Comments Response:

1. *I think the supplement does a good job explaining the tradeoffs you used to select your pairs*

**Response:** We have further clarified this in our main manuscript. Also, considering the tumor heterogeneity, it is even difficult to get a “perfect” match using the normal tissue beside the tumor ones.

1. *It was not clear where the community would go to find them, and whether they would understand the relationship between them and other data/analyses at the portal, including the ENCODE encyclopedia.*

**Response:** We will certainly make this clearer. We're planning to add a section into the paper explaining how we merge our annotations with those in the encyclopedia. We are going back and forth with Zhiping on this, and we'll write back in a little bit with more detail on how this would be done.

1. *I would say cell types; there were plenty of primary cells in ENCODE 2, including NHEK, HMEC.*

**Response:** Yes, a good point. We'll try to use cell type more and try make clear that while we're focusing on top tier cell lines with much data there's many other cells in ENCODE.

1. *At least 7 of the factors in Fig 3 don’t bind DNA in a sequence specific manner; are the results from ChIP in the relevant cell type(s), or predicted from motif analysis?*

**Response:** You're right. We will make this clearer in the figure. Everything is from ChIP-seq. We are not just focusing on TFSS, but also reporting the rewiring status of general TFs.

1. *Good news if you have a model for AML; are they linked to CML?*

**Response:** You're right. We did realize that the pairing is problematical. K562 is from a CML patient. In the mutational analysis part, we have to use CLL data because there is not much whole genome sequencing data on CML. On the patient survival analysis, again there is no good CML data, we have to refer to AML patient information. There is report that IKZF1 deletion is a hallmark of ALL types of BCR-ABL fusion, but not CML. Deletion of IKZF1 was also identified as an acquired lesion at the time of transformation of CML to ALL

We will probably be deemphasized this more in next version

1. *Again, are you finding RBPs, or sequence motifs for RBP families?*

**Response:** We are finding RBPs, NOT the motifs. The binding data is from eCLIP experiments.

1. *Fig 4A RBPs includes TAF15 and GTF2F1; these are part of the Pol II initiation complex, they associate with DNA prior to initiation of transcription, and facilitate initiation.*

**Response:** These are overlaps between RBPs and TFs. We are test every factor with either eCLIP or ChIP-seq data. If there is an association in RBP, we report it as RBP.

1. *Doesn’t the ENCODE encyclopedia have these enhancers, promoters? Are your calculations the same or different? Shouldn’t this be explained either way?*

**Response:** We are carefully merging our annotation

1. *Consistent with oncogenesis model? Reproducible? Consistent with neighboring SNV?*

**Response:** For the SYCP2 gene, there is no clear conclusion of its role. It said “aberrant expression of this protein may contribute to genetic instability during HPV-associated cancer development”. So we did not discussed in detail.

1. *Every locus lies in a TAD; do mean the candidate enhancer and predicted target promoter are in the same TAD?*

**Response:** We just showed that this is quite an active region with lots of interactions within this TAD. We have the enhancer-target prediction results from our tool and from ChIA-pet data.

1. *FWIW, there are both sequence specific TFs and cofactors bound at this location, in Fig 5D, nicely marked by DHS and flanking H3K27ac; one known function for GATA3 is a key factor in mammary epithelia*

 **Response:** Yes, the region looks nice.

1. *It was unclear what the relationship is, if any, between this work and Shirley Liu’s.*

**Response:** Shirley is a co-author of this paper and contributed this analysis

1. *in some cases, I thought multiple comparisons were interesting*

**Response:** This is a good point. We are revising to incorporate the notion of the normal from merge of a number of different related cell types. We are revising the paper to try to incorporate this idea merge normal

1. *PLEASE don’t refer to everything in ENCODE as cell lines*

**Response:** We won't. We will change to cell type in some places.

1. *I’d be curious about how well the tumor cells cluster with ES cells, as opposed to the mature cell type they appear to be derived from.*

**Response:** Very good point. We're doing this calculation

1. *I am most concerned about the A549/IMR-90 pairing, I am next most concerned about the K562/GM12878*

**Response:** We are concerned about the lung one and for the blood cancer a comparison we're trying to use cd34 as appropriate. We have tried to fix this up a bit with a merged normal but we admit the pairing issue is problematic

1. *For HeLa, one could compare it to any primary epithelial cell type, or perhaps a few epithelial cell types. Of course it is thought that HPV integration in the HeLa genome was a critical event in transformation (PMID: 3028716, PMID: 23925245), so viral TF, protein coding genes, and regulatory DNA may play a role in addition to host genes.*

**Response:** Good point we're doing this.