

A Computational Approach inspired by Simulated Annealing to study the Stability of Protein Interaction Networks in Cancer and Neurological Disorders

Data Mining and Knowledge Discovery

Kristina Ibáñez^{1,2}, María Guijarro¹, Gonzalo Pajares¹, Alfonso Valencia^{2,§}

¹Departamento Ingeniería del Software e Inteligencia Artificial, Facultad Informática, Universidad Complutense, 28040 Madrid, Spain

²Structural Biology and Biocomputing Programme, Spanish National Cancer Research Centre (CNIO), 28029 Madrid, Spain

[§] **Corresponding author:** Dr. Alfonso Valencia. Tel: +34 917328000; Fax: +34 912246976; E-mail: avalencia@cniio.es

Gene expression datasets

The following datasets were downloaded from The Gene Expression Barcode ([Barcode, http://barcode.luhs.org](http://barcode.luhs.org)), the human transcriptome repository that provides reliable absolute measures of expression for most annotated genes for 131 human tissue types, including diseased tissue.

Ovary Cancer

- Cancer : GSE9891 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE9891>)

GSM38064, GSM38065 , GSM38066, GSM38070, GSM38071, GSM38088, GSM38095, GSM46815, GSM46821, GSM46830, GSM46831, GSM46839, GSM46853, GSM46886, GSM46897, GSM46898, GSM46910, GSM46911, GSM46918, GSM46925 (GSE2109, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE2109>)

GSM249675, GSM249676, GSM249677, GSM249678 (GSE9890, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE9890>).

- Control: GSM80757, GSM80758, GSM80759, GSM80780 (GSE3526, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE3526>). GSM175789, GSM176131, GSM176136, GSM176237, GSM176318 (GSE7307, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE7307>)

Platform: Affymetrix Human Genome U133 Plus 2.0 Array

Colon Cancer

- Cancer: GSM38055, GSM38061, GSM38074, GSM38075, GSM38089, GSM38105, GSM38107, GSM46823, GSM46832, GSM46841, GSM46845, GSM46857, GSM46861, GSM46864, GSM46865, GSM46877, GSM46878, GSM46887, GSM46895, GSM46899, GSM46915, GSM46921, GSM46924, GSM89103 (GSE2109, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE2109>)

- Normal: GSM95473, GSM95474, GSM95475, GSM95476, GSM95477, GSM95478, GSM95479,

GSM95480, , (GSE4183, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE4183>)

Platform: Affymetrix Human Genome U133 Plus 2.0 Array

● **Cancer stages:** Metastasis Gene Expression Profile Predicts Recurrence and Death in Colon Cancer Patients

(<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE17537>)

Stage I: 4 samples

Stage II: 10 samples

Stage III: 18 samples

Stage IV: 11 samples

Platform: Affymetrix Human Genome U133 Plus 2.0 Array

Liver Cancer

● **Cancer:** GSM38078, GSM38108, GSM46848 (GSE2109,

<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE2109>)

GSM139131, GSM143545, GSM143546, GSM143547, GSM143548, GSM143549, GSM143550, GSM143551, GSM143552, GSM143553 (GSE6222, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE6222>)

GSM248688, GSM248689, GSM248690, GSM248691, GSM248692, GSM248693, GSM248694, GSM248695, GSM248696, GSM248697, GSM248698, GSM248699, GSM248700, GSM248701, GSM248702, GSM248703, GSM248704, GSM248705, GSM248706, GSM248707, GSM248708, GSM248709, GSM248710, GSM248711, GSM248712, GSM248713, GSM248714, GSM248715, GSM248716, GSM248717, GSM248718, GSM248719, GSM248720, GSM248721, GSM248722, GSM248723, GSM248724, GSM248725, GSM248726, GSM248727, GSM248728, GSM248729, GSM248730, GSM248731, GSM248732, GSM248733, GSM248734, GSM248735, GSM248736, GSM248737, GSM248738, GSM248739, GSM248740, GSM248741, GSM248742, GSM248743,

GSM248744, GSM248745, GSM248746, GSM248747, GSM248748, GSM248749, GSM248750, GSM248751, GSM248752, GSM248753, GSM248754, GSM248755, GSM248756, GSM248757, GSM248758, GSM248759, GSM248760, GSM248761, GSM248762, GSM248763, GSM248764, GSM248765, GSM248766, GSM248767, GSM248768, GSM248769, GSM248770, GSM248771, GSM248772, GSM248773, GSM248774, GSM248775, GSM248776, GSM248777, GSM248778 (GSE9829;GSE9843, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE9829>, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE9843>)

- Normal: GSM80728, GSM80729, GSM80730, GSM80739, GSM138595, GSM138596, GSM155919, GSM155926, GSM155927, GSM155928, GSM155947, GSM155948, GSM155961, GSM155964, GSM155988, GSM155989 (GSE11045, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE11045>)

GSM176332, GSM176333, GSM176334, GSM176335, GSM279063, GSM279064, GSM279065 (GSE7307, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE7307>)

Platform: Affymetrix Human Genome U133 Plus 2.0 Array

- **Cancer stages**: Genome-wide molecular profiles of HCV-induced dysplasia and hepatocellular carcinoma GSE6764 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE6764>)

Normal: 13 samples

Very early HCC: 11 samples

Cirrhosis+HCC: 13 samples

Early HCC: 13 samples

Advanced HCC: 20 samples

Platform: Affymetrix Human Genome U133 Plus 2.0 Array

Kidney Cancer

- **Cancer:** GSM38073, GSM46825, GSM46826, GSM46847, GSM46858, GSM46875, GSM46881, GSM46882, GSM46892, GSM89104 (GSE2109, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE2109>)

GSM281278, GSM281279, GSM281280, GSM281281, GSM281282, GSM281283, GSM281284, GSM281285, GSM281286, GSM281287, GSM281288, GSM281289, GSM281290, GSM281291, GSM281292, GSM281293, GSM281294, GSM281295, GSM281296, GSM281297, GSM281298, GSM281299, GSM281300, GSM281301, GSM281302, GSM281303, GSM281304, GSM281305, GSM281306, GSM281307, GSM281308, GSM281309, GSM281310, GSM281313, GSM281317, GSM281318, GSM281319, GSM281320, GSM281321, GSM281322, GSM281323, GSM281324, GSM281325, GSM281326, GSM281327, GSM281328, GSM281329, GSM281330, GSM281331, GSM281332, GSM281333, GSM281334, GSM281335, GSM281336, GSM281337, GSM281338, GSM281339, GSM281340, GSM281341, GSM281342, GSM281343, GSM281344 (GSE11151, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE11151>)

GSM305099, GSM305100, GSM305101, GSM305102, GSM305103, GSM305104, GSM305105, GSM305106, GSM305107, GSM305108, GSM305109, GSM305110, GSM305111, GSM305112, GSM305113, GSM305114, GSM305115, GSM305116 (GSE12090, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE12090>)

- **Normal:** GSM198783, GSM198785 (GSE8050, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE8050>)

GSM240832, GSM240833, GSM240834, GSM240835, GSM240836, GSM240837, GSM240838, GSM240839, GSM240840, GSM240841, GSM240842, GSM240843, GSM240844, GSM240947, GSM240948

(GSE9489;GSE9493, <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE9489>;

<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE9493>)

GSM279060, GSM279061, GSM279062 (GSE11045,

<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE11045>)

GSM281311, GSM281312, GSM281314, GSM281315, GSM281316 (GSE11151,

<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE11151>)

Platform: Affymetrix Human Genome U133 Plus 2.0 Array

Alzheimer's disease

- GSE5281 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE5281>)

Entorhinal Cortex samples: 13 controls, 10 alzheimer

Hippocampus samples: 13 controls, 10 alzheimer

Posterior Singulate samples: 13 controls, 9 alzheimer

Primary Visual Cortex samples: 12 controls, 19 alzheimer

Superior Frontal Gyrus samples: 11 controls, 23 alzheimer

Platform: Affymetrix Human Genome U133 Plus 2.0 Array

- GSE1297 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE1297>)

Hippocampus samples

9 controls

15 alzheimer (Severe and Moderate)

Platform: Affymetrix Human Genome U133 Plus 2.0 Array

Schizophrenia

- GSE4036 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE4036>) 28 postmortem cerebellar tissues from schizophrenia patients and sex-, age-, and PMI-matched controls.

14 schizophrenia

14 controls

Platform: Affymetrix Human Genome U133 Plus 2.0 Array

The following datasets were downloaded from the Stanley Medical Research Institute, Online Genomics Database (SMRI, <https://www.stanleygenomics.org>) that manage 2 collections of brain samples (the Stanley Array collection and the Stanley Consortium collection). The different studies have been conducted using overlapping subsets of these collections.

- Dobrin (https://www.stanleygenomics.org/stanley/standard/studyDetail.jsp?study_id=5)

postmortem frontalBA46 cortice samples

30 schizophrenia

25 controls

Platform: Affymetrix Human Genome U133 Plus 2.0 Array

- Laeng (https://www.stanleygenomics.org/stanley/standard/studyDetail.jsp?study_id=17)41 postmortem

hippocampus (CA1)

20 schizophrenia

21 controls

Platform: Affymetrix Human Genome U133 Plus 2.0 Array

- Kemether (https://www.stanleygenomics.org/stanley/standard/studyDetail.jsp?study_id=16)26 postmortem

thalamus (MD) samples

14 schizophrenia

12 controls

Platform: Affymetrix Human Genome U133 Plus 2.0 Array