Cross-platform gene expression signature for microsatellite instability in colon and gastric cancers

FACULTY OF INFORMATICS
Masaryk University

Research centre for toxic compounds in the environment

MErCuRIC

Anna Pačínková[1,2*], Vlad Popovici[2]

[1] Faculty of Informatics, Masarykova Univerzita, Botanická 68a, 60200 Brno, Czech Republic

[2] Research Centre for Toxic Compounds in the Environment, Faculty of Science, Masarykova Univerzita, Kamenice 5, 62500 Brno, Czech Republic

* Correspondence: **anna.bilkova@recetox.muni.cz**

Objective:

- **To construct a cross-platform gene expression signature for microsatellite instability in both colon and gastric cancers.**

Introduction:

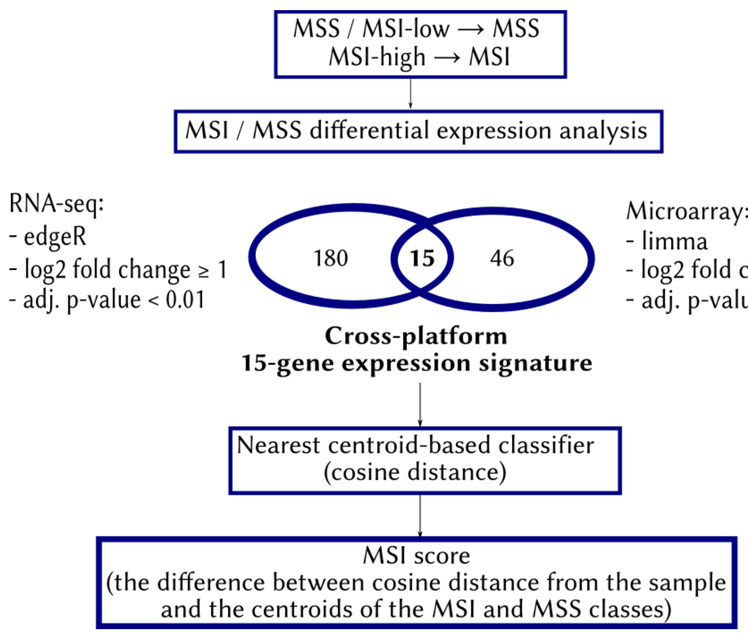
- The dysfunction of DNA mismatch repair system results in microsatellite instability (MSI).
- In colon cacner, MSI is a favourable prognosis marker, despite being associated with resistance to 5-fluorouracil treatment.
- In gastric cancer, its prognostic value is not so well established.
- Recognising the MSI tumours is of clear clinical importance and having **a single cross-platform gene expression signature** for both cancers has immediate practical benefits.

Methods:

Datasets

cohort	N (MSI)	platform	tissue	reference
A1 development	175 (35)	RNA-seq	colon	TCGA
A2 development	369 (56)	microarray	colon	GSE39582
B1 validation	83 (10)	microarray	colon	GSE39582
B2 validation	103 (20)	microarray	colon	GSE41258
C1 validation	335 (54)	RNA-seq	gastric	TCGA
C2 validation	34 (16)	microarray	gastric	GSE13911

Construction of the cross-platform gene expression signature for MSI status



(MSI: microasatellite instability;
MSS: microastellite stability)

Performance evaluation of the cross-platform 15-gene expression signature

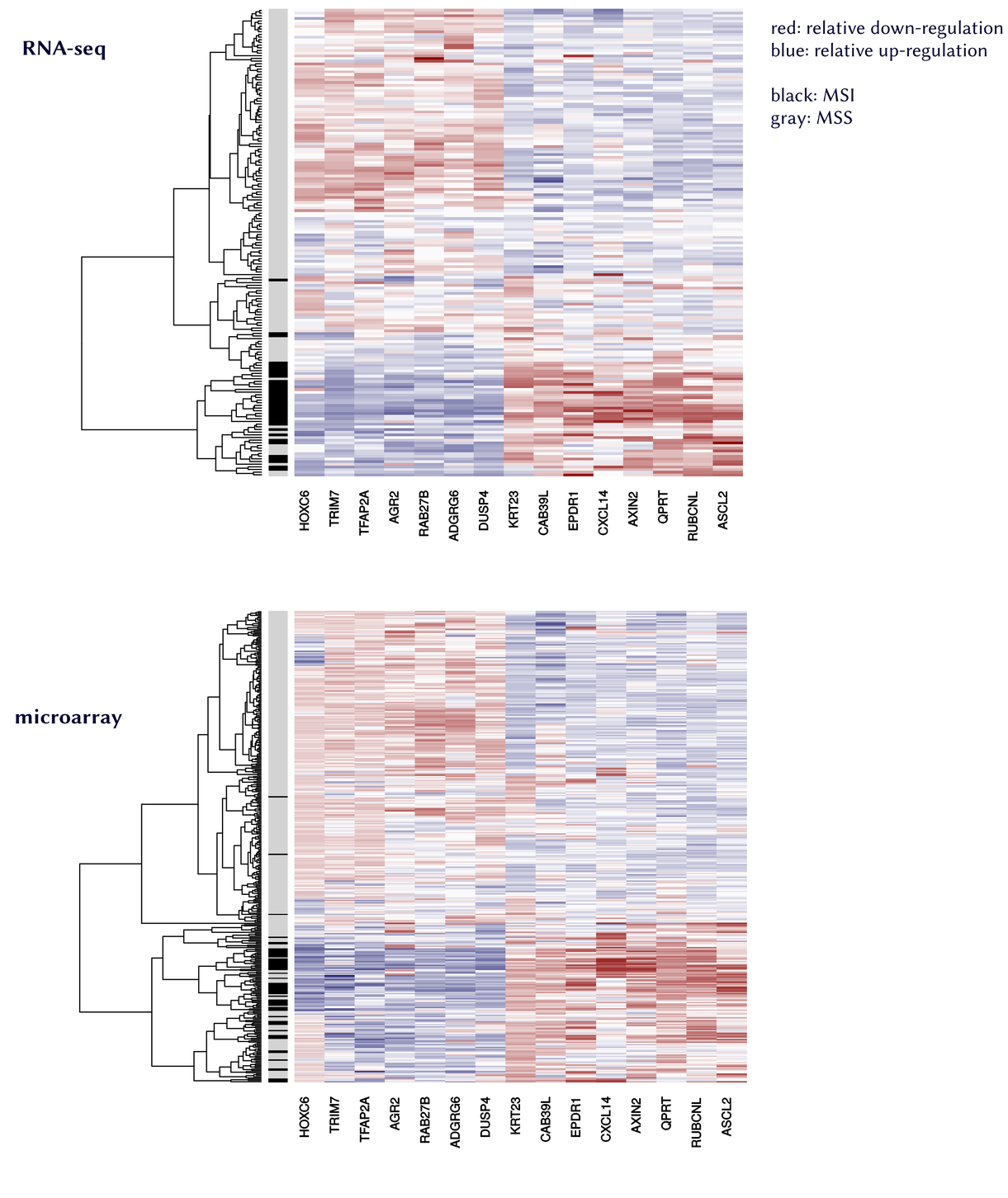
- The performance was estimated using 10-fold cross-validation.
- The performance index ⟶ an area under the receiver operating characteristic curve (AUC); 95 % confidence intervals (CI), normal approximation
- Validation was done using two independent colon cancer datasets (cohort B1 and cohort B2)
- The performance of the cross-platform gene expression signature was also evaluated on 369 **gastric cancer** samples (cohort C1 and cohort C2)
- Comparison with published MSI gene expression signatures trained exclusively on microarray datasets:
  - 8-gene signature[1] (two genes excluded from the nearest centroid classification)
  - 64-gene signature[2] (nine genes excluded from the nearest centroid classification)
- McNemar's test used to compare the accuracy of cross-platform 15-gene expression signature and published MSI gene signatures
- Correlation analysis was performed to detect a great deal of redundancy among genes from signatures (Spearman's correlation)

Pathway enrichment analysis and survival analysis

- The pathway enrichment analysis was performed against MSigDB gene collections
- The prognostic value of the cross-platform 15-gene expression signature was assessed by fitting Cox regression model:
  - relapse-free survival (RFS) in stage II and stage III cohort A2 sub-population (adjusted p-value < 0.1)

Results:

Cross-platform 15-gene expression signature for MSI status



Performance evaluation of the cross-platform 15-gene expression signature

cohort	AUC	95% CI	platform	tissue
A1 development	**0.93**	0.80 - 1	RNA-seq	colon
A2 development	**0.87**	0.62 - 1	microarray	colon
B1 validation	**0.93**	0.83 - 1	microarray	colon
B2 validation	**0.84**	0.73 - 0.96	microarray	colon
C1 validation	**0.86**	0.80 - 0.93	RNA-seq	gastric
C2 validation	**0.87**	0.74 - 1	microarray	gastric

- RNA-seq cohort A1: Significantly higher accuracy of the 15-gene signature than the 8-gene signature (p-value < 0.01)
- RNA-seq cohort A1: The assumptions of McNemar's test was not fulfilled!
  (Significantly higher accuracy of the 64-gene signature than the 15-gene signature (p-value 0.01529))
- 8 genes identified both in the 15-gene signature and in the 64-gene signature
- High redundancy of 64-gene signature in RNA-seq cohort A1 (14 gene-pairs with $|r| > 0.75$)
- Gene AXIN2 (HR=0.66, 95% CI=(0.50 – 0.87)) was significantly associated with RFS (adjusted p-value 0.076).

Pathway enrichment analysis and survival analysis

MsigDB gene set	adj. p-value	genes in overlap
ttgttt foxo4 01 [3]	0.056	DUSP4,ASCL2,KRT23,AGR2,CAB39L,HOXC6,AXIN2,TFAP2A
Ik2 01	0.002	HOXC6,ASCL2,TFAP2A,AXIN2,CXCL14
sansom apc targets [4]	0.010	ASCL2,AXIN2,KRT23,DUSP4
sansom apc targets up [4]	0.054	ASCL2,KRT23,AXIN2
koinuma colon cancer msi dn[5]	0.012	QPRT,AXIN2

Discussion and conclusion:

- We present the cross-platform 15-gene expression signature that yields high accuracy in MSI phenotype prediction in both RNA-seq and microarray studies.
- The pathway enrichment analysis results support the 15-gene expression signature association with colon cancer MSI phenotype.
- Despite being derived from colon cancer datasets, the signature maintained a good performance on gastric cancer datasets.
- From a biologic perspective, this supports the idea that MSI gene expression pattern is comparable across various cancers pointing towards similar regulatory pathways.

References:

[1] Giacomini, C.P., Leung, S.Y., Chen, X., Yuen, S.T., Kim, Y.H., Bair, E., Pollack, J.R.: A gene expression signature of genetic instability in colon cancer. American Association for Cancer Research 65 (2005)
[2] Tian, S., Roepman, P., Popovici, V., Michaut, M., Majewski, I., et al.: A robust genomic signature for the detection of colorectal cancer patients with microsatellite instability phenotype and high mutation frequency. Journal of Pathology 228 (2012)
[3] Xie, X., Kulbokas, E.J., Golub, T.R., Mootha, V., Lindblad-Toh, K., Lander, E.S., Kellis, M., et al.: Systematic discovery of regulatory motifs in human promoters and 3' UTRs by comparison of several mammals. Nature 434 (2005)
[4] Sansom, O.J., Reed K.R., Hayes, A.J., Ireland, H., Brinkmann, H., Newton, I.P., Batlle, E., Simon-Assmann, P., Clevers, H., Nathke, I.S., Clarke, A.R., Winton, D.J.: Loss of Apc in vivo immediately perturbs Wnt signaling, differentiation, and migration. Genes Dev. 18 (2004)
[5] Koinuma, K., Yamashita, Y., Liu, W., Hatanaka, H., Kurashina, K., Wada, T., Takada, S., Kaneda, R., Choi, Y.L., Fujiwara, S.I., Miyakura, Y., Nagai, H., Mano, H.: Epigenetic silencing of axin2 in colorectal carcinoma with microsatellite instability. Oncogene 25 (2006)

Acknowledgements:

This research was supported by the RECETOX Research Infrastructure (LM2015051 and CZ.02.1.01/0.0/0.0/16 013/0001761) and by the European Community's Seventh Framework Programme under grant agreement no. 602901 MerCuRIC. Access to computing and storage facilities owned by parties and projects contributing to the National Grid Infrastructure MetaCentrum provided under the programme "Projects of Large Research, Development, and Innovations Infrastructures" (CESNET LM2015042), is greatly appreciated.