Supplementary Material.

Part A: Additional Simulation Results

A1: Method Comparison

Here are more detailed results for the comparison between BINCO and *stability selection*. The simulation setting is the same as described in Section 3.1. Tables S-1 to S-6 are the results for *stability selection* and Table S-7 is the result for BINCO.

TABLE S-1
Power and FDR of Stability Selection, $l = 1$, $\lambda_{max} = 100$, Strong Signal.

$\lambda_{\min}$	Targeted FDR = 0.05				Targeted FDR = 0.10			
	FDR		Power		FDR		Power	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
≤ 40	FDR control not achievable				FDR control not achievable			
50	0.061 [1]	0.008	0.818	0.013	0.077	0.009	0.836	0.012
60	0.060 [1]	0.007	0.783	0.011	0.063	0.008	0.790	0.011
70	0.054 [1]	0.007	0.725	0.010	0.055	0.007	0.729	0.012
80	0.050	0.007	0.663	0.010	0.051	0.006	0.666	0.011
90	0.050	0.006	0.567	0.014	0.050	0.006	0.572	0.014
100	0.062 [1]	0.008	0.418	0.006	0.061	0.007	0.427	0.016

[1] FDR control failed.

TABLE S-2
Power and FDR of Stability Selection, $l = 0.8$, $\lambda_{max} = 100$, Strong Signal.

$\lambda_{\min}$	Targeted FDR = 0.05				Targeted FDR = 0.10			
	FDR		Power		FDR		Power	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
40	FDR control not achievable				0.099	0.012	0.823	0.011
50	0.077 [1]	0.009	0.785	0.011	0.091	0.012	0.797	0.011
60	0.056 [1]	0.009	0.734	0.011	0.059	0.009	0.739	0.012
70	0.033	0.008	0.668	0.010	0.036	0.009	0.675	0.009
80	0.017	0.006	0.568	0.012	0.018	0.007	0.574	0.013
90	0.010	0.007	0.400	0.015	0.010	0.008	0.408	0.015
100	0.005	0.007	0.207	0.001	0.007	0.008	0.214	0.010

[1] FDR control failed.

TABLE S-3
Power and FDR of Stability Selection, $l = 0.5$, $\lambda_{max} = 100$, Strong Signal.

λ_{min}	Targeted FDR = 0.05				Targeted FDR = 0.10			
	FDR		Power		FDR		Power	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
40	0.007	0.005	0.706	0.011	0.021	0.006	0.747	0.006
50	0.007	0.004	0.662	0.009	0.010	0.005	0.676	0.005
60	0.001	0.002	0.526	0.017	0.002	0.003	0.540	0.003
70	0	0	0.271	0.012	0	0	0.287	0.013
≥ 80	Omitted for too small power.							

TABLE S-4
Power and FDR of Stability Selection, $l = 1$, $\lambda_{max} = 100$, Weak Signal.

λ_{min}	Targeted FDR = 0.05				Targeted FDR = 0.10			
	FDR		Power		FDR		Power	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
≤ 40	FDR control not achievable				FDR control not achievable			
50	FDR control not achievable				0.006	0.006	0.490	0.020
60	0.003	0.004	0.434	0.017	0.004	0.004	0.458	0.013
70	0	0.002	0.288	0.017	< 0.001	0.001	0.298	0.018
80	0	0	0.126	0.015	0	0	0.131	0.014
≥ 90	Omitted for too small power.							

TABLE S-5
Power and FDR of Stability Selection, $l = 0.8$, $\lambda_{max} = 100$, Weak Signal.

λ_{min}	Targeted FDR = 0.05				Targeted FDR = 0.10			
	FDR		Power		FDR		Power	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
40	FDR control not achievable				FDR control not achievable			
50	0.002	0.003	0.407	0.026	0.006	0.006	0.485	0.016
60	0.001	0.003	0.328	0.016	0.001	0.003	0.342	0.014
70	0	0	0.143	0.016	0	0	0.149	0.015
≥ 80	Omitted for too small power.							

TABLE S-6
Power and FDR of Stability Selection, $l = 0.5$, $\lambda_{max} = 100$, Weak Signal.

λ_{min}	Targeted FDR = 0.05				Targeted FDR = 0.10			
	FDR		Power		FDR		Power	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
40	FDR control not achievable				FDR control not achievable			
50	FDR control not achievable				0.001	0.002	0.170	0.020
60	0	0	0.006	0.012	0	0	0.025	0.011
≥ 70	Omitted for too small power.							

Table S-7
Power and FDR of BINCO

Signal Strength	Targeted FDR = 0.05 FDR Mean	SD	Power Mean	SD	Targeted FDR = 0.10 FDR Mean	SD	Power Mean	SD
Strong	0.026	0.016	0.801	0.023	0.056	0.025	0.835	0.016
Weak	0.034	0.011	0.569	0.032	0.059	0.017	0.610	0.028

A2: U-shape Diagnostics for Empty and Power-law Networks.

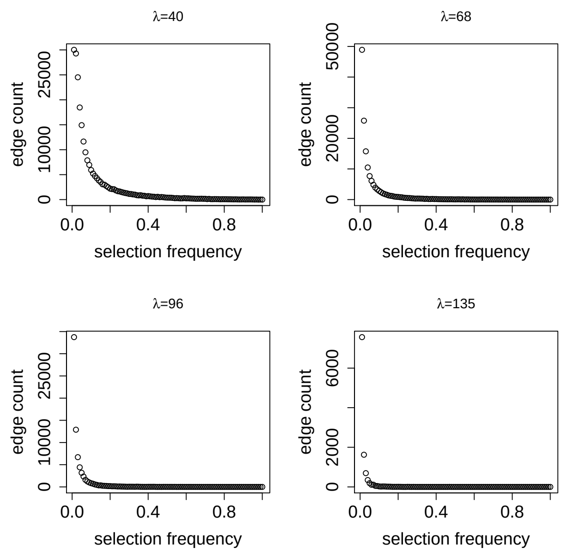

Fig S-1: Diagnostic on the empirical distribution of selection frequencies from empty network: no "U-shape" characteristic is observed for λ in a wide range.

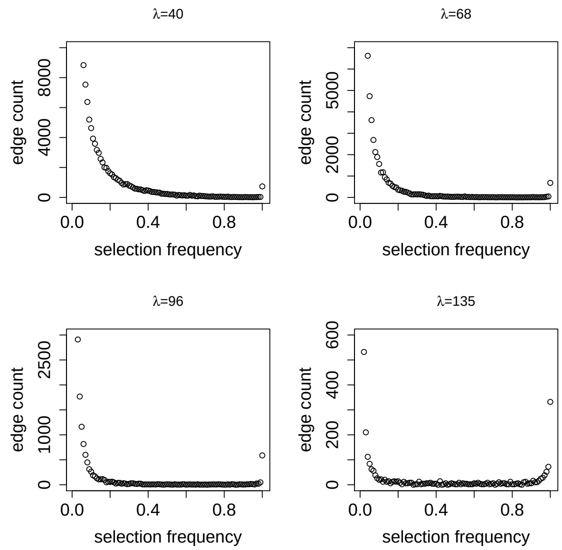

Fig S-2: Diagnostic on the empirical distribution of selection frequencies from power-law network: "U-shape" characteristic is observed for λ in a wide range. Note the "U-shape" characteristic is also observed for the empirical selection frequency distributions from empirical and hub networks. Those diagnostic plots are very similar to this one and hence omitted.

A3: Additional Simulation Results for BINCO.

Here we investigate the impact of the number of components in the networks on BINCO's performance. We consider the power-law network with sample size $n = 200$ and the number of nodes $p = 500$. The signal strength is fixed at the strong level as in Section 3.1. Networks are generated by varying the number of components $C = 5$, 2 to 1.

Since components are independent, the model dimensionality is the size of each component which is smaller for larger C. Thus, as the number of components decreases, it might be more challenging to detect the network due to the increasing dimensionality for each component. Nevertheless, for all three networks with different numbers of components, BINCO provides proper (and slightly conservative) control for FDR and decent power (Table S-8).

TABLE S-8
Investigation of the Impact of Different Number of components in Power-law Networks on BINCO Performance

Number of components	Targeted FDR = 0.05				Targeted FDR = 0.10			
	FDR		Power		FDR		Power	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
5	0.046	0.009	0.810	0.013	0.096	0.013	0.845	0.013
2	0.048	0.012	0.783	0.011	0.096	0.016	0.814	0.011
1	0.039	0.015	0.804	0.011	0.095	0.020	0.836	0.013

Part B: Details of the Hub Genes Detected by BINCO on the Breast Cancer Data

Table S-9
Annotation of Hub Genes and Their Connections to breast cancer (BC).

Rank of Degree [1]	Gene Symbol	Pathway belongs to	Function Summary	Connection to BC
1	MBD4	Brentani-DNA-Methylation-and-Modification; DNA-Binding	encoding methyl-CpG binding domain protein 4	over-expressed and amplified in human BC (Zhu *et al.*, 1999) and differentially expressed in mammary epithelial cells (Jiang *et al.*, 2010);
2	TARDBP	DNA-Binding	encoding TAR DNA binding protein	is well known to be associated to neurodegenerative disorders while its relationship with cancer is discovered recently (Postel-Vinay *et al.*, 2012). The role it plays in BC needs further investigation
3	DDB2	Brentani-Repair; DNA-Damage-Signaling; Damaged-DNA-Binding; DNA-Binding	encoding damage-specific DNA binding protein 2, 48kDa	highly expressed in the human ER-positive breast tumor samples and plays a significant role as an activator of BC cell growth (Kattan, *et al.*, 2008).
4	MAP3K4	Brentani-Signaling	encoding mitogen-activated protein kinase kinase kinase 4	plays a role in the signal transduction pathways of BC cell proliferation, survival, and apoptosis (Bild and Johnson, 2001)
5	ORC3L	Cell-Cycle-KEGG; Cell-Cycle; G1-to-S-Cell-Cycle-Reactome; DNA-Replication-Reactome; HSA04110-Cell-Cycle; DNA-Binding	encoding origin recognition complex, subunit 3-like (yeast)	belongs to the ORC group, which plays important role in the p53 cell cycle pathway where a mutation in the p53 gene is the most common genetic change found in BC. (http://www.genome.jp/kegg/pathway/hsa/hsa04110.html)
6	CDKN1B	Brentani-Cell-Cycle; Cell-Cycle-Regulator; Cell-Cycle; G1-to-S-Cell-Cycle-Reactome; Cell-Cycle-Arrest; HSA04110-Cell-Cycle	encoding cyclin-dependent kinase inhibitor 1B (p27, Kip1), controlling the cell cycle progression at G1	an essential regulators of cell cycle progression where its genetic variants have been verified to be associated to BC risk (Ma *et al.*, 2006 and Canbay *et al.*, 2010)
7	REL	Brentani-Signaling	encoding c-Rel, a transcription factor that is a member of the Rel/NFKB family	may be involved early in the progression of breast epithelial cells towards malignancy (Romieu-Mourez *et al.*, 2001)
8	ATR	DNA-Damage-Signaling; Cell-Cycle-Checkpoint-II; HSA04110-Cell-Cycle	encoding ataxia telangiectasia and Rad3 related	there are potential interaction effects of variations in ATM/ATR/BRCA1/BRCA2 genes for BC (Wang *et al.*, 2010).
9	LGMN	DNA-Damage-Signaling	encoding legumain, a cysteine protease that has a strict specificity for hydrolysis of asparaginyl bonds	may play a role in tumor progression and is important in prognostic for BC (Gawenda *et al.*, 2007)
10	CDKN3	Cell-Cycle-Arrest	encoding cyclin-dependent kinase inhibitor 3	over-expressed in and hence associated with breast and prostate malignancies (Lee *et al.*, 1999)

[1] The rank of the number of connected edges (from the largest to the smallest) for each gene based on the estimated network by BINCO.

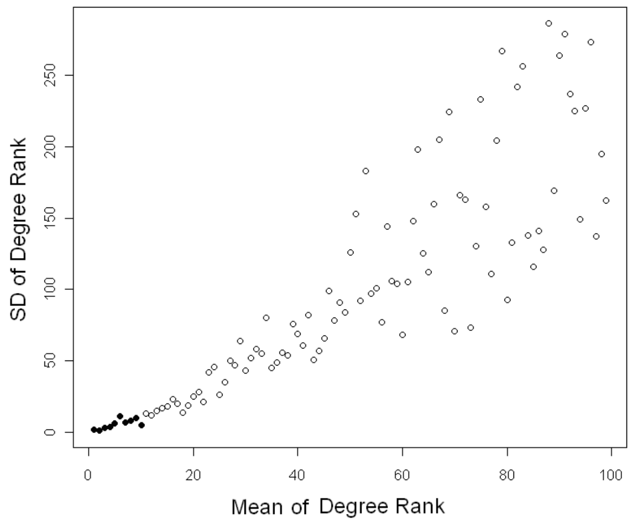

Fig S-3: A scatter-plot of the SD of degree ranks (small to large) versus the mean of degree ranks (large to small) of each gene/probe. The solid circles are the top 10 genes.

Part C: Examples of p_{ij} and $\tilde{p}_{ij}$ being Close

EXAMPLE 1. *Subsampling.*

Since subsampling $Y'_{(m)}$ of size m from a random sample $Y_{(n)}$ is equivalent to directly sampling a random sample $Y_{(m)}$ of size m, the asymptotic behavior of p_{ij} and $\tilde{p}_{ij}$ should be the same. In particular, if $p_{ij}^{(n)}$ has a limit, then $\tilde{p}_{ij}^{(m)}$ converges to the same limit and hence $p_{ij}^{(n)} - \tilde{p}_{ij}^{(m)} \to 0$ as $m, n \to \infty$.

EXAMPLE 2. *Lasso with selection consistency under linear regression settings.*

Consider a linear regression model

$$Y = \mathbf{X}\beta + \epsilon$$

where, for sample size n, Y is an $n \times 1$ response, $\mathbf{X} = (X_1, \ldots, X_p)$ is the $n \times p$ design matrix and ϵ is the random error with mean $\mathbf{0}$ and covariance $\mathbf{I}$. β is the coefficient vector that needs to estimate.

Denote the covariance matrix of $\mathbf{X}$ by $C = E(\mathbf{X}'\mathbf{X})$ and write C as

$$C = \begin{pmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{pmatrix},$$

where C_{11} is the covariance matrix of relevant variables in $\mathbf{X}$, C_{22} is the covariance matrix of irrelevant variables in $\mathbf{X}$ and $C_{12} = C_{21}$ is the matrix of covariance between relevant and irrelevant variables in $\mathbf{X}$. When p is fixed, the selection consistency of the Lasso procedure is equivalent to the irrepresentable condition (Zhao and Yu, 2006)

$$|C_{12}(C_{11})^{-1} sign(\beta(1))| < \mathbf{1} - \eta \tag{S-1}$$

where $\beta(1)$ is the non-zero coefficients for the relevant variables in the linear model, η is a positive constant vector and $sign(\cdot)$ maps positive entry to 1, negative entry to -1 and zero to zero. Denote the Lasso estimator of β by $\hat{\beta}$ and the one based on bootstrap data by $\tilde{\hat{\beta}}$. Also define the event

$$S \equiv \{sign(\hat{\beta}) = sign(\beta)\}$$

and

$$\tilde{S} \equiv \{sign(\tilde{\hat{\beta}}) = sign(\beta)\}.$$

Note that $\tilde{\cdot}$ *is used to represent the counterpart in the bootstrap sample space to that in the sample space. Below we will show that* $P(S) \to 1$ *implies* $P(\tilde{S}) \to 1$*, i.e., the Lasso procedure is also consistent on bootstrap resample data. Thus, denote the selection probability of the* i^{th} *feature w.r.t. the sample space by* p_i *and that w.r.t. the bootstrap resample space by* $\tilde{p}_i$*, then* p_i *and* $\tilde{p}_i$ *converge to the same limit (1 or 0, depending on whether the* i^{th} *feature is a true or irrelevant one) for all* $1 \leq i < p$.

We use the notation consistent with Zhao and Yu (2006). First we see that, under the finite-moment assumption of $\mathbf{X}$*, both the sample covariance* C^n *and bootstrap resample covariance* $\tilde{C^n}$ *converge to the same limit* C *(Arenal-Gutierrez et al. 1996), which means Proposition 1 in Zhao and Yu (2006) can be applied to the bootstrap resample data. Then,*

$$1 - P(\tilde{S}) \leq \sum_{i=1}^{q} P\left(|\tilde{z_i^n}| \geq \sqrt{n}(|\beta_i| - \frac{\lambda_n}{2n}\tilde{b_i^n})\right) + \sum_{i=1}^{p-q} P\left(|\tilde{\zeta_i^n}| \geq \frac{\lambda_n}{2\sqrt{n}}\tilde{\eta}_i\right)$$

where $(\tilde{z_1^n}, \ldots, \tilde{z_q^n}, \tilde{\zeta_1^n}, \ldots, \tilde{\zeta_{p-q}^n})' = \tilde{D^n}\tilde{W^n}$ *with*

$$\tilde{D^n} = \begin{pmatrix} (\tilde{C_{11}^n})^{-1} & \mathbf{0} \\ \tilde{C_{21}^n}(\tilde{C_{11}^n})^{-1} & -\mathbf{1} \end{pmatrix},$$

$\tilde{W^n} = \tilde{\mathbf{X}}'\tilde{\epsilon}/\sqrt{n}$*, and* $\tilde{b} = (\tilde{b_1^n}, \ldots, \tilde{b_q^n}) = (\tilde{C_{11}^n})^{-1} sign\beta(1)$.

Denote the counterpart of $\tilde{W}^n$ *and* $\tilde{D}^n$ *w.r.t. the sample space by* W^n *and* D^n*. Note that* $W^n \to_d N(\mathbf{0}, C)$ *and by Theorem 2.2 of Bickel and Freedman (1981)* $\tilde{W^n} - W^n \to_d N(\mathbf{0}, C)$*. Also note that* $\tilde{D}^n - D^n \to 0$ *and* $D^n \to D$ *where*

$$D = \begin{pmatrix} (C_{11})^{-1} & \mathbf{0} \\ C_{21}(C_{11})^{-1} & -\mathbf{1} \end{pmatrix}.$$

By the Slutsky's Theorem,

$$D^n W^n \to_d D \cdot N(0, C)$$

and

$$\tilde{D}^n \tilde{W}^n - D^n W^n = \tilde{D^n}(\tilde{W}^n - W^n) + (\tilde{D^n} - D^n)W^n \to_d D \cdot N(0, C),$$

which implies that $\tilde{z_i^n} - z_i^n$ *and* z_i^n *have the same limiting distribution. Thus, for* λ_n *such that*

$\lambda_n/n \to 0,\ \lambda_n/n^{\frac{1+c}{2}} \to \infty$ *with* $0 \le c < 1$,

$$\begin{aligned}
& \sum_{i=1}^{q} P\left(|\tilde{z_i^n}| \ge \sqrt{n}(|\beta_i| - \frac{\lambda_n}{2n}\tilde{b_i^n})\right) \\
= \;& \sum_{i=1}^{q} P\left(|\tilde{z_i^n} - z_i^n + z_i^n| \ge \sqrt{n}(|\beta_i| - \frac{\lambda_n}{2n}\tilde{b_i^n})\right) \\
\le \;& \sum_{i=1}^{q} P\left(|\tilde{z_i^n} - z_i^n| + |z_i^n| \ge \sqrt{n}(|\beta_i| - \frac{\lambda_n}{2n}\tilde{b_i^n})\right) \\
\le \;& \sum_{i=1}^{q} \left[P\left(|\tilde{z_i^n} - z_i^n| \ge \frac{1}{2}\sqrt{n}(|\beta_i| - \frac{\lambda_n}{2n}\tilde{b_i^n})\right) + P\left(|z_i^n| \ge \frac{1}{2}\sqrt{n}(|\beta_i| - \frac{\lambda_n}{2n}\tilde{b_i^n})\right)\right] \quad \text{(S-2)} \\
= \;& o(e^{-n^c}), \quad \text{(S-3)}
\end{aligned}$$

where (S-3) uses the result from the proof of Theorem 1 in Zhao and Yu (2006) while for (S-2) it is because $P(Z_1 + Z_2 \ge t) \le P(max(Z_1, Z_2) + max(Z_1, Z_2) \ge t) = P(Z_1 \ge t/2 \text{ or } Z_2 \ge t/2) \le P(Z_1 \ge t/2) + P(Z_2 \ge t/2)$. *Similarly, it can be shown that*

$$\sum_{i=1}^{p-q} P\left(|\tilde{\zeta_i^n}| \ge \frac{\lambda_n}{2\sqrt{n}}\tilde{\eta}_i\right) = o(e^{-n^c}).$$

Therefore, $P(\tilde{S}) \to 1$.

Analogues to the above, it can be shown that lasso is consistent w.r.t. both the sample and the bootstrap resample space under (S-1) and additional regularity conditions when p *is allowed to grow as* n *grows.*

Example 3. *Space procedure (Peng et al., 2009) with selection consistency for network construction.*

Similar to the irrepresentable condition for the lasso regression case, the selection consistency for the space procedure is implied by a condition imposed on the second derivative of the objective loss function which converges to the same limit for both sample data and bootstrap resample data. Under this condition and additional regularity conditions, the probability of space procedure being inconsistent based on bootstrap resamples can be bounded above by a small number in a similar way as the bound based on samples, which implies that the space procedure is also consistent w.r.t. the bootstrap resmaple space and hence $p_{ij} - \tilde{p}_{ij} \to 0$ *for all* $(i, j) \in \Omega$.

Part D: Other Simulation Results

D1: An Example where BINCO's FDR Cannot be Controlled at Stringent Levels when the Valley Point Value is too Large.

We simulate data from the power-law network as in Fig 5(a), but the signal strength is set at the "very weak" level (see details in the simulation where we investigate the effect of signal strength on BINCO in Section 3.2). The selection frequencies are generated by applying *space* on bootstrap resamples from the simulated data. In this example, the empirical selection frequency distribution is not U-shaped (Fig S-4) with a valley point value (0.96) greater than the threshold (0.8) set in Step 2.3 in the U-shape Detection Procedure. The smallest FDR of an aggregation-based selection $S_c^\lambda = \{(i,j) : X_{ij}^\lambda \geq c\}$ is 0.07, achieved at $c = 1$, which means the FDR of BINCO's slelction can not be less than 0.07.

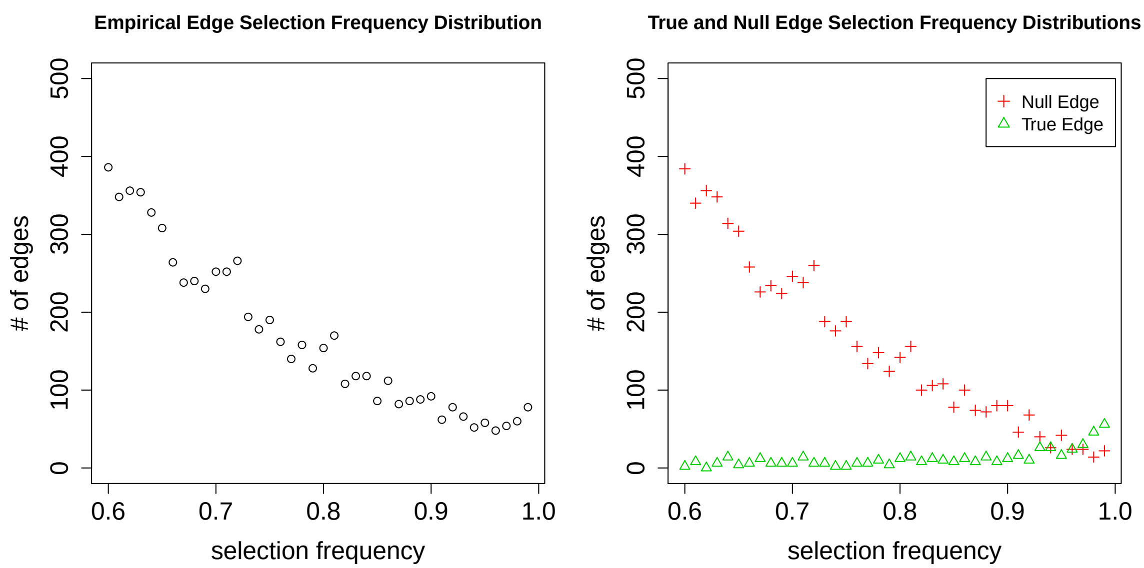

Fig S-4: Non-U-shaped empirical distribution of selection frequencies where the FDR of BINCO's selection can not be less than 0.07. The selection frequencies are generated under $\lambda = 30$

D2: An Example where Stability Selection Fails to Control False Positives.

In this example we simulate selection frequencies from a setting where the exchangeability assumption needed by *stability selection* is violated. Specifically, we generate 20 independent samples of selection frequencies under the following setting:

(a) Consider 100,000 candidate variables, of which 99,500 are null variables and 500 are true variables.

(b) The selection probabilities for first 90,000 null variables are set to be $p_0 = 0.01$. The distribution of selection probabilities (denoted by $F(p)$) of the other 9,500 null variables is $F(p) = Pr(p_{ij} \le p) = \sqrt{p}$, such that, e.g., 10% of 9,500 null variables (less than 1% of total null edges) have selection probabilities greater than 0.81 but less than 1.

(c) The selection probabilities for all 500 true variables are set to be $p_1 = 0.99$.

(d) We repeat this for $B = 100$.

The empirical mixture densities are all U-shaped (Fig S-5), although the valley point value is large. We observe (Fig S-6) that BINCO provides reasonable estimates for the null distribution and hence controls the FDR well, but the *stability selection* fails to control the false positives (Fig S-7).

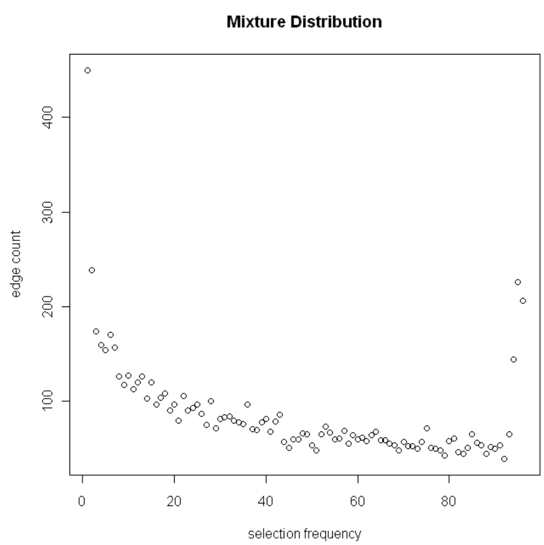

Fig S-5: A typical empirical mixture distribution of selection frequencies.

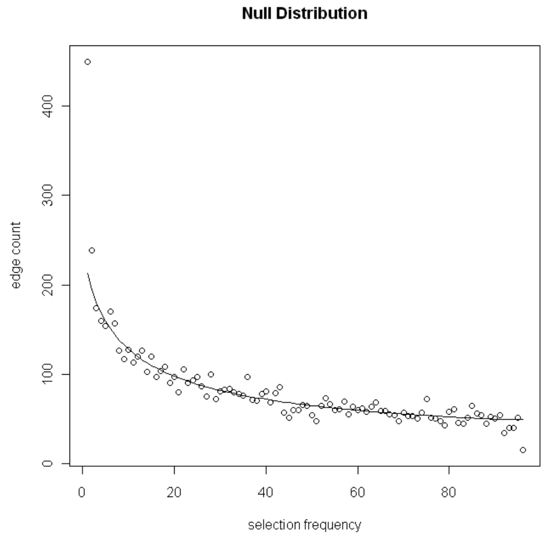

Fig S-6: The null distribution (in dots) is well estimated by BINCO (solid line).

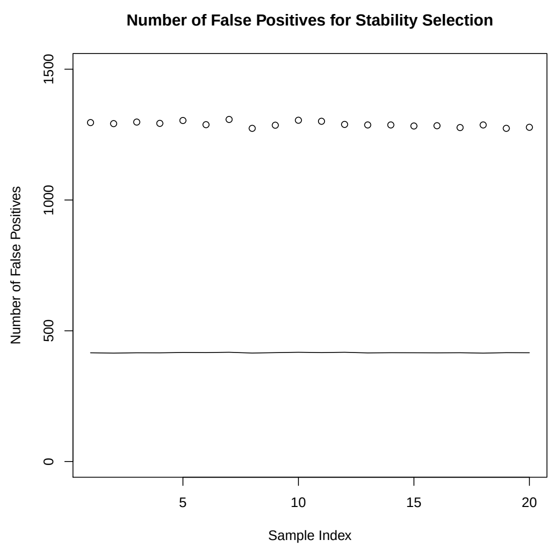

Fig S-7: The actual number of false positives (around 1300, dots) is significantly larger than the theoretical upper bound (below 500, solid line) suggested by the *stability selection* method, for all 20 samples.

D3: Correlations between Edges are Present in both Simulated and Real Data.

There are correlations between edges in both simulated and real data (see Fig S-8), and the correlation distributions for both are similar.

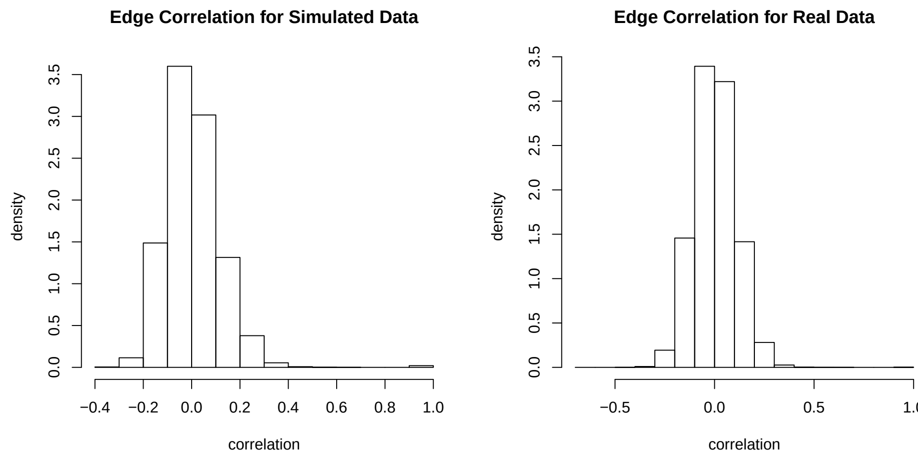

Fig S-8: Distributions of correlation between edges for a simulated data (left panel, which is the same data set used in Figs 1-3 of the main text) and for the real data (right panel, the BC data analyzed in Section 4 of the main text). The mean and MAD for the simulated data are 0.005 and 0.111, respectively, and the mean and MAD for the real data are 0.002 and 0.111, respectively.

D4: U-Shaped Empirical Distributions of Selection Frequencies Generated from Non-normal Data.

We simulate two non-normal data sets as follows. First we apply Cholesky decomposition on the correlation matrix (Σ), which we used to generate the normal data for the power-law network in Section 3, i.e., we decompose Σ as $\Sigma = LL^*$ where L is an lower triangular matrix with strictly positive diagonal entries and L^* is the conjugate transpose of L. Then we simulate uncorrelated vector u following some non-normal marginal distribution (we use the t-distribution (df=5) for one data set and the uniform distribution on [-1,1] for the other) and apply L to this u to obtain the jointly non-normal data Lu. We apply *space* on the two simulated non-normal data sets and both yield U-shaped empirical distributions of selection frequencies (Figs S-9, S-10). Applying BINCO on these U-shaped distributions yields edge selection results with well-controlled FDR and decent power (details omitted).

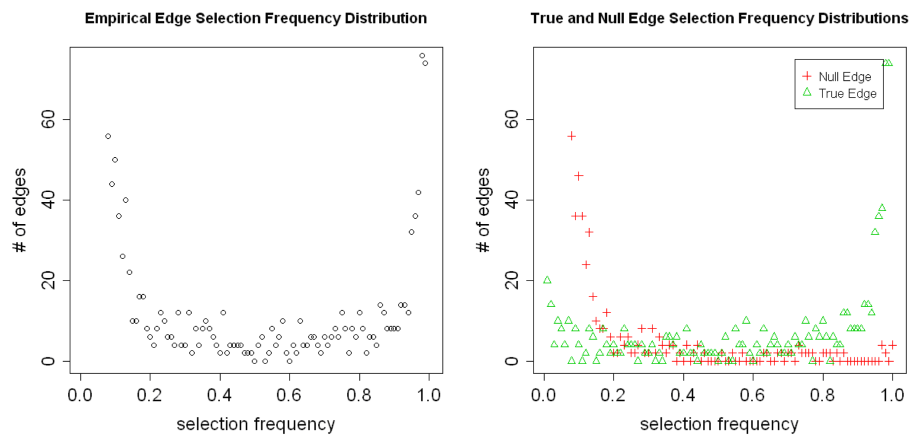

Fig S-9: Selection frequency distribution for data simulated from a t-distribution (df=5).

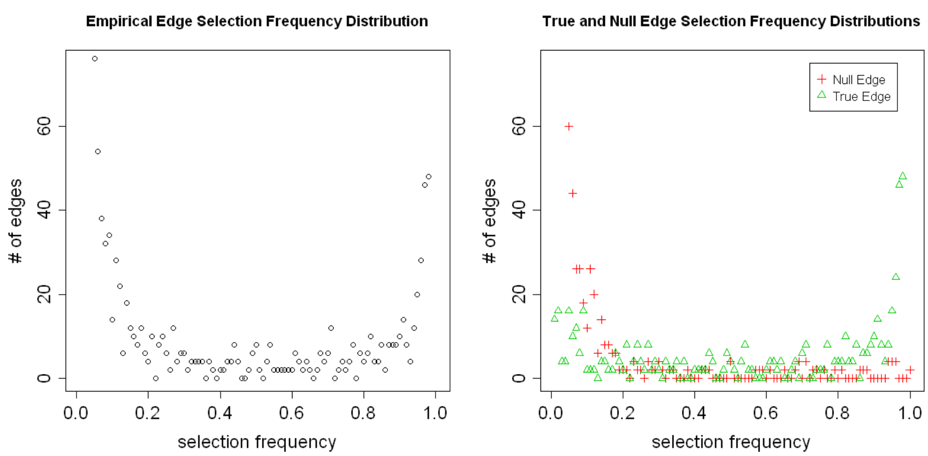

Fig S-10: Selection frequency distribution for data simulated based on uniform distribution.

References.

[1] Arenal-Gutierrez, E., Matran, C. and Cuesta-Albertos, J. (1996). On the unconditional strong law of large numbers for the bootstrap mean. *Statistics and Probability Letters*, **27**, 49-60.

[2] Bickel, P and Freedman, D (1981). Some asymptotic theory for the bootstrap. *Ann. Stat.*, **9, No. 6**, 1196-1217.

[3] Bild, A. and Johnson G (2001). Signaling by erbB receptors in breast cancer: regulation by compartmentalization of heterodimetric receptorComplexes. *Annual summary rept.*, **15 Sep. 1998 - 14 Sep. 2001**, http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA400019.

[4] Canbay E, Eraltan IY, Cercel A, Isbir T, Gazioglu E, Aydogan F, Cacina C, Cengiz A, Ferahman M, Zengin E and Unal H. (2010). CCND1 and CDKN1B polymorphisms and risk of breast cancer. *Anticancer Research*, **30**, 3093-3098.

[5] Gawenda J., Traub F., Lck H. J., Kreipe H. and von Wasielewski R. (2007). Legumain expression as a prognostic factor in breast cancer patients. *Breast Cancer Res Treat.* **102(1)**, 1-6.

[6] Jiang, P., Hu, Q., Ito, M., Meyer, S., Waltz, S., Khan, S., Roeder, R. G. and Zhang, X. (2010). Key roles for MED1 LxxLL motifs in pubertal mammary gland development andluminal-cell differentiation. *Proc Natl Acad Sci USA.* **107(15)**, 6765-70.

[7] Kattan, Z., Marchal, S., Brunner, E., Ramacci, C., Leroux, A., Merlin, J. L., Domenjoud, L., Daua, M. and Becuwe, P. (2008). Damaged DNA binding protein 2 plays a role in breast cancer cell growth. *PLoS ONE*, **3(4)**, e2002. doi:10.1371/journal.pone.0002002.

[8] Lee, S., Reimer, C., Fang, L., Iruela-Arispe, M. and Aaronson. S. (1999). Overexpression of Kinase-Associated Phosphatase (KAP) in breast and prostate cancer and inhibition of the transformed phenotype by antisense KAP expression. *Molecular and Cellular Biology*, **20, No.5**, 1723-1732.

[9] Ma, H., Jin, G., Hu, Z., Zhai, X., Chen, W., Wang, S., Wang, X., Qin, J., Gao, J., Liu, J., Wang, X., Wei, Q. and Shen, H. (2006). Variant genotypes of CDKN1A and CDKN1B are associated with an increased risk of breast cancer in Chinese women. *Int. J. Cancer*, **119**, 2173-2178.

[10] Peng, J., Wang, P., Zhou, N. and Zhu, J. (2009). Partial correlation estimation by joint sparse regression models, *JASA.*, **104(486)**, 735-746.

[11] Postel-Vinay, S., Véron, A., Tirode, F., Pierron, G., Reynaud, S., Kovar, H., Oberlin, O., Lapouble, E., Ballet, S., Lucchesi, C., Kontny, U., Gonzlez-Neira, A., Picci, P., Alonso, J., Patino-Garcia, A., Bressac de Paillerets, B., Laud, K., Dina, C., Froguel, P., Clavel-Chapelon, F., Doz, F., Michon, J., Chanock, S., Thomas, G., Cox, D. and Delattre, O. (2012). Common variants near TARDBP and EGR2 are associated with susceptibility to Ewing sarcoma. *Nature Genetics* **44**, 323-327.

[12] Romieu-Mourez, R., Landesman-Bollag, E., Seldin, D., Traish, A., Mercurio, F. and Sonenshein, G. (2001). Roles of IKK kinases and protein kinase CK2 in activation of nuclear factor-kB in breast cancer. *Cancer Res.*, **2001,61**, 3810-3818.

[13] Wang, K., Ye, Y., Xu, Z., Zhang, X., Hou, Z., Cui, Y. and Song, Y. (2010). Interaction between BRCA1/BRCA2 and ATM/ATR associate with breast cancer susceptibility in a Chinese Han population. *Cancer Genetics and Cytogenetics*, **200(1)**, 40-6.

[14] Zhao, P. and Yu, B. (2006). On model selection consistency of lasso. *Journal of Machine Learning Research*, **7**, 2541-2563.

[15] Zhu, Y., Qi, C., Jain, S., Le-Beau, M., Espinosa III, R., Atkins, B., Lazar, M., Yeldandi, A., Rao, S. and Reddy, J. (1999). Amplification and overexpression of peroxisome proliferator-activated receptor binding protein (PBP/PPARBP) gene in breast cancer. *Proc Natl Acad Sci USA.*, **96**, 10848-10853.