Online Supplementary

Cluster derivation

We applied an unsupervised ML approach to develop clinical phenotypes of kidney transplant recipients with prolonged dialysis duration before transplant in the UNOS/OPTN database by conducting unsupervised consensus clustering.[1] We performed consensus clustering analysis on the whole study population. We initially assessed the distribution and missingness in phenotyping variables. Subsequently, missing data were imputed through multiple imputation using multivariate imputation by chained equations (MICE),[2] and non-normal data were z-score normalized. Multiple imputation is a widely used approach to estimate variables when data are missing at random. MICE is optimal when less than 30% of a variable's data are missing.[3-8] All of the extracted variables in our study had missing data less than 5%. We subsequently applied clustering using the consensus cluster algorithm. The algorithm begins by subsampling a proportion of items and a proportion of features from a data matrix. Each subsample is then partitioned into up to groups (k) by a user-specified clustering algorithm. This process is repeated for a specified number of times. Pairwise consensus values, defined as 'the proportion of clustering runs in which two items are grouped together, are calculated and stored in a consensus matrix (CM) for each cluster. Clustering settings used were as follows: maximum number of clusters, 10; number of iterations, 100; subsampling fraction, 0.8; clustering algorithm, K-means; Euclidean distance).[1] The number of potential clusters ranges from 2 to 10, to avoid producing an excessive number of clusters that would not be clinical useful. Pairwise consensus values, defined as 'the proportion of clustering runs in which two items are [grouped] together[1], are calculated and stored in a CM for each k. Then for each k, a final agglomerative hierarchical consensus clustering using distance of 1−consensus values is completed and pruned to k groups, which are called consensus clusters.

The clustering algorithm is to maximize the potential number of clusters while maintaining high cluster consensus. The optimal number of clusters was determined by examining the CM heat map, cumulative distribution function, cluster-consensus plots with the within-cluster consensus scores, and the proportion of ambiguously clustered pairs (PAC).[9, 10] The within-cluster consensus score, ranging between 0 and 1, is defined as the average consensus value for all pairs of individuals belonging to the same cluster.[10] A value closer to one indicates better cluster stability.[10] PAC, ranging between 0 and 1, is calculated as the proportion of all sample pairs with consensus values falling within the predetermined boundaries.[9] A value closer to zero indicates better cluster stability.[9] To examine the cluster profile, we calculated and graphically displayed the standardized mean differences of the variables between each cluster and the overall study population. Calculation of the standardized difference of each parameter used the cutoff of ±0.3 to show subgroup features with the key features for each cluster.

All cluster derivation analyses were performed using R, version 4.0.3 (RStudio, Inc., Boston, MA; http://www.rstudio.com/), with the packages of ConsensusClusterPlus (version 1.46.0)[10]. We imputed missing data through multivariable imputation by chained equation (MICE) method.[2] All analyses were two-tailed, and P value < .05 was considered statistically significant.

Table S1 Clinical characteristics assciated with death-censored graft failure and patient death

	Death-censored graft failure		Patient death	
	HR (95% CI)	p-value	HR (95% CI)	p-value
Recipient age (per 10-year increase)	0.93 (0.86-1.01)	0.10	1.79 (1.64-1.95)	<0.001
Recipient male sex	1.34 (1.11-1.61)	0.002	1.36 (1.14-1.63)	0.001
Recipient race				
- White	1 (ref)	-	1 (ref)	-
- Black	1.77 (1.31-2.37)	<0.001	0.83 (0.66-1.05)	0.12
- Hispanic	1.02 (0.72-1.44)	0.93	0.73 (0.50-1.04)	0.06
- Other	0.99 (0.63-156)	0.97	0.74 (0.51-1.08)	0.12
Body mass index (per 5-kg/m2 increase)	1.12 (1.05-1.21)	0.002	1.14 (1.07-1.22)	<0.001
Dialysis duration	1.01 (0.98-1.03)	0.53	1.02 (1.01-1.04)	0.02
Cause of end-stage kidney disease				
- Diabetes mellitus	1 (ref)	-	1 (ref)	-
- Hypertension	1.16 (0.89-1.52)	0.28	0.47 (0.39-0.58)	<0.001
- Glomerular disease	1.20 (0.89-1.62)	0.22	0.28 (0.21-0.37)	<0.001
- PKD	0.55 (0.31-0.96)	0.04	0.52 (0.35-0.74)	0.001
- Other	0.98 (0.71-1.38)	0.94	0.35 (0.26-0.47)	<0.001
Comorbidity				
- Diabetes mellitus	1.09 (0.87-1.35)	0.45	2.51 (2.10-3.00)	<0.001
- Malignancy	0.90 (0.63-1.25)	0.56	1.37 (1.03-1.80)	0.03
- Peripheral vascular disease	1.36 (1.03-1.77)	0.03	2.24 (1.78-2.80)	<0.001
PRA	1.00 (0.99-1.01)	0.43	0.997 (0.994-1.00)	0.052
Positive HCV serostatus	1.03 (0.75-1.42)	0.85	1.19 (0.88-1.58)	0.24
Positive HBs antigen	1.07 (0.65-1.66)	0.77	1.02 (0.64-1.61)	0.94
Positive HIV serostatus	1.75 (1.26-2.39)	0.001	1.00 (0.66-1.45)	1.00
Karnofsky performance scale index below 80%	1.01 (0.85-1.21)	0.91	1.28 (1.08-1.52)	0.004
Serum albumin (per 1-g/dL increase)	0.95 (0.81-1.11)	0.53	0.80 (0.69-0.93)	0.003
Kidney donor status				
- Non-ECD deceased	1.36 (0.80-2.32)	0.26	1.35 (0.82-2.22)	0.24
- ECD deceased	3.25 (1.85-5.71)	<0.001	3.00 (1.76-5.11)	<0.001
- Living	1 (ref)	-	1 (ref)	-
Donor age (per 10-year increase)	1.19 (1.12-1.27)	<0.001	1.21 (1.14-1.29)	<0.001
Donor male sex	0.76 (0.64-0.91)	0.003	0.85 (0.71-1.01)	0.06
Donor race				
- White	1 (ref)	-	1 (ref)	-
- Black	1.33 (1.06-1.64)	0.01	1.24 (1.00-1.54)	0.05
- Hispanic	0.97 (0.76-1.23)	0.81	1.08 (0.86-1.34)	0.52
- Other	1.28 (0.85-1.84)	0.21	1.05 (0.68-1.55)	0.81
History of hypertension in donor	1.75 (1.45-2.10)	<0.001	1.64 (1.37-1.96)	<0.001
KDPI				
- Living donor	1 (ref)	-	1 (ref)	-
- KDPI<85	1.42 (0.83-2.42)	0.20	1.38 (0.84-2.28)	0.20
- KDPI≥85	4.28 (2.36-7.77)	<0.001	4.39 (2.51-7.66)	<0.001
HLA mismatch	1.13 (1.05-1.21)	0.002	1.04 (0.98-1.12)	0.19
Cold ischemia time (per 6-hour increase)	1.08 (1.03-1.13)	0.003	1.04 (0.98-1.09)	0.18
Delay graft function	2.36 (1.98-2.82)	<0.001	1.85 (1.56-2.19)	<0.001
High risk EBV status	1.25 (0.94-1.66)	0.12	1.20 (0.91-1.58)	0.19
High risk CMV status	1.09 (0.85-1.39)	0.48	1.24 (0.98-1.55)	0.06

Figure S1. Consensus matrix heat map (k = 2) depicting consensus values on a white to blue color scale of each cluster

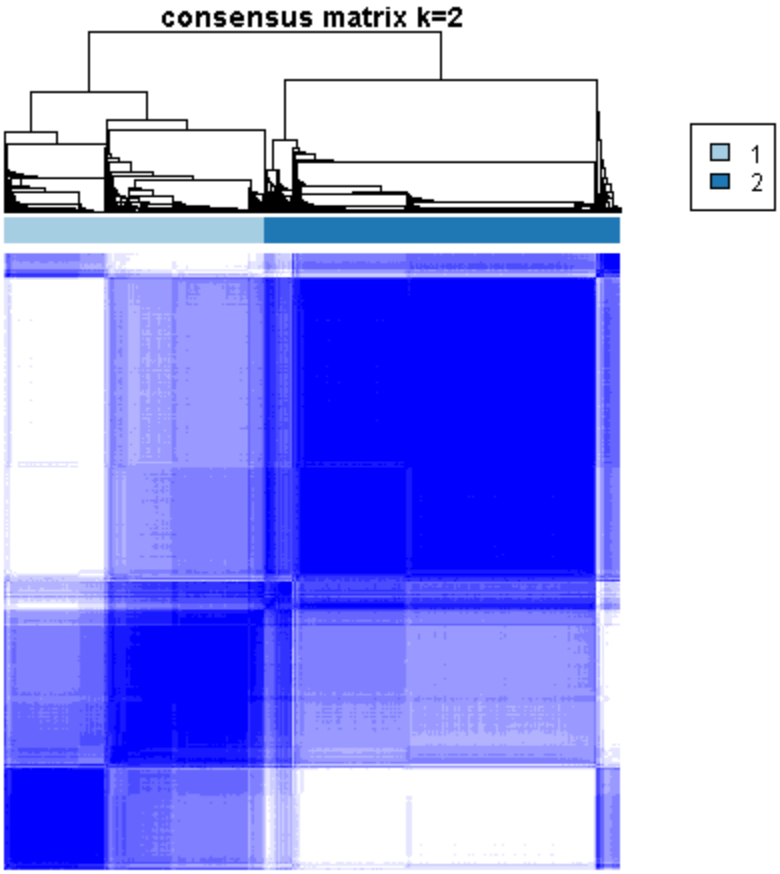


Figure S2. Consensus matrix heat map (k = 3) depicting consensus values on a white to blue color scale of each cluster

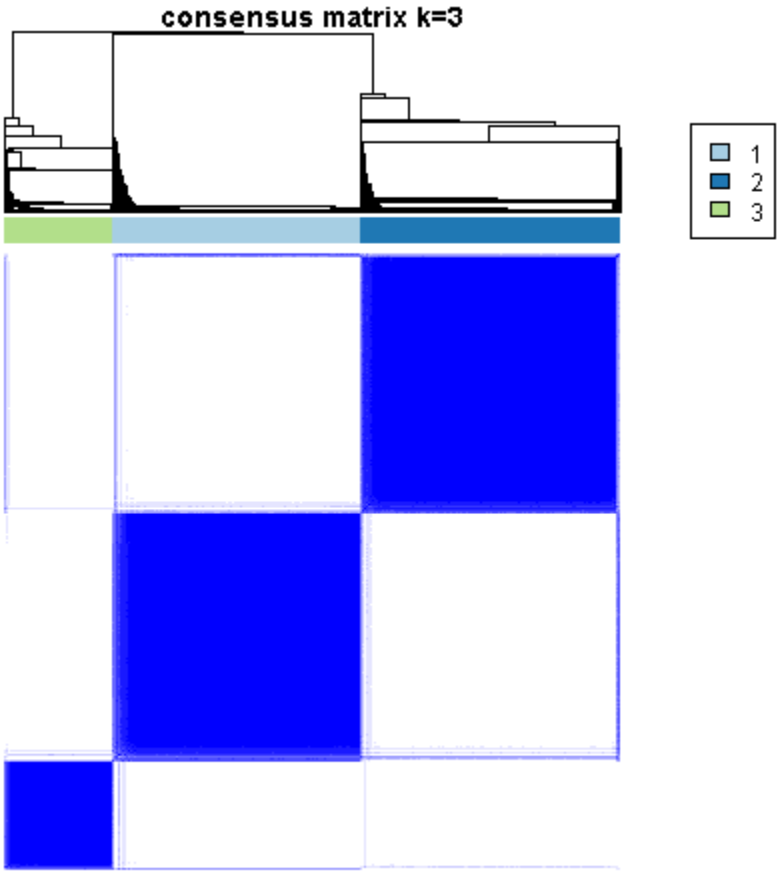


Figure S3. Consensus matrix heat map (k = 4) depicting consensus values on a white to blue color scale of each cluster

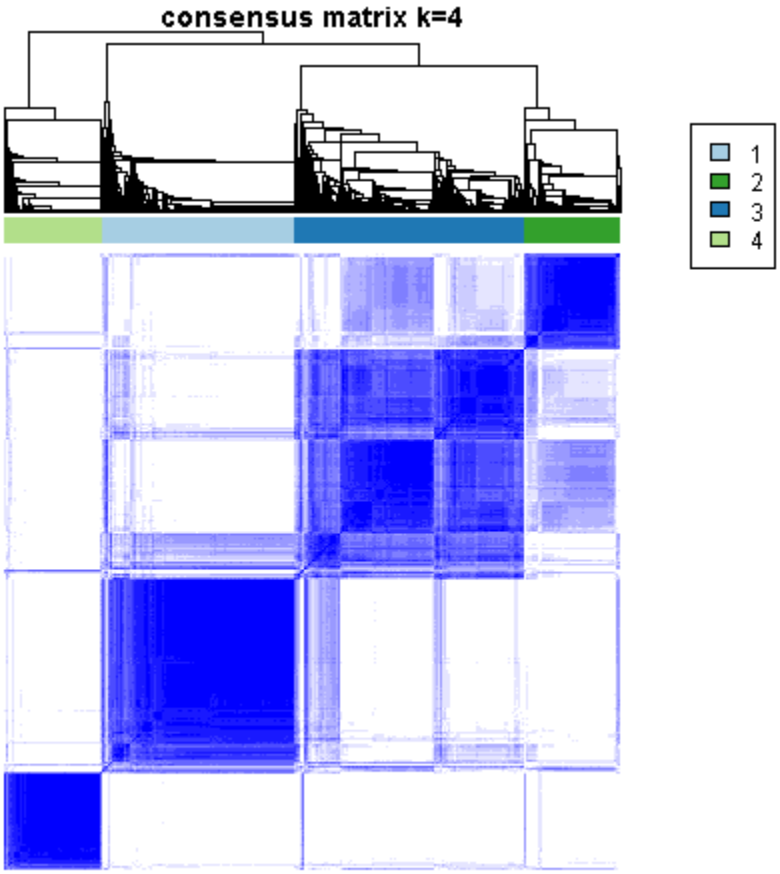


Figure S4. Consensus matrix heat map (k = 5) depicting consensus values on a white to blue color scale of each cluster

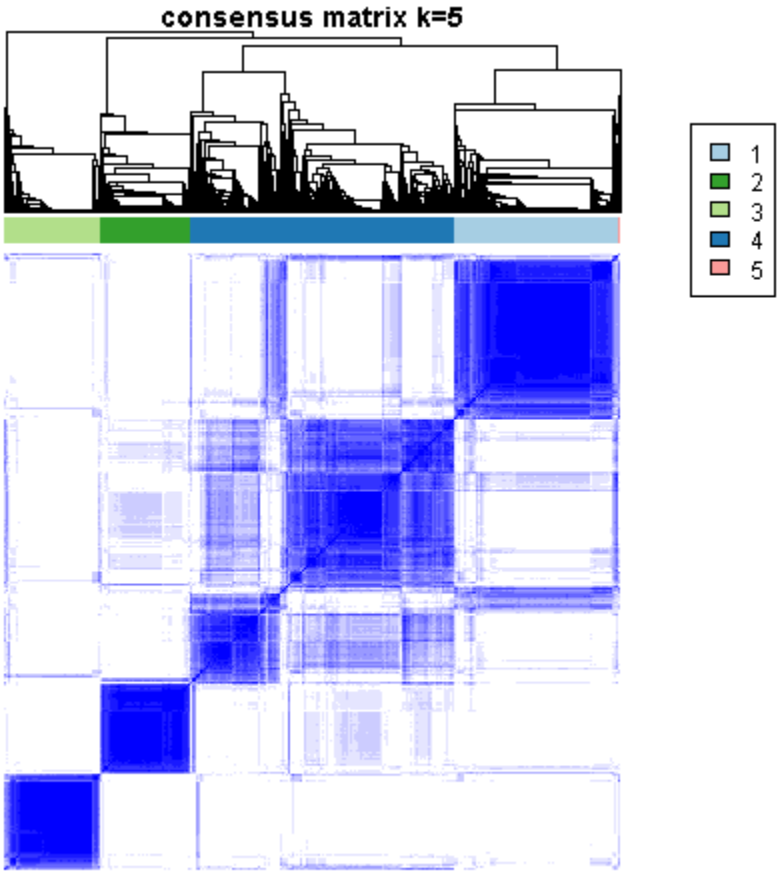


Figure S5. Consensus matrix heat map (k = 6) depicting consensus values on a white to blue color scale of each cluster

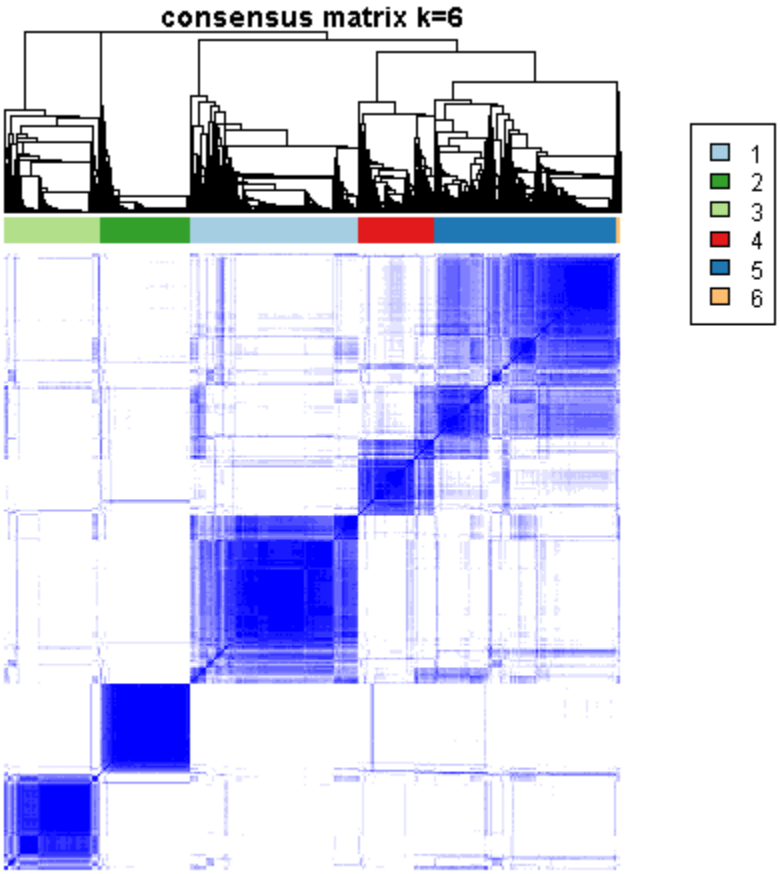


Figure S6. Consensus matrix heat map (k = 7) depicting consensus values on a white to blue color scale of each cluster

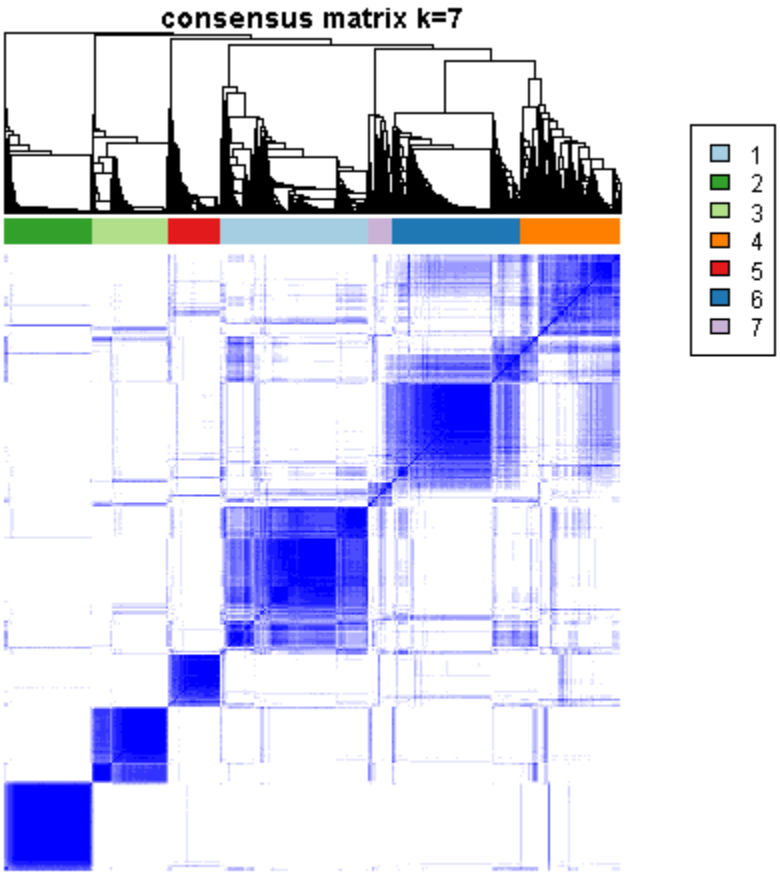


Figure S7. Consensus matrix heat map (k = 8) depicting consensus values on a white to blue color scale of each cluster

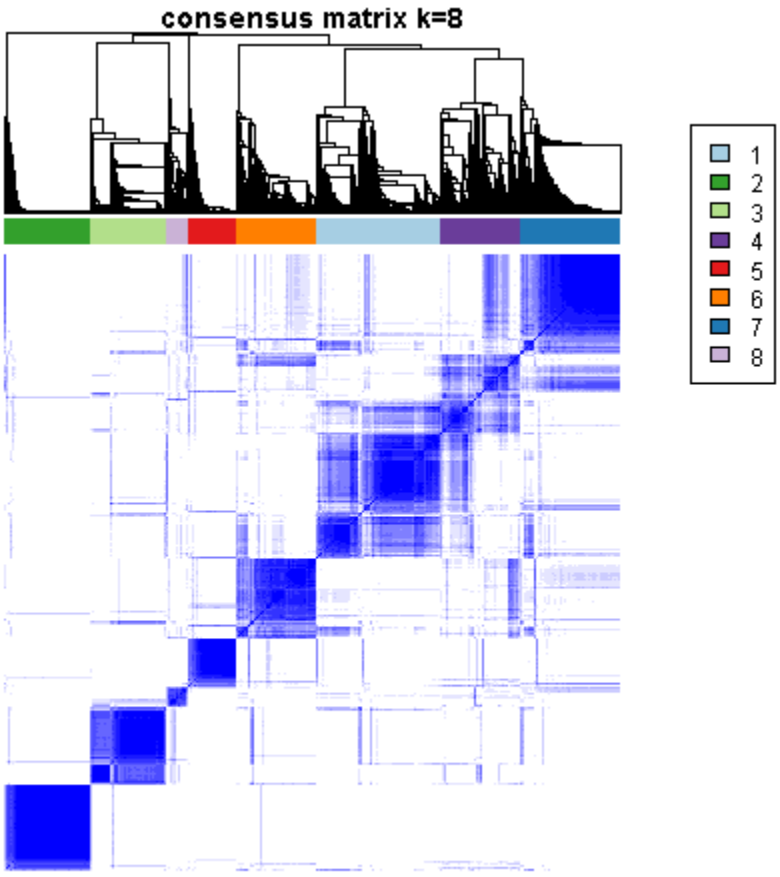


Figure S8. Consensus matrix heat map (k = 9) depicting consensus values on a white to blue color scale of each cluster

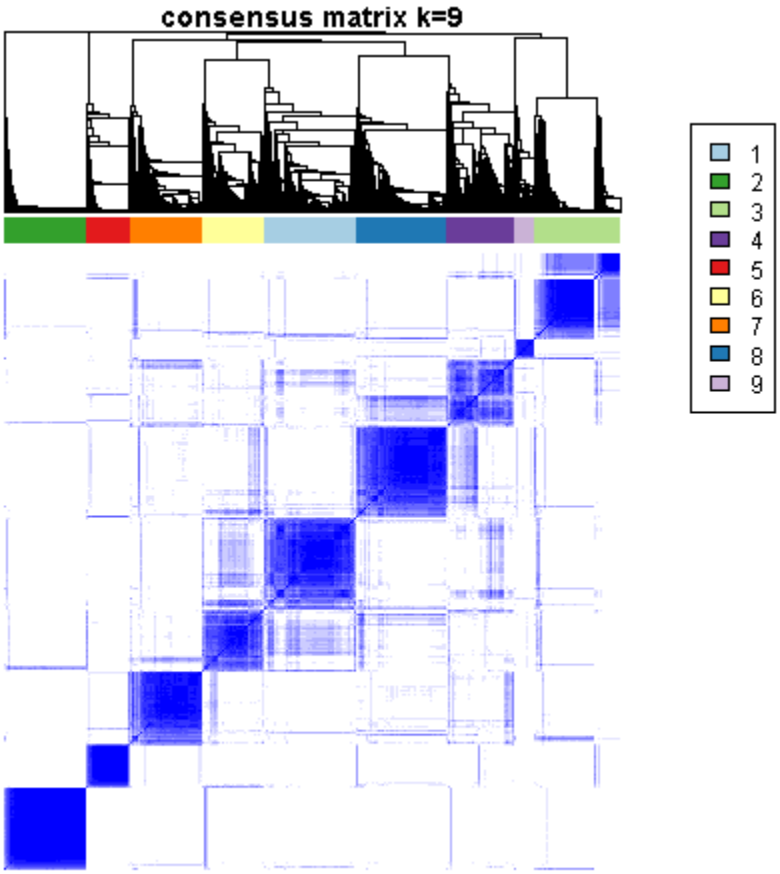
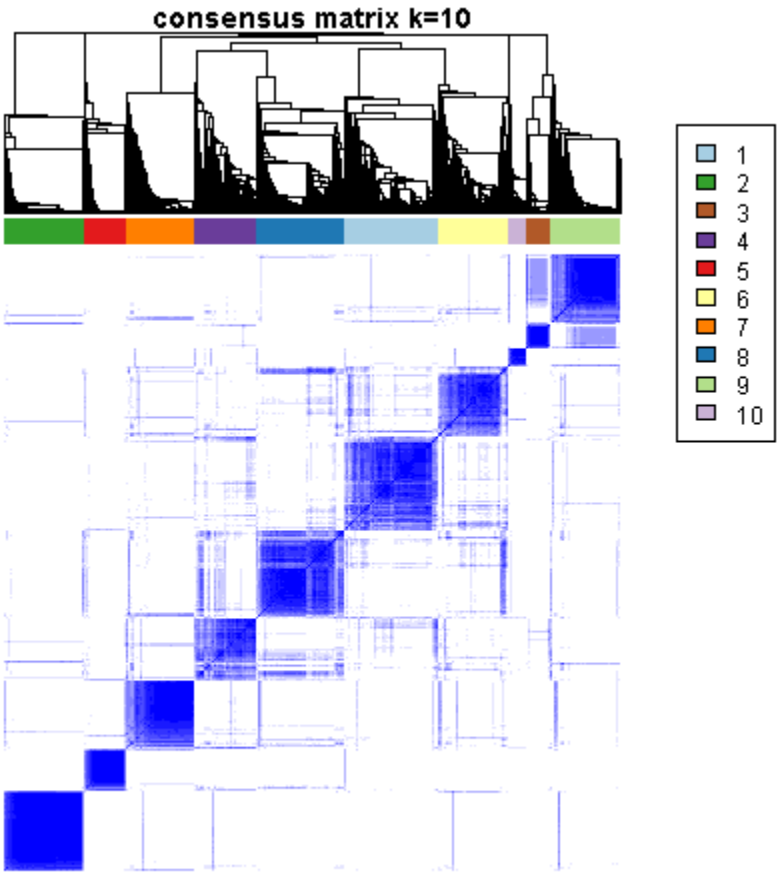


Figure S9. Consensus matrix heat map (k = 10) depicting consensus values on a white to blue color scale of each cluster



References

1 Monti S, Tamayo P, Mesirov J, Golub T. Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. Machine learning. 2003: 91
2 Van Buuren S, Groothuis-Oudshoorn K. mice: Multivariate imputation by chained equations in R. Journal of statistical software. 2011: 1
3 Jannat-Khah DP, Unterbrink M, McNairy M, Pierre S, Fitzgerald DW, Pape J, Evans A. Treating loss-to-follow-up as a missing data problem: a case study using a longitudinal cohort of HIV-infected patients in Haiti. BMC Public Health. 2018: 1269 [PMID: 30453995 10.1186/s12889-018-6115-0: 10.1186/s12889-018-6115-0]
4 Knol MJ, Janssen KJ, Donders AR, Egberts AC, Heerdink ER, Grobbee DE, Moons KG, Geerlings MI. Unpredictable bias when using the missing indicator method or complete case analysis for missing confounder values: an empirical example. J Clin Epidemiol. 2010: 728 [PMID: 20346625 10.1016/j.jclinepi.2009.08.028: 10.1016/j.jclinepi.2009.08.028]
5 White IR, Carlin JB. Bias and efficiency of multiple imputation compared with complete-case analysis for missing covariate values. Stat Med. 2010: 2920 [PMID: 20842622 10.1002/sim.3944: 10.1002/sim.3944]
6 White IR, Royston P, Wood AM. Multiple imputation using chained equations: Issues and guidance for practice. Stat Med. 2011: 377 [PMID: 21225900 10.1002/sim.4067: 10.1002/sim.4067]
7 Hedden SL, Woolson RF, Carter RE, Palesch Y, Upadhyaya HP, Malcolm RJ. The impact of loss to follow-up on hypothesis tests of the treatment effect for several statistical methods in substance abuse clinical trials. J Subst Abuse Treat. 2009: 54 [PMID: 19008067 10.1016/j.jsat.2008.09.011: 10.1016/j.jsat.2008.09.011]
8 Donders ART, Van Der Heijden GJ, Stijnen T, Moons KG. A gentle introduction to imputation of missing values. Journal of clinical epidemiology. 2006: 1087
9 Şenbabaoğlu Y, Michailidis G, Li JZ. Critical limitations of consensus clustering in class discovery. Sci Rep. 2014: 6207 [PMID: 25158761 10.1038/srep06207: 10.1038/srep06207]
10 Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010: 1572