The nomination for research award is for following research:

Over the years, using genome-wide and locus-specific approaches, her group has identified uniquely marked enhancers that are responsible for reproducible and robust gene expression upon repetitive signaling events such as hormone signaling. We have identified unconventional principles of enhancer actions that challenge the prevalent, '*one-size-fits-all approach*' of their molecular understanding. Further, we have revealed how rare genetic mutations can give rise to active enhancers that increase the genetic predisposition of a population to disease. Ongoing efforts in my lab seek to explore the effects of transcription/RNA/cell cycle and other active processes on enhancer functions and genome organization whose perturbations are directly linked to patho-physiologies.

1. Interdependent network of functional enhancers within dense enhancer cluster

Enhancers were originally identified from locus-specific studies. Spurred by the advancements in next-generation sequencing technologies, genome-wide application of ChIP-seq revealed ~1 million putative enhancers in the human genome across tissues and conditions, galvanizing a broad effort to functionally prioritize and characterize these putative enhancers. The enhancers present within short distances from each other and those that exhibited greater threshold of certain marks (e.g., H3K27ac, MED1, TF binding etc.) were referred to as super enhancers (SE)[2-3]. As a dense cluster of enhancers, SEs exhibit disproportionately high signals of enhancer marks, and robustly induce the expression of target gene at a substantially higher level than a single enhancer. These observations, combined with the fact that cell identity genes typically harbour SEs, have helped the status of SEs as "Functional enhancers" and has dominated the enhancer field.

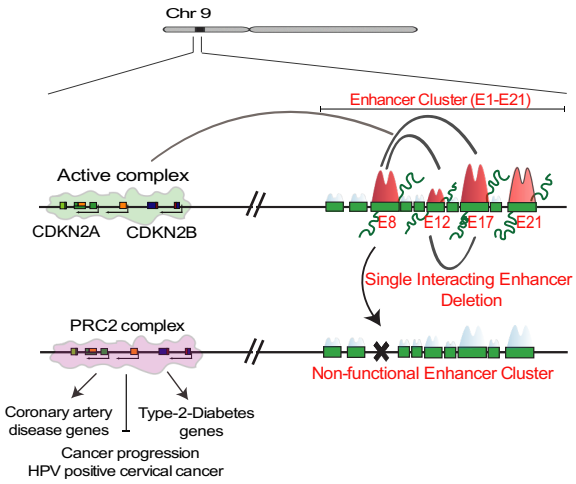

Figure 2: Only a few enhancers in a cluster regulate the target genes (Farooq et al., 2021).

Unfortunately, such 'one size fits all' approach to SE annotations can miss the identification of *bona fide* enhancers that are non-canonically marked or exceed the arbitrary distance cut-off prescribed to demarcate SEs. We first observed the proof for this caveat in ERα-bound enhancer clusters where the function of entire enhancer cluster was dependent on single book-marked enhancer[8]. The work emphasizes that assigning superlative roles to the clusters based on the number of constituent enhancers and their biochemical marking is not valid.

Further, individual enhancers in a cluster can have various regulatory arrangements (co-operative, hierarchical, mutual, redundant/non-redundant) with each other. To understand these enhancer relationships, dissection of one locus at a time approach is required. Towards this, we chose one of the densest enhancer cluster (24 enhancers) in the genome that is adjacent to *INK4/ARF* locus[13]. The locus codes for p14, p15 and p16 by virtue of alternative splicing of *CDKN2A* and *CDKN2B* loci. Importantly, genetic variation in the gene desert region beside *INK4a/ARF* locus has been associated with diseases such as cancers, coronary artery disease, type-2-diabetes, atherosclerosis and Alzheimer's[16-17]. Understanding of the enhancer cluster is crucial in linking the disease associated variants in this region with pathologies.

We used chromatin conformation capture (3C) techniques and a series of enhancer deletions within the cluster and observed that the promoter of *CDKN2A* physically interacted with only five enhancers (E5, E8, E12, E17 and E19) (Figure 2). Further, the perturbation or deletion of any of the interacting enhancers individually silenced the transcription of INK4a, ARF and INK4b promoters whereas deletion of promoter non-interacting enhancer, did not exhibit any effects on target genes. The loading

of EZH2 (Component of polycomb protein complex that catalyzes H3K27me3) and subsequent increase in H3K27me3 levels on promoters was observed upon such enhancer perturbations suggesting that at least one of the functions of the enhancer cluster is to protect the promoters from EZH2 loading. Our data suggests that there exists an inter-dependent enhancer network within the super enhancer cluster where all enhancers rely on each other for their transcriptional activity and the deletion of any single enhancer results in silencing of other intact enhancers as well as the promoters.

- Farooq U, Saravanan B, Islam Z, Walavalkar K, Singh AK, Jayani RS, Meel S, Swaminathan S, Notani D. An interdependent network of functional enhancers regulates transcription and EZH2 loading at the INK4a/ARF locus. Cell Rep. 2021 Mar 23;34(12):108898. doi: 10.1016/j.celrep.2021.108898.

- Farooq U, Notani D. Optimized protocol to create deletion in adherent cell lines using CRISPR/Cas9 system. STAR Protoc. 2021 Oct 22;2(4):100857. doi: 10.1016/j.xpro.2021.100857.

- Blobel GA, Higgs DR, Mitchell JA, Notani D, Young RA. Testing the super-enhancer concept. Nat Rev Genet. 2021 Dec;22(12):749-755. doi: 10.1038/s41576-021-00398-w. Epub 2021 Sep 3.

- Farooq U, Notani D. Transcriptional regulation of INK4/ARF locus by cis and trans mechanisms. Front Cell Dev Biol. 2022 Sep 9;10:948351.

By functionally dissecting enhancer cluster at 9p21 locus and, signaling driven transient enhancer cluster, we reveal the unexpected dependence of enhancer clusters on single to few functional enhancers emphasising that the effect of super enhancer on target genes is not the sum of constituent enhancers. **Our studies underscore the inadequacy of super enhancer based assigning of functionality as currently practiced and emphasises that, even though laborious, a careful enhancer-by-enhancer dissection of SE function is the only way forward towards a comprehensive understanding of SE functional diversity.**

2. <u>Functions of enhancers in genome organization</u>

The genome is partitioned into TADs[4]. The promoter search process is supposedly restricted to these local chromatin domains in which both the enhancers and its promoters are located. About half of the boundaries of these TADs exhibit transcriptional activity and are correlated with better TAD insulation[5-6]. However, the role of these transcripts *per se* in TAD insulation, enhancer:promoter interactions and transcription remains unknown.

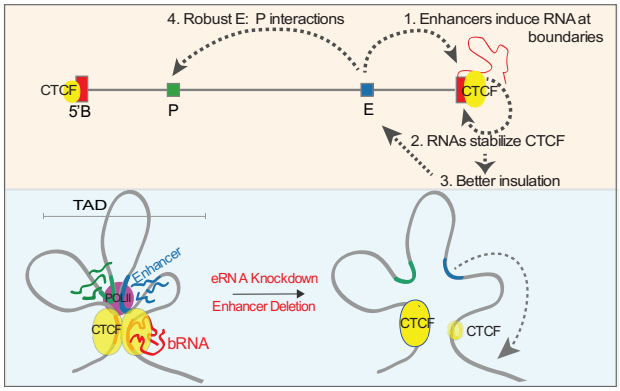

Figure 3: eRNAs within the boundaries recruit CTCF to enhance the TAD insulation (Islam and Saravanan et al, 2021).

We investigated the functional roles of these bRNAs (boundary-RNAs) in boundary insulation and subsequent effects on enhancer-promoter interactions and TAD transcription genome-wide and on disease relevant *INK4a/ARF* TAD. We have extensively studied the enhancer cluster mediated regulation of this locus[9]. In this study, using a series of deletions of CTCF sites at boundary, enhancer deletions and bRNA knockdown approaches at this TAD, we observed that indeed, enhancers interact with boundaries and positively regulate the RNA transcription at TAD boundaries (bRNA). In return, the bRNAs recruit/stabilize CTCF even on weaker motifs within these boundaries and support CTCF binding in clusters. Robust CTCF binding enhances TAD insulation during loop extrusion process

which then favours the intra-TAD enhancer:promoter interactions and robust gene transcription. Functionally, eRNAs within the boundaries are repurposed as more stable bRNAs and their knockdown exactly mimics the boundary loss.

Together, these results show that active enhancers directly mediate better insulation of TADs by activating the transcription at TAD boundaries. Furthermore, our results highlight the noncanonical role of enhancers in chromatin organization which is completely novel.

- Islam Z*, Saravanan B*, Walavalkar K, Farooq U, Singh AK, Thakur J, Pandit A, Sabarinathan R, Henikoff S, Notani D. Active enhancers strengthen TAD insulation by bRNA mediated CTCF enrichment at the TAD boundaries. *bioRxiv* 2021.07.13.452118; doi: https://doi.org/10.1101/2021.07.13.452118 (* Equal Contribution).

3. <u>Understanding Enhancer mechanisms through germline mutations</u>

Over 3000 genome-wide association studies (GWASs) published until 2017 have identified susceptibility loci to over 1800 unique traits and common diseases[18]. Most of these susceptibility loci are associated with non-coding genetic variation, especially in the enhancers. Such genetic variations can be exploited to discover mechanisms of enhancer action and in the long term can be used to develop therapeutic approaches that are driven by such enhanceropathies.

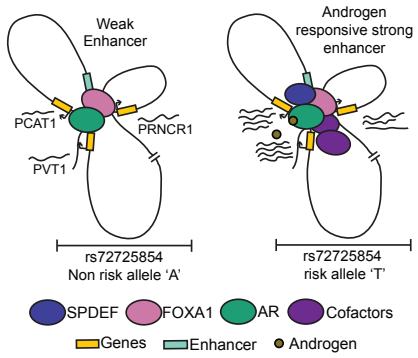

Figure 4: The rare SNPs *rs72725854* in an enhancer allows enhancers to respond to androgens (Walavalkar et al., 2020).

In order to understand the heightened genetic predisposition of African men to prostate cancer, my collaborator Dr. Christopher Haiman (University of Southern California, Los Angeles) had identified a rare SNP, *rs72725854,* that was prevalent in 12% of prostate cancer patients of the African- ancestry as compared to 6% in the control cohort[20]. We were interested in exploring how this rare SNP might be contributing to prostate cancer susceptibility in men of African ancestry as rare SNPs are not known for their causal functions. We got even more intrigued by the initial analyses which revealed that this SNP was actually present in the non-coding DNA on 8q24, a region identified as an enhancer in prostate cancer cells. The risk SNP (T) lead to the gain of the binding site for a transcriptional activator, SPDEF, within the enhancer. SPDEF expression is a mark of aggressive prostate tumours. We observed that, SPDEF binding at 'T' allele rendered the enhancer sensitive to androgens (key hormone in prostate cancers) thereby, the enhancer with risk allele was hyper-activated[14] (Figure 4).

Once activated, the enhancer physically interacted with multiple oncogenic genes (*PCAT1*, *PRNCR1* and distant gene *PVT1*) that do not code for any protein but are long non-coding RNAs themselves. This led to the upregulation of these oncogenic ncRNAs. Thus, in the individuals with the risk allele, the enhancer gets hyper-activated upon androgen stimulation which results in higher expression of these target oncogenic non-coding RNAs. These RNAs are known to increase the expression of downstream cell proliferation genes and thus lead to a tumorous phenotype[21].

The study is the first report where a rare mutation in an enhancer has been functionally shown to lead to the susceptibility to a disease. The workflow described in the article can be adopted to understand the susceptibilities of the Indian population to several metabolic diseases like type-2-diabetes.

- Walavalkar K, Saravanan B, Singh AK, Jayani RS, Nair A, Farooq U, Islam Z, Soota D, Mann R, Shivaprasad PV, Freedman ML, Sabarinathan R, Haiman CA, Notani D. A rare variant of

African ancestry activates 8q24 lncRNA hub by modulating cancer associated enhancer. Nat Commun. 2020 Jul 17;11(1):3598. doi: 10.1038/s41467-020-17325-y.

- Walavalkar K, Notani D. Beyond the coding genome: Non-coding mutations and cancer. Frontiers in Bioscience, Landmark. 2020. 25:1825-1835. doi: 10.2741/4879

Implication of our findings in development of therapeutic avenues:

My findings on crucial roles of enhancer-RNA (eRNA) in gene regulation is already en route to development of therapy to correct the faulty gene expression levels by manipulating the enhancer RNA. The eRNA being tissue-type specific serve as an effective strategy to manipulate the gene in target tissue specifically without the off-target effects on other genes or the same gene in other tissues.

Further, our recent observations that perturbation of single enhancer in INK4/ARF locus is enough to completely silence the genes in this locus, has opened the door to suppress this locus in aging and cervical cancer. The reason being, the cell cycle regulator p16, p15 that are encoded by this locus are severely upregulated during aging. The higher levels of these proteins hamper the cell division and regeneration potential of stem cells that triggers the early onset of aging. Furthermore, these proteins are highly upregulated in cervical cancer. The reports have suggested the perturbation of these proteins can achieve 90% of cellular reprogramming efficiency and cervical cancer cell death. However, the avenues to target these genes in target tissues has been a bottleneck.

Towards this my lab has identified a single enhancer and its eRNA that controls this locus. We are currently looking to partners with pharmaceuticals companies to develop RNA therapy targeting this eRNA. If successful, the therapy will bring tremendous hope in delaying aging and treating advance stages of cervical cancer.

Furthermore, my lab is the only lab in India that is functionally exploiting the genetic variants in non-coding/regulatory genome to understand the genetic predisposition of Indian population to a disease. Our success in underpinning the role of the rare SNP ‘rs72725854’ with prostate cancer has emphasized the need of whole genome sequencing in order to understand the genetic predisposition of the Indian population to diabetes and cardio-vascular diseases. We have forged collaborations with clinicians and scientists in India and abroad towards this goal.

Dimple Notani
Associate Professor
NCBS, Bangalore