

Beyond the coding genome: non-coding mutations and cancer

Kaivalya Walavalkar¹, Dimple Notani¹

¹Department of Cellular Organization and Signaling, National Centre for Biological Sciences, Tata Institute for Fundamental Research, Bangalore 560065, India

TABLE OF CONTENTS

1. Abstract
2. Introduction
3. Link between mutations in enhancers and cancer
4. Mutations in insulators/boundaries
5. Mutations in chromatin architecture proteins
6. Mutations in transcriptional machineries
7. eRNAs in cancer
8. Conclusions
9. Acknowledgments
10. References

1. ABSTRACT

Latest advancements in genomics involving individuals from different races and geographical locations has led to the identification of thousands of common as well as rare genetic variants and copy number variations (CNVs). These studies have surprisingly revealed that the majority of genetic variation is not present within the coding region but rather in the non-coding region of the genome, which is also termed as "Medical Genome". This short review describes how mutations/variations within; regulatory sequences, architectural proteins and transcriptional regulators give rise to the aberrant gene expression profiles that drives cellular transformations and malignancies.

2. INTRODUCTION

Unlike other patho-physiologies, understanding the role of mutations in cancer requires the knowledge of both germline as well as somatic mutations in the given cancer type. To this regard, several consortia have extensively catalogued both germline and somatic mutations using a variety of tools like genotyping arrays, exome-sequencing and whole-genome

sequencing (WGS). To identify the functional mutations amongst the thousands of mutations discovered from such studies, the presence of mutation being in coding or non-coding region is crucial further, knowledge of chromatin signatures is necessary in case of non-coding mutations as, the functional mutations in the non-coding genome affect coding gene regulation. Towards this, there are ever-growing and evolving tools to assess such signatures including but not restricted to, DNase-hypersensitivity (DHS-seq), histone marks, transcription factor binding alterations (ChIP-seq), transcriptional changes (RNA-seq and GRO-seq) and involvement of chromatin looping (ChIA-PET and 3C-based tools) are routinely used. Further functional characterization by which these variants mediate their effects on target genes is done using several tools such as reporter assays in mouse or relevant cell lines and/or genetic manipulations of target sequences to decipher the mechanism in which the variant affects target gene expression. Systems approaches have also been undertaken on these mutations obtained from GWAS and they been linked to core metabolic processes which can help understand the underlying principles of cancer progression and design therapeutic

Non-coding mutations in cancer

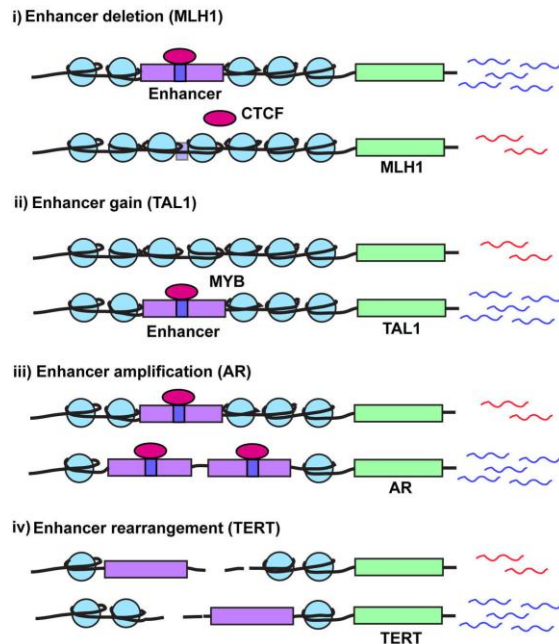


Figure 1. Structural mutations in enhancers: (i) IGFBP5 gene is regulated by its enhancer. When a small stretch in the enhancer is deleted which harboured the CTCF motif, the enhancer activity from the region is lost and gene is down-regulated. IGFBP5 being a tumour-suppressor gene leads to increased risk of cancer. (ii) An oncogene TAL1, has some basal expression in the wild type condition. Mutations lead to the gain of Myb motifs which facilitate the birth of a super-enhancer which up-regulates TAL1 expression and thus leading to increased susceptibility to cancer. (iii) Mutations in the enhancer of AR gene lead to the amplification of the region harbouring the enhancer which leads to increased expression of androgen receptor leading to increased prostate cancer risk. (iv) An enhancer is present physically far from the TERT gene. Rearrangement places the enhancer in close physical proximity to the TERT oncogene leading to its activation.

approaches (1,2). The latest advances in sequencing techniques have allowed the genome-wide profiling of regulatory regions and associated protein cargos, accelerating the identification of functional mutations that was otherwise synonymous to “finding a needle in a haystack”.

3. LINK BETWEEN MUTATIONS IN ENHANCERS AND CANCER

Employment of such tools on several cancers has revealed that enhancers carry the highest density of variants. For example, in prostate cancer, less than 20% of the variants were present in promoters or the coding region (3)

and 88% of the SNPs fell in putative enhancers (4). Enhancers are at the heart of regulated gene transcription; they mediate the transcriptional regulation of genes in a spacio-temporal manner to drive cell fate during development. Numerous enhancers are born, erased, activated, repressed during the cell fate choices to give rise to tightly controlled gene expression pattern in a committed lineage (5). It is believed that enhancers undergo similar transitions as oncogenes, during the cellular transformation from “normal” to “cancer-like”. Thus, perturbations in enhancer functions can have potentially harmful pleiotropic effects on the cell. Although coding variation is more deleterious as it can directly affect the protein in terms of its stability, composition, localization, interacting partners, function and activity, the variations within enhancers may affect the transcription rate of several genes in response to an environmental stimuli or developmental stage(s) resulting in the dysregulation in protein levels. Several reports now have strongly linked malignancies with mutation in enhancers (6).

Variations within the disease-causing enhancers could be of different types namely, an insertion or a deletion (indel), rearrangements or single nucleotide mutations. All these types of variations can lead to the loss of existing enhancer or de novo formation of an enhancer or can lead to a gain or loss of transcription factor binding sites resulting in alteration in enhancer activity (Figure 1). For example, it has been reported that individuals with single nucleotide deletions in the c-myc gene desert gains MYB binding site leading to the formation of a super enhancer to activate the TAL1 expression in T-cell acute lymphoblastic leukemia (T-ALL) (7). This study suggests, how an organism can gain a new enhancer during evolution to target the expression of important genes. Not only the deletion but also enhancer duplications aberrantly activate or represses target genes associated with monogenic diseases. For instance, duplication of an enhancer element upstream of the myc oncogene leads to the gain of a super enhancer (8). In gliomas, small somatic rearrangement hijack myb enhancer thus leading to the over expression of myb (9).

Non-coding mutations in cancer

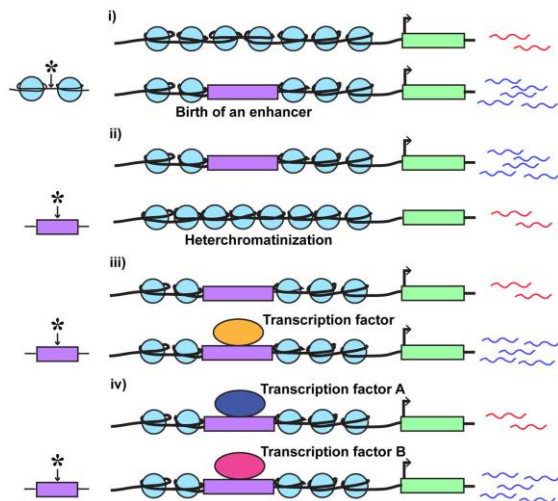


Figure 2. Single nucleotide polymorphism in enhancers: (i) A Single nucleotide polymorphism (SNP) leads to the birth of an enhancer. In the non-risk allele, the target gene is not regulated by any enhancer, but upon the gain of the risk allele of the SNP, an enhancer is born which targets the gene leading to its higher expression than normal levels. When such a gene is an oncogene, its upregulation might prove to be tumorigenic. (ii) An enhancer which targets its cognate gene possesses a SNP, the risk allele of which leads to the death of the enhancer and heterochromatinization of the enhancer region. This leads to the down-regulation of the gene and when this gene is a tumour-suppressor gene, it increases the susceptibility to cancer. (iii) An enhancer which targets its cognate gene, if possesses a risk SNP which leads to the gain of binding of a transcription activator leading to an up-regulation of the oncogene hence, susceptibility to cancer. (iv) An enhancer which targets its cognate gene, if possesses a risk SNP which leads to the gain of a different transcription factor than the one which was binding when the wild-type allele of the SNP was present. The new transcription factor is more potent in its activation of the enhancer and the target oncogene gets upregulated leading to cancer susceptibility.

The third kind of variation, single nucleotide polymorphisms (SNPs) within the enhancers is the most interesting but least understood. The risk SNP often affects the transcription factor binding in the core region of an enhancer which is where the eRNA transcription machinery also assembles. Thus lack or gain of TF binding on an enhancer, can not only affect the polII loading but also the rate at which eRNA is transcribed. Together, they affect the magnitude of target gene activation associated with a disease pathogenesis (Figure 2). For example, the gene desert region beside *myc* is a hotspot for genetic variation associated with several cancers in different human races and many GWAS have identified SNPs in this region which ultimately increases the expression of *myc* or numerous lncRNAs such as

PCAT1, *PRNCR1*, *CCAT1* and *PVT1* within the same TAD as *myc*. The lncRNA *PCAT1* in 8q24 is upregulated by rs7463708 which gains the binding for *ONECUT2* transcription factor leading to activating metastatic genes in trans (10). Similarly, risk SNP rs11672691 allows for the switching from *PCAT19* lncRNA's short to long isoform which increases the prostate cancer susceptibility (11). Similarly, there are plenty of examples now to conclude that common genetic variation within the enhancer constitutes a significant portion of functional genomics that is linked to malignancies (Table 1).

4. MUTATIONS IN INSULATORS/BOUNDARIES

The genome is compartmentalized into megabase-sized bins known as topologically associating domains (TADs). Generally, TADs exhibit two functional features namely, high intra-TAD interactions and low inter-TAD interactions. The low inter-TAD interactions are the result of the presence of a strong physical barrier between the TADs known as boundaries or insulators. During evolution, it has been seen that break points around syntenic regions are enriched at boundaries suggesting that the break points within TADs are strongly selected against, indicating the self-regulated or modular nature of TADs. Experimental deletions or mutations at boundaries merging two neighboring TADs results in the dysregulation of genes in the affected TADs. Likewise, the mutations within and around insulator regions are enriched in tumors resulting from alterations in TADs and ultimately gene dysregulation (12). Most TAD boundaries have two-fold higher binding of the transcription factor CTCF and CTCF knockdown severely affects the functions of boundaries. Thus, it is not surprising that a loss or gain of CTCF binding site(s) at TADs affects the genes in the TADs as evident in some cancers (Figure 3) (13).

5. MUTATIONS IN CHROMATIN ARCHITECTURE PROTEINS

A precise chromatin structure is the key to regulated transcription. Just like the mutations in boundary elements cause activation or

Non-coding mutations in cancer

Table 1. Examples of reported mutations in regulatory regions, their target genes and associated cancer

Type of Mutation	Gene	Cancer type	Reference
Gain of super enhancer	TAL1	T-cell Acute Lymphoblastic Leukemia	18
rs339331	RFX6	Prostate	19
rs8072254/rs1859961	SOX9	Prostate	20
rs67491583	MYC	Colorectal	21
Enhancer Translocation	MYB-QK1	Glioma	9
rs7463708	PCAT1 lncRNA	Prostate	10
Insulator deletions	TAL1	Lymphoblastic Leukemia	12
rs965513	FOXE1 and PTCSC2	Thyroid	22
rs554219, rs78540526, rs75915166	CCND1	Breast	23
rs6983267	MYC	Colorectal	24
Enhancer deletion	MLH1	Colorectal	7
Enhancer amplification	AR	Castrate-resistant prostate cancer	25
Enhancer invasion	MYCN-target genes	Neuroblastoma	26
Enhancer amplification, translocation	MYC	Pediatric neuroblastomas	27
Super-enhancer translocation	MYB	Cystic carcinoma	28
rs1167291	PCAT19 lncRNA	Prostate	11
rs2981578, rs35054928, rs45631563	FGFR2	Breast	13
rs920778	HOTAIR	Esophageal squamous cell carcinoma	29
rs12203592	IRF4	Acute Lymphocytic Leukemia	30
Super-enhancer rearrangements	TERT	Pheochromocytomas	31

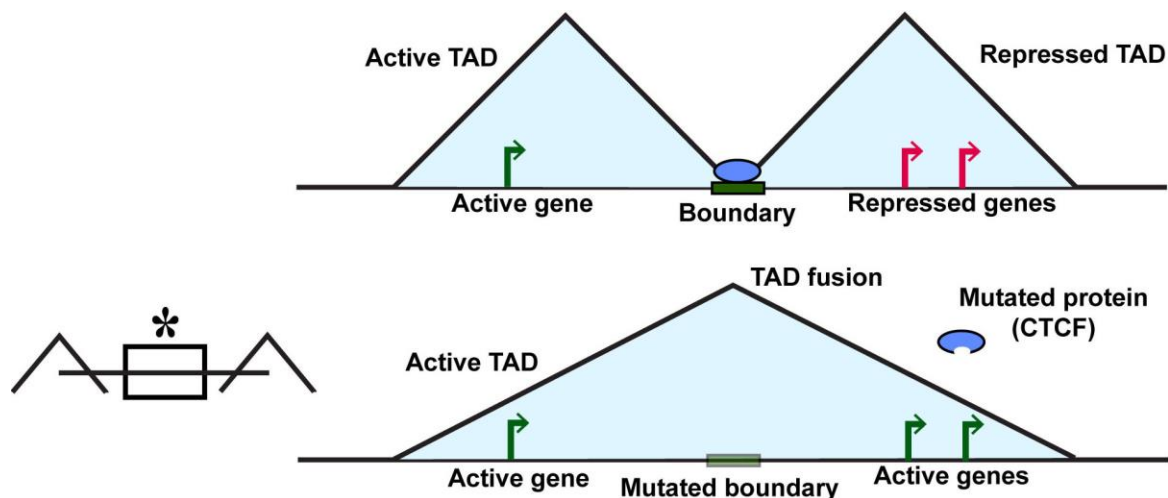


Figure 3. Mutations and insulator: An insulator separates two TADs; one active and the other repressed, harbouring active and repressed genes respectively. When there is a mutation in the insulator and/or a transcription factor (like CTCF) which binds to the insulator such that the insulator function is lost, the adjacent TADs merge and all the genes show a similar pattern of expression, active expression in this case.

repression of key tumor suppressors or activators, mutations in architectural proteins contribute to

genome-wide modifications in chromatin arrangement affecting regulation of several genes.

Non-coding mutations in cancer

Table 2. Cancer associated mutations in transcription regulators and architecture proteins.

Gene	Cancer	Reference
IRF4	Chronic lymphocytic leukemia	32
TET2	Myeloid Cancer	33
EZH2	Myelodysplastic syndromes	34
DNMT3A	Acute myeloid leukemia	35
SMARCE1	Familial multiple spinal meningiomas	36
ARID1B	Childhood neuroblastoma	37
BAF180	Breast Cancer	38
PBRM1	Renal cell carcinoma	39
ATRX	Glioblastoma	40
DOT1	Leukemia	41
MLL2 and MLL3	Medulloblastoma	42
NUP98-NSD1	Acute myeloid leukemia	43
RUNX1	Acute myeloid leukemia	44
JMJD3	Brainstem glioma	45
BRD3 and BRD4	Prostate Cancer	46
CREBBP	Oesophageal Cancer	47
EP300	Acute lymphoblastic leukemia	48
HDAC4	Acute lymphoblastic leukemia	49
SETD2	Leukemia	50
MLL1	Acute lymphoblastic leukemia	51
CTCF	Acute myeloid leukemia	52
STAG2	Several cancers	53
RAD21	Myeloid neoplasms	54
SPDEF	Gastric cancer	55
MYC	Several cancers	56

For example, mutations in the STAG2 subunit of the cohesin complex, which along with CTCF establishes the insulated neighborhood of TADs, has been reported in AML, Ewing-Sarcoma, bladder, melanoma, cervical and glioblastoma malignancies (14). Cohesin also plays a role in enhancer:promoter looping and thus mutations in these proteins could lead to pleiotropic effects. Since STAG2 is located on ChrX, a single mutation could lead to the loss of protein. Similarly, CDK8, CYCLIN C, MED12, MED13 and MED23 subunits of mediator complex are also frequently mutated in several cancers (15, 16).

6. MUTATIONS IN TRANSCRIPTIONAL MACHINERIES

Cancer cells employ wide arrays of mechanisms to alter the highly coordinated gene expression in normal cells which leads to cancerogenesis. Thus, not surprisingly, transcription factors, coactivators, corepressors and chromatin modifiers are mutated the most in cancers (12). Since chromatin regulators are highly cell-type specific, mutations in them give rise to specific cancers for example, mutations in nuclear receptors are specific

Non-coding mutations in cancer

to solid tumors. On the other hand, mutations in non-cell type specific or global regulators is a common theme in most tumors, such as mutations in zinc fingers of several transcription factors, particularly of Kruppel associated box (KRAB) domain containing factors are widespread across several cancers. Some reported mutations in transcriptional regulators in various cancer types are mentioned in Table 2.

7. eRNAs IN CANCER

The most fundamental aspects of functional enhancers and their associated eRNAs is that they are highly cell- and tissue-type specific. Organismal development has also been largely governed by these spatially and temporally regulated enhancers thus clearly bringing corresponding eRNAs to the fore of these aspects. Though mutations in enhancers affect corresponding eRNA expression, direct link between cancer and mutations in eRNAs has not been established yet. eRNAs themselves can serve many roles or just the act of their transcription can also be important, for instance, divergent transcription from super enhancer in the gene body of an oncogene can create stalling of PolII whereby these stalled polymerases create ssDNA that are recognized by Activation-Induced Cytidine Deaminase (AID) which in turn can cause the DSB leading to the translocations (17). There is a further need of locus based as well as genome-wide studies to fully appreciate the breadth and depth of their functions.

8. CONCLUSIONS

At present, hundreds of thousands of regulatory elements have been identified by ENCODE, Roadmap, FANTOM5 and a very recent project, Blueprint/IHEC. However, large scale measures to unravel the functionality of these genomic elements in different tissues and cancers are still lacking. Once these enhancers and their functions are identified, finding the causative mutations can be accelerated which will then pave the ways for correctional enhanceropathies.

9. ACKNOWLEDGMENTS

The authors apologize to all authors whose work could not be cited due to space constraints. The

authors are grateful to all the lab members for the fruitful discussions. KW acknowledges the SPM fellowship from Council of Scientific and Industrial Research (CSIR), India. DN acknowledges the financial support from the Wellcome Trust/India Alliance as well as the intramural funds from NCBS-TIFR. DN acknowledges support from EMBO global investigator program. Authors declare no competing interests.

10. REFERENCES

1. Zhang, W., Bojorquez-Gomez, A., Velez, D.O., Xu, G., Sanchez, K.S., Shen, J.P., Chen, K., Licon, K., Melton, C., Olson, K.M. and Yu, M.K. A global transcriptional network connecting noncoding mutations to changes in tumor gene expression. *Nat genet*, 50(4), 613. (2018)
DOI: 10.1038/s41588-018-0091-2
2. Jolly, M.K., Ware, K.E., Gilja, S., Somarelli, J.A. and Levine, H. EMT and MET: necessary or permissive for metastasis?. *Mol Oncol*. 11(7):755-769. (2017)
DOI: 10.1002/1878-0261.12083
3. Corradin, O. and Scacheri, P.C. Enhancer variants: evaluating functions in common disease. *Genome med*, 6(10), p.85 (1-14). (2014)
DOI: 10.1186/s13073-014-0085-3
4. Hazelett, D.J., Rhie, S.K., Gaddis, M., Yan, C., Lakeland, D.L., Coetzee, S.G., Henderson, B.E., Noushmehr, H., Cozen, W., Kote-Jarai, Z. and Eeles, R.A. Comprehensive functional annotation of 77 prostate cancer risk loci. *PLoS genet*, 10(1), p.e1004102. (2014)
DOI: 10.1371/journal.pgen.1004102
5. Bonifer, C., & Cockerill, P. N. Chromatin priming of genes in development: concepts, mechanisms and consequences. *Exp hematol*, 49, 1-8. (2017)

Non-coding mutations in cancer

- DOI: 10.1016/j.exphem.2017.01.003
6. Herz, H. M., Hu, D., & Shilatifard, A. Enhancer malfunction in cancer. *Mol cell*, 53(6), 859-866. (2014)
DOI: 10.1016/j.molcel.2014.02.033
 7. Liu, Q., Thoms, J.A., Nunez, A.C., Huang, Y., Knezevic, K., Packham, D., Poulos, R.C., Williams, R., Beck, D., Hawkins, N.J. and Ward, R.L. Disruption of a- 35 kb Enhancer Impairs CTCF Binding and MLH1 Expression in Colorectal Cells. *Clinical Cancer Res*, 24(18), pp.4602-4611. (2018)
DOI: 10.1158/1078-0432.CCR-17-3678
 8. Gabay, M., Li, Y. and Felsher, D.W. MYC activation is a hallmark of cancer initiation and maintenance. *CSH perspectives in medicine*, 4(6), p.a014241. (2014)
DOI: 10.1101/cshperspect.a014241
 9. Bandopadhyay, P., Ramkissoon, L.A., Jain, P., Bergthold, G., Wala, J., Zeid, R., Schumacher, S.E., Urbanski, L., O'Rourke, R., Gibson, W.J. and Pelton, K. MYB-QKI rearrangements in angiocentric glioma drive tumorigenicity through a tripartite mechanism. *Nat genet*, 48(3), pp.273-282. (2016)
DOI: 10.1158/1538-7445.AM2016-4372
 10. Guo, H., Ahmed, M., Zhang, F., Yao, C.Q., Li, S., Liang, Y., Hua, J., Soares, F., Sun, Y., Langstein, J. and Li, Y. Modulation of long noncoding RNAs by risk SNPs underlying genetic predispositions to prostate cancer. *Nat genet*, 48(10), pp.1142-1150. (2016)
DOI: 10.1038/ng.3637
 11. Hua, J.T., Ahmed, M., Guo, H., Zhang, Y., Chen, S., Soares, F., Lu, J., Zhou, S., Wang, M., Li, H. and Larson, N.B. Risk SNP-mediated promoter-enhancer switching drives prostate cancer through lncRNA PCAT19. *Cell*, 174(3), pp.564-575. (2018)
DOI: 10.1016/j.cell.2018.06.014
 12. Hnisz, D., Weintraub, A.S., Day, D.S., Valton, A.L., Bak, R.O., Li, C.H., Goldmann, J., Lajoie, B.R., Fan, Z.P., Sigova, A.A. and Reddy, J. Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science*, 351(6280), pp.1454-1458. (2016)
DOI: 10.1126/science.aad9024
 13. Campbell, T.M., Castro, M.A., de Santiago, I., Fletcher, M.N., Halim, S., Prathalingam, R., Ponder, B.A. and Meyer, K.B. FGFR2 risk SNPs confer breast cancer risk by augmenting oestrogen responsiveness. *Carcinogenesis*, 37(8), pp.741-750. (2016)
DOI: 10.1093/carcin/bgw065
 14. Hill, V.K., Kim, J.S. and Waldman, T. Cohesin mutations in human cancer. *Biochimica Et Biophysica Acta (BBA)-Reviews on Cancer*, 1866(1), pp.1- 11. (2016)
DOI: 10.1016/j.bbcan.2016.05.002
 15. Rovnak, J., Brewster, C.D. and Quackenbush, S.L. Retroviral cyclin enhances cyclin-dependent kinase-8 activity. *Journal of virol*, 86(10), pp.5742-5751. (2012)
DOI: 10.1128/JVI.07006-11
 16. Guo X.1. and Wang X.F. A mediator lost in the war on cancer. *Cell*. 21;151(5), pp. 927-9299. (2012)
DOI: 10.1016/j.cell.2012.11.003
 17. Meng, F.L., Du, Z., Federation, A., Hu, J., Wang, Q., Kieffer-Kwon, K.R., Meyers, R.M., Amor, C., Wasserman, C.R., Neuberg, D. and Casellas, R. Convergent transcription at intragenic super-enhancers

Non-coding mutations in cancer

- targets AID-initiated genomic instability. *Cell*, 159(7), pp.1538-1548. (2014)
DOI: 10.1016/j.cell.2014.11.014
18. Mansour, M.R., Abraham, B.J., Anders, L., Berezovskaya, A., Gutierrez, A., Durbin, A.D., Etchin, J., Lawton, L., Sallan, S.E., Silverman, L.B. and Loh, M.L. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science*, 346(6215), pp.1373-1377. (2014)
DOI: 10.1126/science.1259037
19. Huang, Q., Whittington, T., Gao, P., Lindberg, J.F., Yang, Y., Sun, J., Väisänen, M.R., Szulkin, R., Annala, M., Yan, J. and Egevad, L.A. A prostate cancer susceptibility allele at 6q22 increases RFX6 expression by modulating HOXB13 chromatin binding. *Nat genet*, 46(2), pp.126-135. (2014)
DOI: 10.1038/ng.2862
20. Zhang, J., Ding, L., Holmfeldt, L., Wu, G., Heatley, S.L., Payne-Turner, D., Easton, J., Chen, X., Wang, J., Rusch, M. and Lu, C. The genetic basis of early T- cell precursor acute lymphoblastic leukaemia. *Nature*, 481(7380), pp.157-163. (2012)
21. Tuupanen, S., Yan, J., Turunen, M., Gylfe, A.E., Kaasinen, E., Li, L., Eng, C., Culver, D.A., Kalady, M.F., Pennison, M.J. and Pasche, B. Characterization of the colorectal cancer-associated enhancer MYC-335 at 8q24: the role of rs67491583. *Cancer genet*, 205(1-2), pp.25-33. (2012)
DOI: 10.1016/j.cancergen.2012.01.005
22. He, H., Li, W., Liyanarachchi, S., Srinivas, M., Wang, Y., Akagi, K., Wang, Y., Wu, D., Wang, Q., Jin, V. and Symer, D.E. Multiple functional variants in long-range enhancer elements contribute to the risk of SNP rs965513 in thyroid cancer. *PNAS*, 112(19), pp.6128-6133. (2015)
DOI: 10.1073/pnas.1506255112
23. French, J.D., Ghoussaini, M., Edwards, S.L., Meyer, K.B., Michailidou, K., Ahmed, S., Khan, S., Maranian, M.J., O'Reilly, M., Hillman, K.M. and Betts, J.A. Functional variants at the 11q13 risk locus for breast cancer regulate cyclin D1 expression through long-range enhancers. *AJHG*, 92(4), pp.489-503. (2013)
24. Pomerantz, M.M., Ahmadiyeh, N., Jia, L., Herman, P., Verzi, M.P., Doddapaneni, H., Beckwith, C.A., Chan, J.A., Hills, A., Davis, M. and Yao, K. The 8q24 cancer risk variant rs6983267 shows long-range interaction with MYC in colorectal cancer. *Nat genet*, 41(8), pp.882-884. (2009)
DOI: 10.1038/ng.403
25. Takeda, D.Y., Spisák, S., Seo, J.H., Bell, C., O'Connor, E., Korthauer, K., Ribli, D., Csabai, I., Solymosi, N., Szállási, Z. and Stillman, D.R. A somatically acquired enhancer of the androgen receptor is a noncoding driver in advanced prostate cancer. *Cell*, 174(2), pp.422-432. (2018)
DOI: 10.1016/j.cell.2018.05.037
26. Zeid, R., Lawlor, M.A., Poon, E., Reyes, J.M., Fulciniti, M., Lopez, M.A., Scott, T.G., Nabet, B., Erb, M.A., Winter, G.E. and Jacobson, Z. Enhancer invasion shapes MYCN-dependent transcriptional amplification in neuroblastoma. *Nat genet*, 50(4), pp.515-523. (2018)
DOI: 10.1038/s41588-018-0044-9
27. Zimmerman, M.W., Liu, Y., He, S., Durbin, A.D., Abraham, B.J., Easton, J., Shao, Y., Xu, B., Zhu, S., Zhang, X. and Li, Z. MYC drives a subset of high-risk pediatric neuroblastomas and is activated through mechanisms including enhancer hijacking and focal enhancer amplification. *Cancer*

Non-coding mutations in cancer

- disc, 8(3), pp.320-335. (2018)
DOI: 10.1158/2159-8290.CD-17-0993
28. Drier, Y., Cotton, M.J., Williamson, K.E., Gillespie, S.M., Ryan, R.J., Kluk, M.J., Carey, C.D., Rodig, S.J., Sholl, L.M., Afrogheh, A.H. and Faquin, W.C. An oncogenic MYB feedback loop drives alternate cell fates in adenoid cystic carcinoma. *Nat genet*, 48(3), p.265. (2016)
DOI: 10.1038/ng.3502
29. Zhang, X., Zhou, L., Fu, G., Sun, F., Shi, J., Wei, J., Lu, C., Zhou, C., Yuan, Q. and Yang, M. The identification of an ESCC susceptibility SNP rs920778 that regulates the expression of lncRNA HOTAIR via a novel intronic enhancer. *Carcinogenesis*, 35(9), pp.2062-2067. (2014)
DOI: 10.1093/carcin/bgu103
30. Visser, M., Palstra, R.J. and Kayser, M. Allele-specific transcriptional regulation of IRF4 in melanocytes is mediated by chromatin looping of the intronic rs12203592 enhancer to the IRF4 promoter. *Human mol genet*, 24(9), pp.2649-2661. (2015)
DOI: 10.1093/hmg/ddv029
31. Dwight, T., Flynn, A., Amarasinghe, K., Benn, D.E., Lupat, R., Li, J., Cameron, D.L., Hogg, A., Balachander, S., Candiloro, I.L. and Wong, S.Q. TERT structural rearrangements in metastatic pheochromocytomas. *Endocrine-related cancer*, 25(1), pp.1-9. (2018)
DOI: 10.1530/ERC-17-0306
32. Havelange, V., Pekarsky, Y., Nakamura, T., Palamarchuk, A., Alder, H., Rassenti, L., Kipps, T. and Croce, C.M. IRF4 mutations in chronic lymphocytic leukemia. *Blood*, 118(10), pp.2827-2829. (2011)
DOI: 10.1182/blood-2011-04-350579
33. Delhommeau, F., Dupont, S., Valle, V.D., James, C., Trannoy, S., Masse, A., Kosmider, O., Le Couedic, J.P., Robert, F., Alberdi, A. and Lecluse, Y. Mutation in TET2 in myeloid cancers. *NEJM*, 360(22), pp.2289-2301. (2009)
DOI: 10.1056/NEJMoa0810069
34. Nikoloski, G., Langemeijer, S.M., Kuiper, R.P., Knops, R., Massop, M., Tönnissen, E.R., Van der Heijden, A., Scheele, T.N., Vandenberghe, P., de Witte, T. and van der Reijden, B.A. Somatic mutations of the histone methyltransferase gene EZH2 in myelodysplastic syndromes. *Nat genet*, 42(8), pp.665-667. (2010)
DOI: 10.1038/ng.620
35. Ley, T.J., Ding, L., Walter, M.J., McLellan, M.D., Lamprecht, T., Larson, D.E., Kandoth, C., Payton, J.E., Baty, J., Welch, J. and Harris, C.C. DNMT3A mutations in acute myeloid leukemia. *NEJM*, 363(25), pp.2424-2433. (2010)
36. Smith, M.J., O'Sullivan, J., Bhaskar, S.S., Hadfield, K.D., Poke, G., Caird, J., Sharif, S., Eccles, D., Fitzpatrick, D., Rawluk, D. and du Plessis, D. Loss- of-function mutations in SMARCE1 cause an inherited disorder of multiple spinal meningiomas. *Nat genet*, 45(3), pp.295-298. (2013)
DOI: 10.1038/ng.2552
37. Kadoch, C., Hargreaves, D.C., Hodges, C., Elias, L., Ho, L., Ranish, J. and Crabtree, G.R. Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat genet*, 45(6), pp.592-601. (2013)
DOI: 10.1038/ng.2628
38. Xia, W., Nagase, S., Montia, A.G., Kalachikov, S.M., Keniry, M., Su, T., Memeo, L., Hibshoosh, H. and Parsons, R.

Non-coding mutations in cancer

- BAF180 is a critical regulator of p21 induction and a tumor suppressor mutated in breast cancer. *Cancer res*, 68(6), pp.1667-1674. (2008)
DOI: 10.1158/0008-5472.CAN-07-5276
39. Varela, I., Tarpey, P., Raine, K., Huang, D., Ong, C.K., Stephens, P., Davies, H., Jones, D., Lin, M.L., Teague, J. and Bignell, G. Exome sequencing identifies frequent mutation of the SWI/SNF complex gene PBRM1 in renal carcinoma. *Nature*, 469(7331), pp.539-542. (2011)
40. Koschmann, C., Calinescu, A.A., Nunez, F.J., Mackay, A., Fazal-Salom, J., Thomas, D., Mendez, F., Kamran, N., Dzaman, M., Mulpuri, L. and Krasinkiewicz, J. ATRX loss promotes tumor growth and impairs nonhomologous end joining DNA repair in glioma. *Science transl med*, 8(328), pp. (1-11). (2016)
DOI: 10.1126/scitranslmed.aac8228
41. Bernt, K.M., Zhu, N., Sinha, A.U., Vempati, S., Faber, J., Krivtsov, A.V., Feng, Z., Punt, N., Daigle, A., Bullinger, L. and Pollock, R.M. MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L. *Cancer cell*, 20(1), pp.66-78. (2011)
DOI: 10.1016/j.ccr.2011.06.010
42. Parsons, D.W., Li, M., Zhang, X., Jones, S., Leary, R.J., Lin, J.C.H., Boca, S.M., Carter, H., Samayoa, J., Bettegowda, C. and Gallia, G.L. The genetic landscape of the childhood cancer medulloblastoma. *Science*, 331(6016), pp.435- 439. (2011)
DOI: 10.1126/science.1198056
43. Wang, G.G., Cai, L., Pasillas, M.P. and Kamps, M.P. NUP98-NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. *Nat cell biol*, 9(7), pp.804-812. (2007)
- DOI: 10.1038/ncb1608
44. Papaemmanuil, E., Doehner, H. and Campbell, P.J., Genomic Classification in Acute Myeloid Leukemia. *NEJM*, 375(9), pp.900-901. (2016)
DOI: 10.1056/NEJMc1608739
45. Hashizume, R., Andor, N., Ihara, Y., Lerner, R., Gan, H., Chen, X., Fang, D., Huang, X., Tom, M.W., Ngo, V. and Solomon, D. Pharmacologic inhibition of histone demethylation as a therapy for pediatric brainstem glioma. *Nat med*, 20(12), pp.1394-1396. (2014)
DOI: 10.1038/nm.3716
46. Asangani, I.A., Dommeti, V.L., Wang, X., Malik, R., Cieslik, M., Yang, R., Escara-Wilke, J., Wilder-Romans, K., Dhanireddy, S., Engelke, C. and Iyer, M.K. Therapeutic targeting of BET bromodomain proteins in castration-resistant prostate cancer. *Nature*, 510(7504), 278. (2014)
DOI: 10.1038/nature13229
47. Gao, Y.B., Chen, Z.L., Li, J.G., Hu, X.D., Shi, X.J., Sun, Z.M., Zhang, F., Zhao, Z.R., Li, Z.T., Liu, Z.Y. and Zhao, Y.D. Genetic landscape of esophageal squamous cell carcinoma. *Nat genet*, 46(10), pp.1097-1102. (2014)
DOI: 10.1038/ng.3076
48. Gocho, Y., Kiyokawa, N., Ichikawa, H., Nakabayashi, K., Osumi, T., Ishibashi, T., Ueno, H., Terada, K., Oboki, K., Sakamoto, H. and Shioda, Y. A novel recurrent EP300-ZNF384 gene fusion in B-cell precursor acute lymphoblastic leukemia. *Leukemia*, 29(12), pp.2445-2448. (2015)
DOI: 10.1038/leu.2015.111
49. Gruhn, B., Naumann, T., Gruner, D., Walther, M., Wittig, S., Becker, S., Beck, J.F. and Sonnemann, J. The expression of

Non-coding mutations in cancer

- histone deacetylase 4 is associated with prednisone poor-response in childhood acute lymphoblastic leukemia. *Leukemia res*, 37(10), pp.1200-1207. (2013)
DOI: 10.1016/j.leukres.2013.07.016
50. Wang, Q. and Cheng, T. Evidences for mutations in the histone modifying gene SETD2 as critical drivers in leukemia development. *Science China Life Sciences*, 57(9), pp.944-946. (2014)
DOI: 10.1007/s11427-014-4702-6
51. Figueroa, M.E., Chen, S.C., Andersson, A.K., Phillips, L.A., Li, Y., Sotzen, J., Kundu, M., Downing, J.R., Melnick, A. and Mullighan, C.G. Integrated genetic and epigenetic analysis of childhood acute lymphoblastic leukemia. *The Journal of clinical investment*, 123(7), pp.3099-3111. (2013)
DOI: 10.1172/JCI66203
52. Dolnik, A., Engelmann, J.C., Scharfenberger-Schmeer, M., Mauch, J., Kelkenberg-Schade, S., Haldemann, B., Fries, T., Krönke, J., Kühn, M.W., Paschka, P. and Kayser, S. Commonly altered genomic regions in acute myeloid leukemia are enriched for somatic mutations involved in chromatin remodeling and splicing. *Blood*, 120(18), pp.e83-e92. (2012)
DOI: 10.1182/blood-2011-12-401471
53. Solomon, D.A., Kim, T., Diaz-Martinez, L.A., Fair, J., Elkahoul, A.G., Harris, B.T., Toretsky, J.A., Rosenberg, S.A., Shukla, N., Ladanyi, M. and Samuels, Y. Mutational inactivation of STAG2 causes aneuploidy in human cancer. *Science*, 333(6045), pp.1039-1043. (2011)
DOI: 10.1126/science.1203619
54. Kon, A., Shih, L.Y., Minamino, M., Sanada, M., Shiraishi, Y., Nagata, Y., Yoshida, K., Okuno, Y., Bando, M., Nakato, R. and Ishikawa, S. Recurrent mutations in multiple components of the cohesin complex in myeloid neoplasms. *Nat genet*, 45(10), pp.1232-1237. (2013)
DOI: 10.1038/ng.2731
55. Wu, J., Qin, W., Wang, Y., Sadik, A., Liu, J., Wang, Y., Song, P., Wang, X., Sun, K., Zeng, J. and Wang, L., SPDEF is overexpressed in gastric cancer and triggers cell proliferation by forming a positive regulation loop with FoxM1. *Journal of cellular biochem*, 119(11), pp.9042-9054. (2018)
DOI: 10.1002/jcb.27161
56. Kalkat, M., De Melo, J., Hickman, K., Lourenco, C., Redel, C., Resetca, D., Tamachi, A., Tu, W. and Penn, L. MYC deregulation in primary human cancers. *Genes*, 8(6), p.151 (1-30). (2017)
DOI: 10.3390/genes8060151

Abbreviations: CNVs- Copy number variations, WGS- Whole-genome sequencing, DHS- DNase hypersensitivity, ChIP- Chromatin immunoprecipitation, GRO-Global run-on, ChIA-PET- Chromatin Interaction Analysis by Paired-End Tag, 3C- Chromatin conformation capture, GWAS- Genome-wide association study, SNP- Single nucleotide polymorphism, TF-Transcription factor, TAD- Topologically associating domain and DSB- Double strand break.

Key Words: Genome, Mutation, Enhancer, Cancer, SNP, Oncogene, GWAS, eRNA, Insulator, Boundary, TAD, CTCF, Chromatin architecture, Enhanceropathy and Review.

Send correspondence to: Dimple Notani, 1National Centre for Biological Sciences, Bangalore, India, Tel: 080 2366 6664, E-mail: dnotani@ncbs.res.in