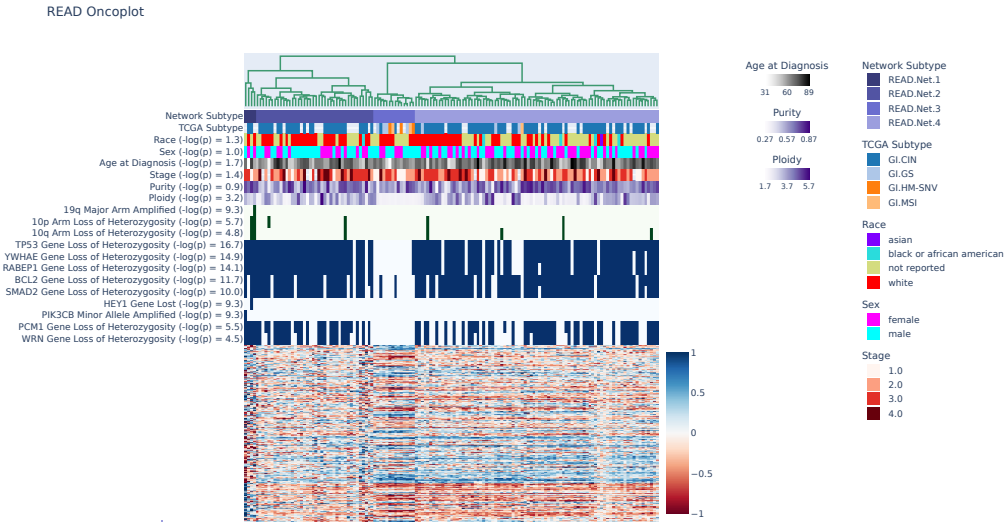
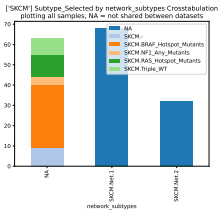


Figure S25: Exploration of network subtypes for Rectum adenocarcinoma (READ), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Network subtypes crosstabulation with known subtype

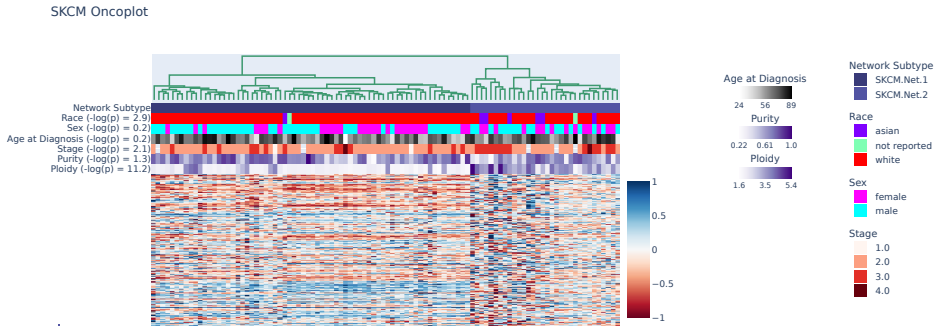
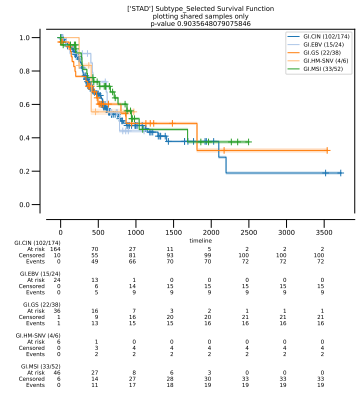
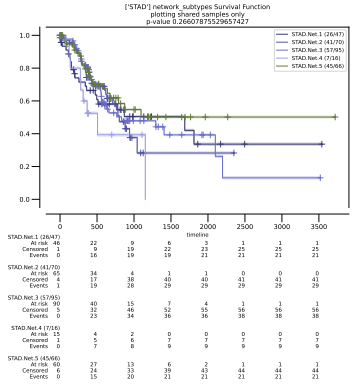


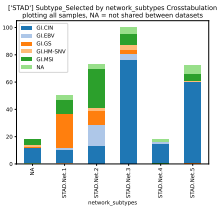
Figure S26: Exploration of network subtypes for Skin Cutaneous Melanoma (SKCM), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function



(b) Network subtypes survival function



(c) Network subtypes crosstabulation with known subtype

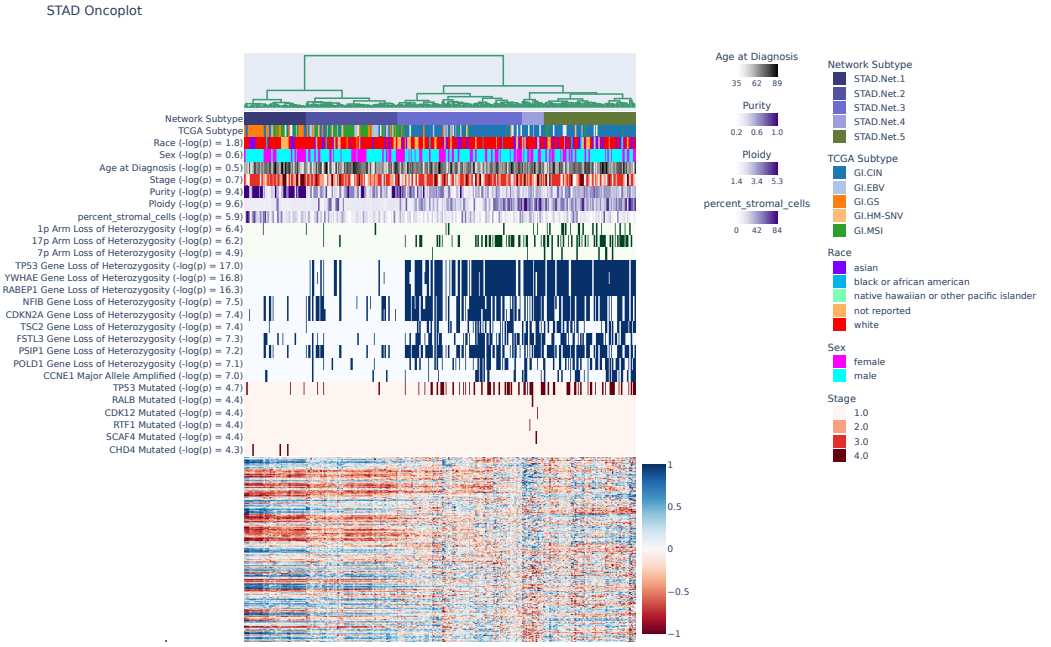
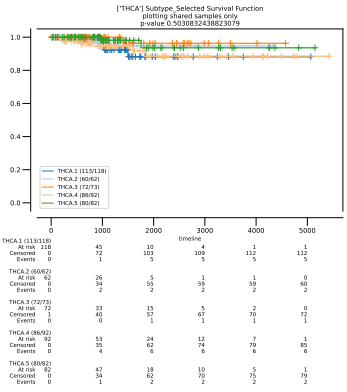
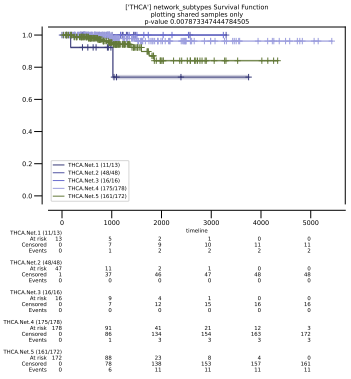


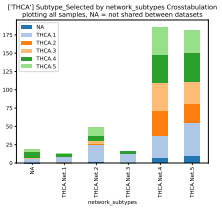
Figure S27: Exploration of network subtypes for Stomach adenocarcinoma (STAD), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function



(b) Network subtypes survival function



(c) Network subtypes crosstabulation with known subtype

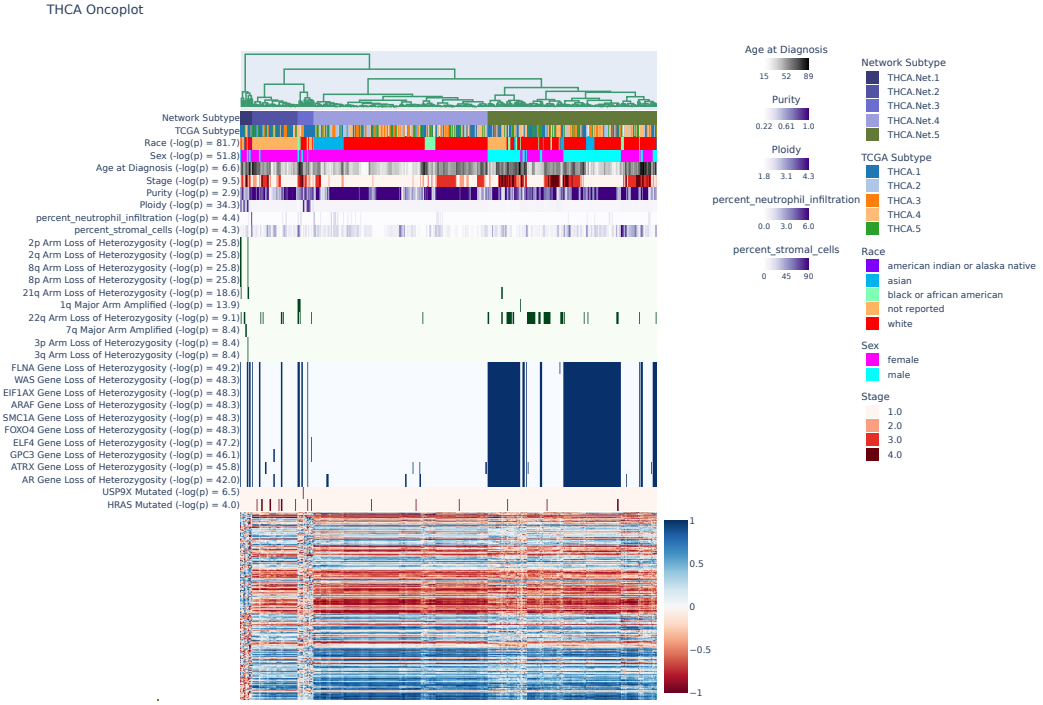
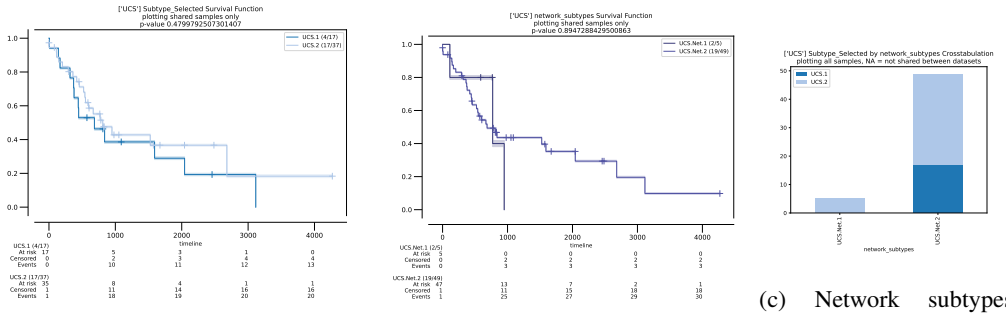


Figure S28: Exploration of network subtypes for Thyroid carcinoma (THCA), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function (b) Network subtypes survival function (c) Network subtypes crosstabulation with known subtype

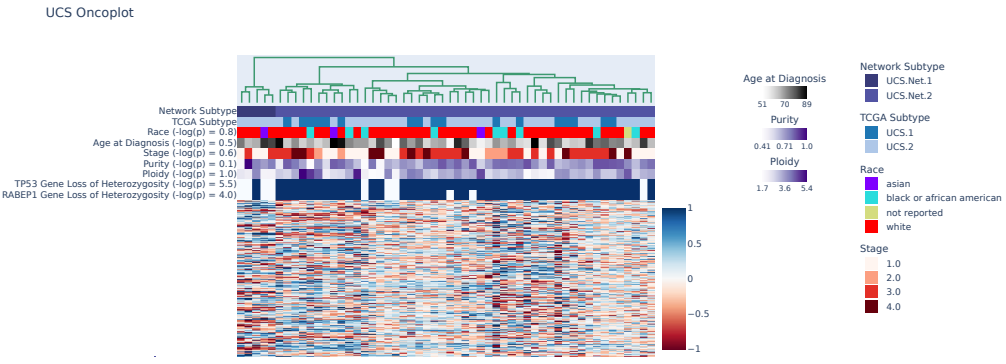
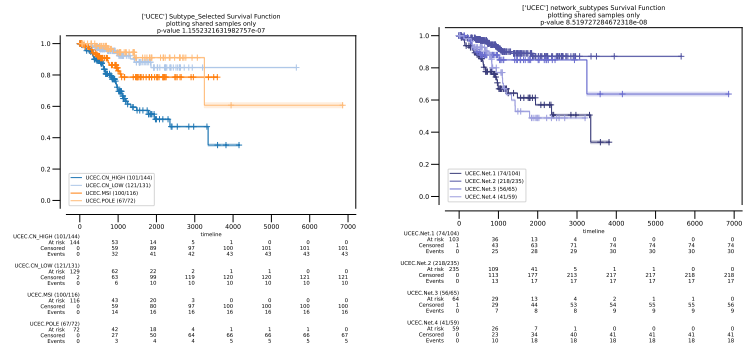


Figure S29: Exploration of network subtypes for Uterine Carcinosarcoma (UCS), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function (b) Network subtypes survival function (c) Network subtypes crosstabulation with known subtype

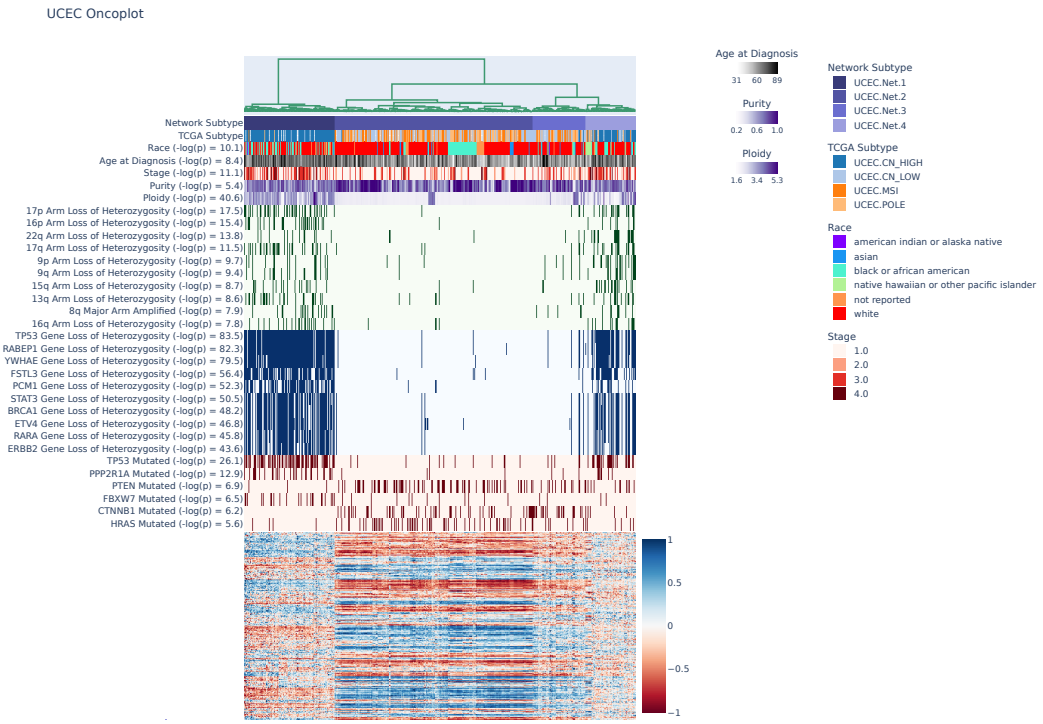
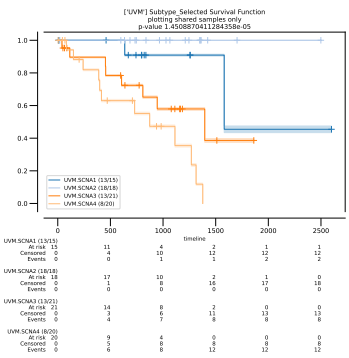
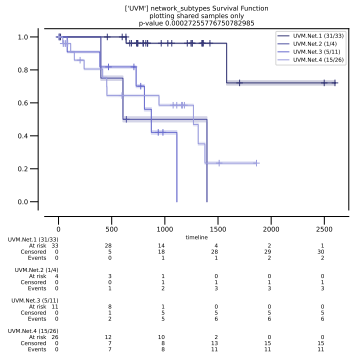


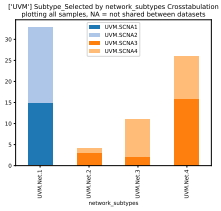
Figure S30: Exploration of network subtypes for Uterine Corpus Endometrial Carcinoma (UCEC), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.



(a) Known subtypes survival function



(b) Network subtypes survival function



(c) Network subtypes crosstabulation with known subtype

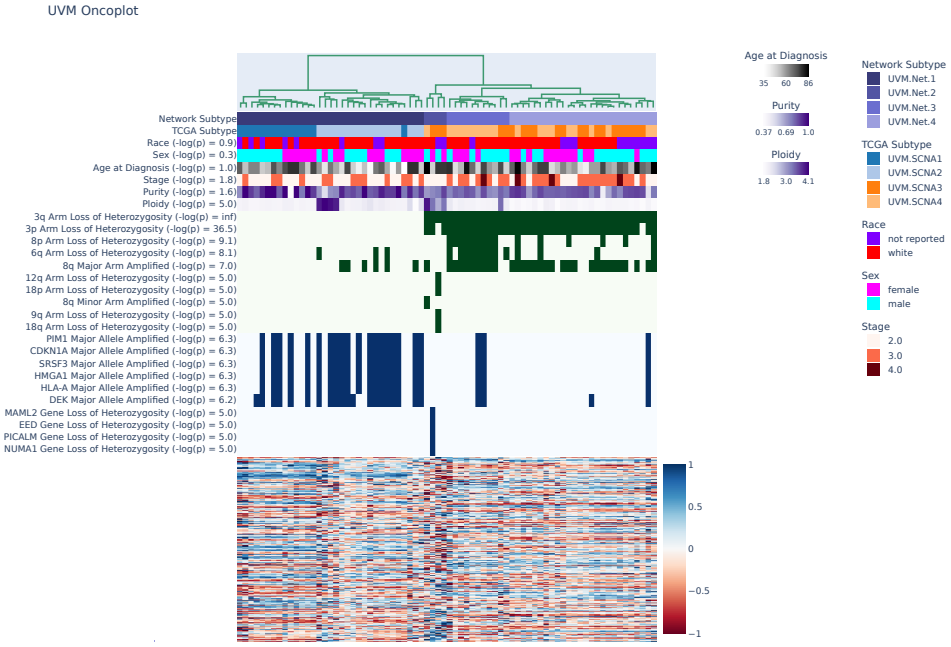


Figure S31: Exploration of network subtypes for Uveal Melanoma (UVM), looking at correlated clinical information, arm-level copy alterations, gene-level copy alterations, and gene-level single nucleotide variations.