

Expression Correlation and Copy Number Alteration(CNA) Prevalence of Human RNASEH2A in cancer supports a role for *RNASEH2A* in cancer proliferation



Deepali L. Kundnani, Stefania Marsili, Ailone Tichon and Francesca Storici
School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, USA

INTRODUCTION

Ribonuclease (RNase) H2, a key enzyme for the removal of RNA found in DNA-RNA hybrids and is composed of three subunits with **RNASEH2A** as its catalytic subunit.

Differential cell-cycle regulation of the RNASEH2A orthologous gene has also been observed in yeast *Saccharomyces cerevisiae*. Previous studies also report increase in levels of RNASEH2A expression in multiple cancers, and increased expression with overexpression of several oncogenes in mesenchymal stem cells.

Finding **co-expressed genes** with RNASEH2A can reveal other genes involved with RNASEH2A either in common or interconnected biological processes. In this study, we work with a selected list of genes involved different cell cycle and cancer related markers for correlation analysis in large datasets to ask specific questions. Use of clustering methods further helps us group genes with similar expression trends closer to each other.

Copy Number Alterations(CNAs) are another way to confirm increases/decrease in expression in cancers. CNAs can be categorized into the following;

- Deep deletion (-2) indicating a deep loss/homozygous deletion;
- Shallow deletion (-1) indicates shallow loss/heterozygous deletion;
- Diploid (0) indicates homozygous genes;
- Gain indicates a low-level gain (a few additional copies, often broad);
- Amplification (2) indicate a high-level amplification (more copies, often focal).

OBJECTIVES

1. Find suggestive role of RNASEH2A in human tissues by using RNASEH2A co-expressed genes for gene ontology analysis
2. To elucidate the role of RNASEH2A in different cell cycle phases and in cancer via expression correlation analysis with marker genes
3. To validate RNASEH2A involvement in cancer by measuring the prevalence of CNAs of RNASEH2A in different cancer subtypes

MATERIAL (DATASETS)



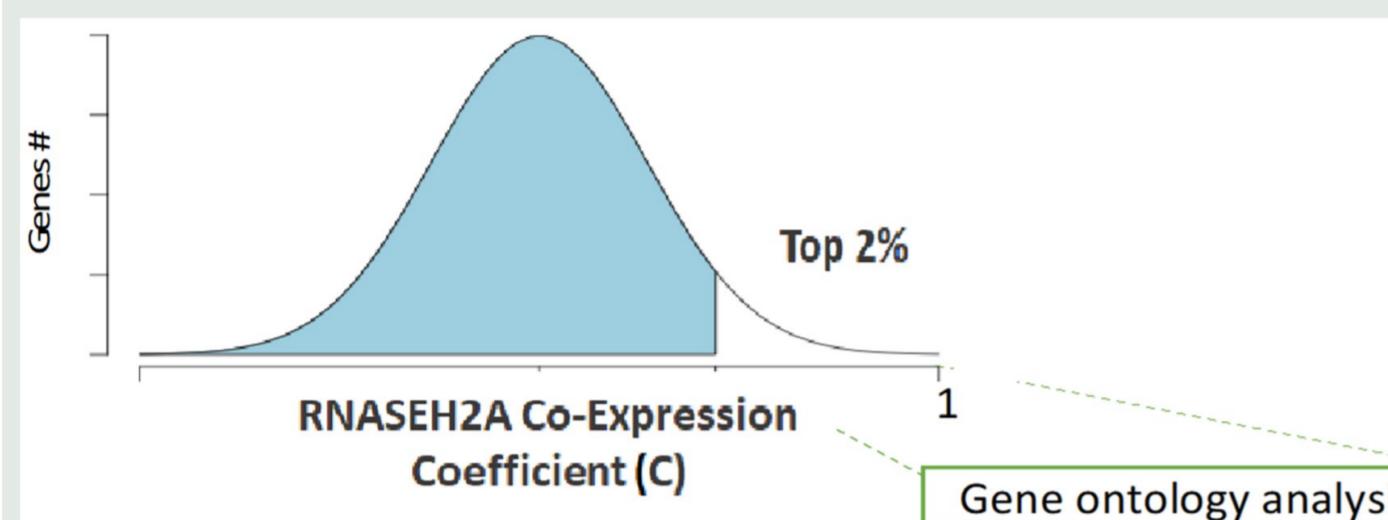
1. RNA-seq expression data in Transcripts per million from **Genotype-Tissue Expression (GTEx)** portal v7 is available for 53 different human tissues



2. RNA-Seq data expression from **Broad Institute Cancer Cell Line Encyclopedia (CCLE)** dataset in 1019 cancer cell lines from 26 different tissues of origin



3. Copy number alterations (CNAs) and RNA-Seq data from **The Cancer Genome Atlas (TCGA) Pan Cancer studies** involving 32 studies and 10,967 patients



Selection of gene pool of high positive correlation with RNASEH2A to find RNASEH2A associated processes in Human Tissues

Pearson's Correlation coefficient was calculated for expressions values of each gene with expression values of RNASEH2A, termed as RNASEH2A Co-expression Coefficient(C). Top 2% of genes showing highest co-expression coefficient were used for Gene Ontology analysis

METHODOLOGY

Gene name(s)	Function/Role	References
CDKN2A, CDKN2B, CCND1, DHFR, CCNE1	G1 Cell Cycle Phase	A. Subramanian et al. (GSEA Database)
AKT1-3, E2F4-5	G1/S Cell Cycle Phase	A. Subramanian et al. (GSEA Database)
CDKN2D, MDM2	S Cell Cycle Phase	A. Subramanian et al. (GSEA Database)
CCNB2, TOPBP1	G2/M Cell Cycle Phase	A. Subramanian et al. (GSEA Database)
APC, BUB1	M Cell Cycle Phase	A. Subramanian et al. (GSEA Database)
E2F2, E2F3, CCNB1	M/G1 Cell Cycle Phase	A. Subramanian et al. (GSEA Database)
MYBL2, FOXM1, BUB1, AURKA, AURKB	Upregulated in cancer	A. Subramanian et al. (GSEA Database)
SCARAS, MYOM1	Downregulated in cancer	A. Subramanian et al. (GSEA Database)
PCNA, MKI67(K167), MCM2-MCM6, E2F1	Proliferative markers in cancer	M. Li et al
CCNE1, CCND1, CCNB1	Cell cycle markers associated with cancer	M. Li et al
RNASEH2A, RNASEH2B, RNASEH2C RNASEH1	Target genes in this study	M. L. Whitfield et al.
CCT8, DNAAJ1, AIFM1	Predicted Binding partners of RNASEH2A	This study

Selected list of genes related to cancer proliferation and cell cycle phases

The above list contains 1. Markers involved/associated in different cell cycle phases, 2. Markers up/down regulated in cancer, 3. Cell cycle associated cancer markers, 4. Genes of our interest(RNASEH1, subunits of RNASEH2)

Correlation analysis in CCLE and TCGA expression datasets

Correlation coefficient analysis was between all 40 genes using Pearson correlation and hierarchical clustering to group genes based on similar trend of expression in the given set of cell lines/samples.

Prevalence of Copy Number Alterations calculation in TCGA Pan Cancer data

Prevalence of Copy Number Variations and average RNASEH2A was calculated for patients with every CNA type in every cancer subtype.

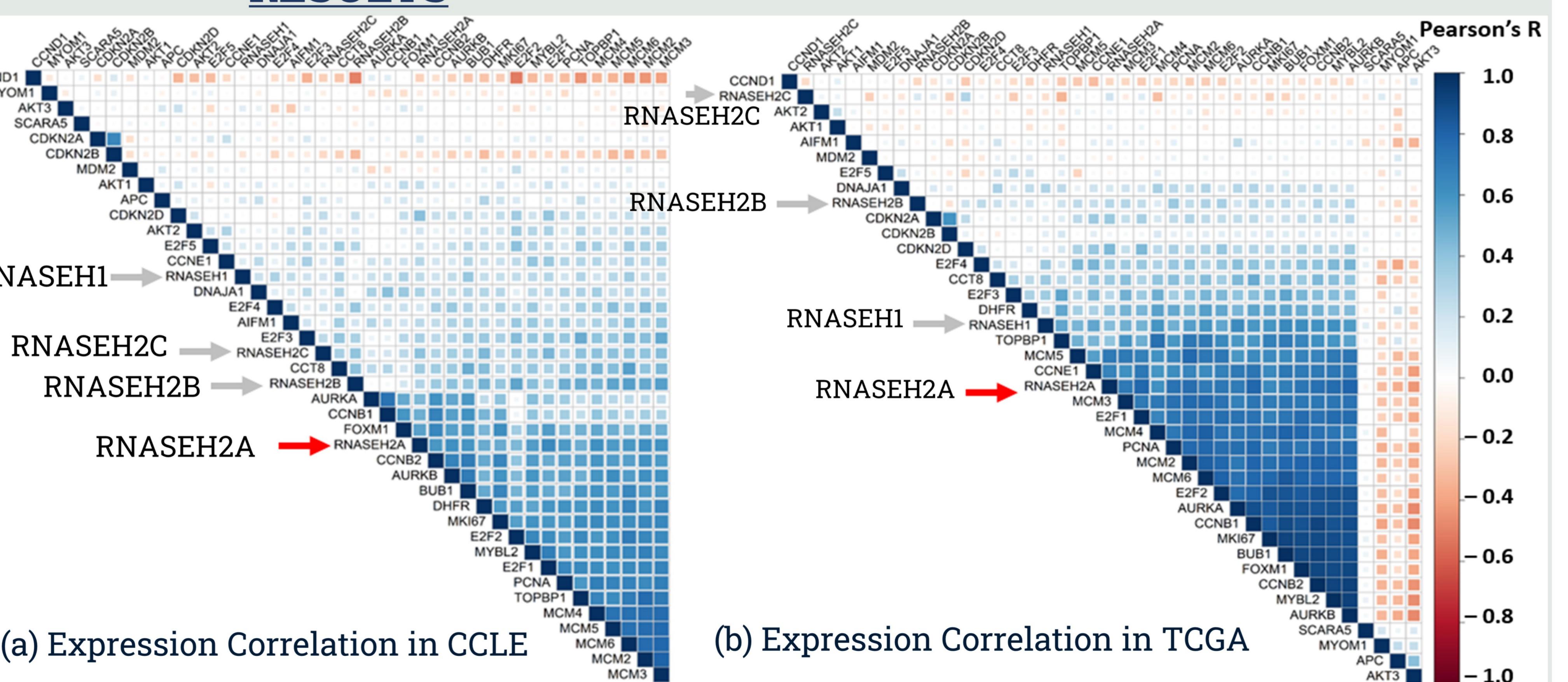
CONCLUSION

1. GO analysis of highly co-expressed with RNASEH2A in human tissues and correlation analysis in CCLE and TCGA concludes **RNASEH2A expression association with mitotic cell cycle regulation**
2. Correlation analysis of RNASEH2A with 39 other genes shows that RNASEH2A positively correlates up regulated genes in cancer, and negative correlation with down regulated genes in cancer implying **RNASEH2A plays a role in cancer proliferation**
3. Higher prevalence of RNASEH2A amplification vs Deep Deletion in various cancer subtypes support and validate **higher level of RNASEH2A is supportive of cancer**

REFERENCES

1. Marsili, S.; Tichon, A.; Kundnani, D.; Storici, F. Gene Co-Expression Analysis of Human RNASEH2A Reveals Functional Networks Associated with DNA Replication, DNA Damage Response, and Cell Cycle Regulation. *Biology* 2021, 10, 221. [\[CrossRef\]](#)
2. Reijns, M.A.; Rabe, B.; Rigby, R.E.; Mill, P.; Astell, K.R.; Lettice, L.A.; Boyle, S.; Leitch, A.; Keighren, M.; Kilanowski, F.; et al. Enzymatic removal of ribonucleotides from DNA is essential for mammalian genome integrity and development. *Cell* 2012, 149, 1008–1022. [\[CrossRef\]](#)
3. GTEx Consortium. The genotype-tissue expression (GTEx) project. *Nat. Genet.* 2013, 45, 580–585. [\[CrossRef\]](#)
4. Liu, J.; Lichtenberg, T.; Hoadley, K.A.; Poisson, L.M.; Lazar, A.J.; Cherniack, A.D.; Kovatch, A.J.; Benz, C.C.; Levine, D.A.; Lee, A.V.; et al. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell* 2018, 173, 400–416.e11. [\[CrossRef\]](#)
5. Ghandi, M.; Huang, F.W.; Jané-Valbuena, J.; Kryukov, G.V.; Lo, C.C.; McDonald, E.R., 3rd; Barretina, J.; Gelfand, E.T.; Bielski, C.M.; Li, H.; et al. Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* 2019, 569, 503–508. [\[CrossRef\]](#)
6. Shao, X.; Lv, N.; Liao, J.; Long, J.; Xue, R.; Ai, N.; Xu, D.; Fan, X. Copy number variation is highly correlated with differential gene expression: A pan-cancer study. *BMC Med. Genet.* 2019, 20, 175. [\[CrossRef\]](#)
7. Subramanian, A.; Tamayo, P.; Mootha, V.K.; Mukherjee, S.; Ebert, B.L.; Gillette, M.A.; Paulovich, A.; Pomery, S.L.; Golub, T.R.; Lander, E.S.; et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA* 2005, 102, 15545–15550. [\[CrossRef\]](#)
8. Whitfield, M.L.; George, L.K.; Grant, G.D.; Perou, C.M. Common markers of proliferation. *Nat. Rev. Cancer.* 2006, 2, 99–106. [\[CrossRef\]](#)

RESULTS



2. Correlation Plots of selected genes in CCLE(a) and TCGA Pan Cancer(b) Datasets with hierarchical clustering to reveal **close association of RNASEH2A (in red)** (as opposed to RNASEH2B, RNASEH2C and RNASEH1) **with cancer Proliferation and G2 and M Cell cycle markers**

Cancer Subtype(s)	Prevalence of Copy Number Alterations(CNAs) of RNASEH2A gene					Average RNASEH2A mRNA expression each CNA group, RSEM				
	Deep Deletion	Shallow Deletion	Diploid	Gain	Amplification	Deep Deletion	Shallow Deletion	Diploid	Gain	Amplification
Ovarian Epithelial Tumor	0.34%	29.15%	26.10%	36.27%	8.14%	649.61	730.15	1150.31	1444.92	2435.81
Endometrial Carcinoma		12.13%	69.67%	14.90%	3.29%		838.32	912.12	1721.50	2087.09
Adrenocortical Carcinoma		1.32%	34.21%	61.84%	2.63%		707.14	480.63	795.96	1114.81
Pleural Mesothelioma		8.05%	73.56%	16.09%	2.30%		631.88	534.82	975.79	1203.17
Esophageal Squamous Cell Carcinoma		34.04%	44.68%	19.15%	2.13%		775.80	697.66	974.64	2097.29
Cervical Squamous Cell Carcinoma	0.41%	26.23%	57.79%	13.52%	2.05%	1700.49	1289.19	1556.32	2079.49	7219.93
Diffuse Glioma	0.20%	3.33%	73.92%	20.78%	1.76%	205.72	413.52	351.50	461.54	404.56
Sarcoma		10.36%	50.20%	37.85%	1.59%		812.55	808.58	1356.59	1326.82
Invasive Breast Carcinoma	0.09%	21.25%	60.11%	17.23%	1.31%	384.40	618.31	615.99	940.40	1485.75
Ocular Melanoma		3.75%	92.50%	2.50%	1.25%		731.66	617.38	1234.06	798.60

3. Prevalence and average mRNA expression in patient samples with different Copy number variations for RNASEH2A gene in cancer subtypes having greater than 1% amplification. Average expression increases in patients found with amplified copy of RNASEH2A gene in comparison to average expression of patients with diploid RNASEH2A gene subtype. **Prevalence of Amplification of RNASEH2A gene copy is seen to be higher than prevalence of Deep Deletion.**

ACKNOWLEDGEMENTS/CONTACT

This research was funded by the National Institute of Health, NIH, NIGMS R01 GM115927 (F.S.); NIH NIEHS R01ES026243 (F.S.); the National Science Foundation, NSF, MCB-1615335 (F.S.) and the Howard Hughes Medical Institute Faculty Scholar grant 55108574 (F.S.).

Deepali Kundnani: dkundnani3@gatech.edu (D.K.)
Francesca Storici: storici@gatech.edu (F.S.)