Identification of driver mutations in distal regulatory elements in breast cancer

Matteo D'Antonio, Agnieszka D'Antonio-Chronowska, Christopher DeBoever, Frauke Drees, Angelo Arias, Hiroko Matsui, Kristen Jepsen, Florence Coulet, Andrea L. Richardson, Richard B. Schwab, Kelly A. Frazer

Summary

The presence of hundreds to thousands of background somatic mutations in a tumor makes it difficult to identify driver mutations that confer a growth advantage. Efforts to identify driver mutations have largely focused on the ~3% of the genome that encodes for genes whereas non-coding sequences remain relatively unexplored.

Here we identify distal regulatory elements harboring driver mutations in breast cancer by applying statistical and filtering approaches developed to identify driver mutations in genes. A pan-cancer analysis shows that some of these elements are drivers across multiple cancer types and have mutation rates similar to driver genes. Characterization using CRISPR genome editing and in silico methods show that the most mutated element has long-range control over chromatin structure and gene expression and the second most mutated element drives overexpression of TERT and other neighboring cancer genes.

We demonstrate for the first time that distal regulatory elements harboring driver mutations induce long-range epigenetic and gene expression changes, providing novel insights into the role non-coding elements play in cancer.

Introduction

Aim: To detect distal regulatory elements that harbor driver mutations in breast cancer

Breast cancer is a highly heterogeneous disease characterized by four major clinically relevant phenotypes (Gradishar 2014).

Efforts to identify driver mutations in breast cancer have largely focused on coding genes.

Recent pan-cancer analyses examining non-coding sequences have identified promoters of two genes (PLEKHS1 and WDR74) that harbor putative driver mutations in breast cancer (Weinhold 2014).

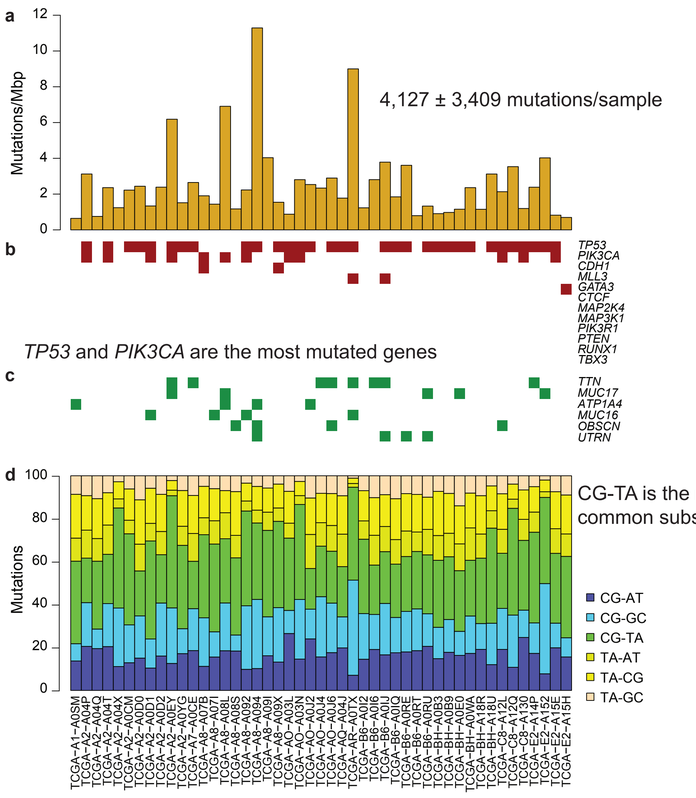
Regions of active chromatin have decreased somatic mutation rates compared with the rest of the non-coding genome (Polak, 2015), suggesting a likely driver role for mutations in these regions in cancer.

Limited data pinpoints specific distal regulatory elements that contain driver mutations in cancer and the functional role these elements play in cancer is unknown.

Mutational landscape of breast cancer

We downloaded whole genome sequencing data of 47 breast cancer samples and their matched normal DNA from TCGA (Wilks 2014).

We detected 193,958 somatic mutations using MuTect (Cibulskis 2013, ≥ 14X coverage in the tumor; ≥8X in the normal; ≥10% allelic fraction; not overlapping with repeat elements or known SNPs).

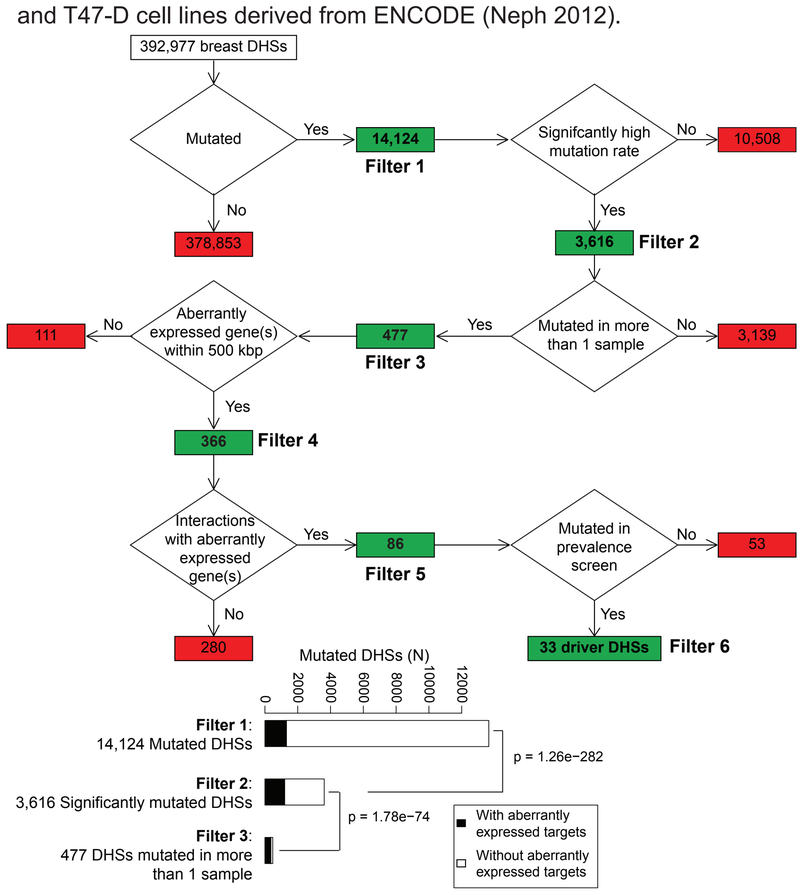


The mutational rates and patterns in the 47 discovery samples are generally consistent with breast cancer samples analyzed in previous genome-wide experiments (TCGA 2012, Ding 2010).

Identifying driver regulatory elements

Recent studies have shown that the local mutation density is lower in regulatory elements that are active in the cell-type of origin of the tumor compared to regulatory elements in other cell types (Polak 2015).

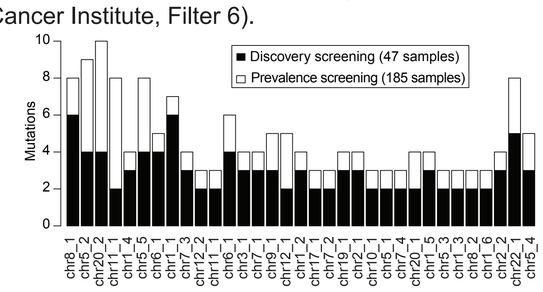
We focused on 392,977 DNase I hypersensitive sites (DHSs) in HMF and T47-D cell lines derived from ENCODE (Neph 2012).



The series of filters we applied to the breast DHSs enrich for those carrying driver mutations that effect the expression of neighboring genes.

Prevalence screening

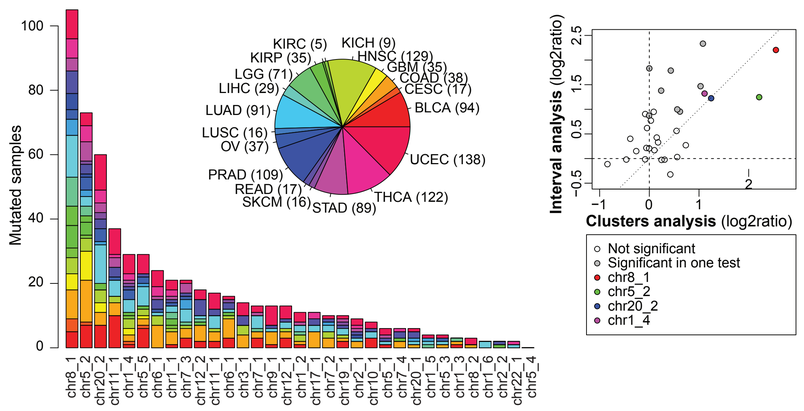
To further examine the 86 putative driver DHSs we determined whether or not they were mutated in 185 additional breast cancer samples (50 genomes from TCGA and 135 tumor samples from UCSD and Dana Farber Cancer Institute, Filter 6).



Prevalence screening allowed us to detect 33 driver DHSs in breast cancer.

Driver DHSs in multiple cancer types

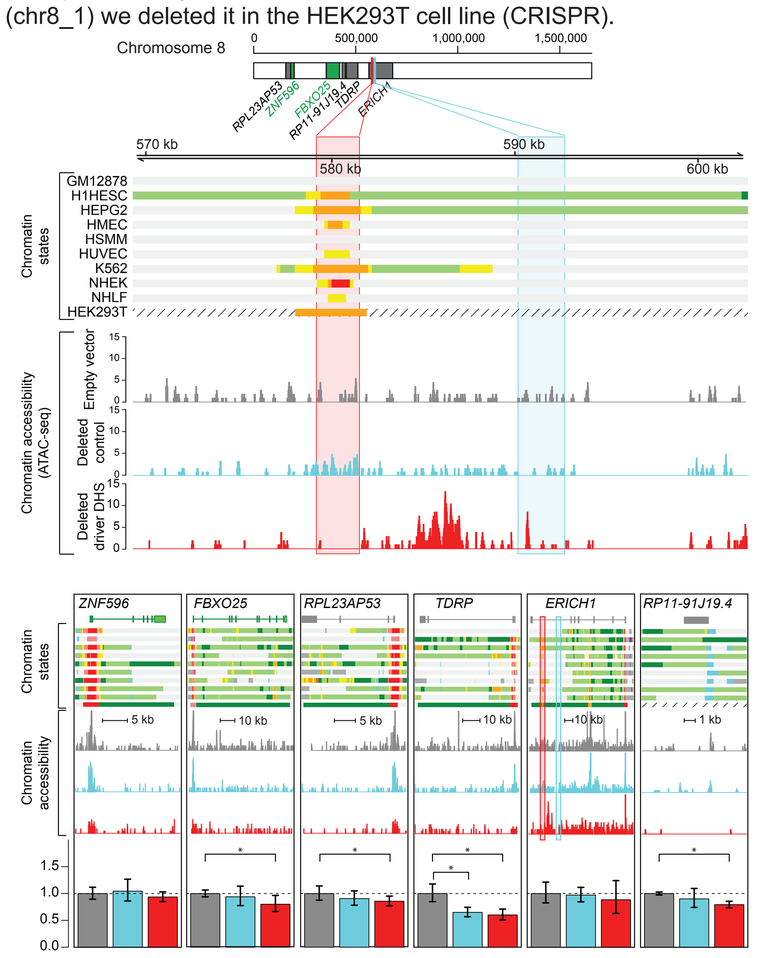
We examined the sequences of the 33 DHSs in 1,097 TCGA cancer genomes from 19 tumor types for mutations.



Four DHSs harbor driver mutations in multiple cancer types.

Long-range chromatin remodeling

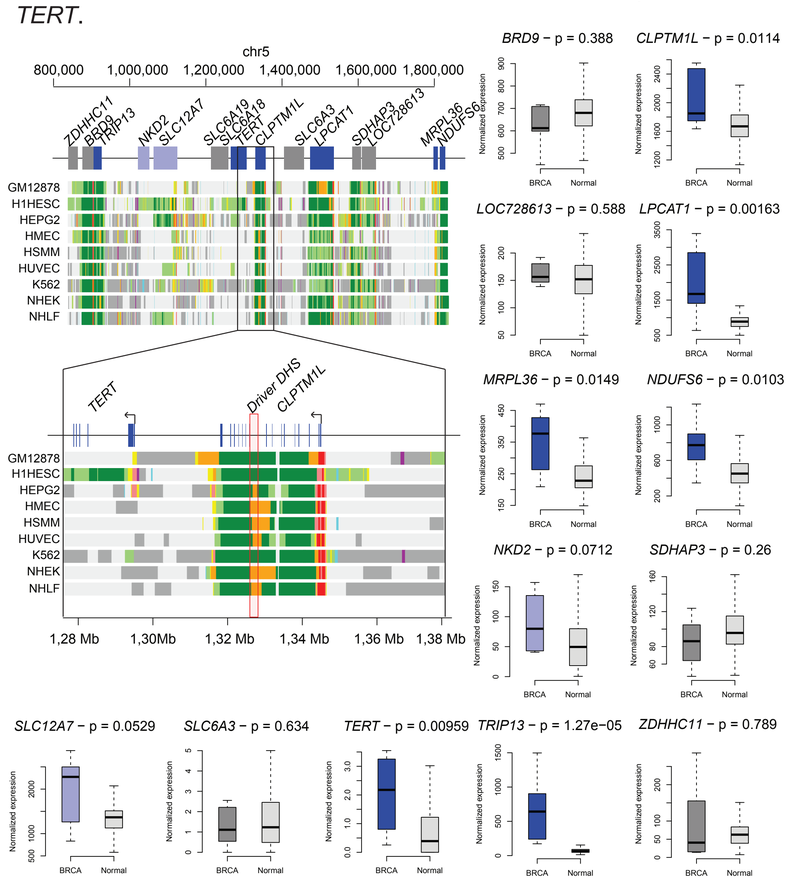
To experimentally examine the function of the most mutated driver DHS (chr8_1) we deleted it in the HEK293T cell line (CRISPR).



Deletion of the driver DHS resulted in increased chromatin accessibility of a region 3-kb downstream and decreased accessibility to the promoter regions of four nearby genes, coupled with decrease in expression levels of three of these genes.

Overexpression of cancer genes

The second-most mutated pan-cancer driver DHS (chr5_2) acts as an enhancer located within an intron of CLPTM1L and 30 kb upstream of *TERT*.



Mutations in this driver DHS may contribute to the development of cancer by driving overexpression of several different genes.

Conclusions

We adapted methods used to identify recurrently mutated cancer genes to distal regulatory elements.

We filtered on functional mutations by only retaining regulatory elements associated with aberrantly expressed gene(s).

We detected four DHSs with likely driver role in several cancer types.

We experimentally validated loss-of-function mutations in the most mutated regulatory element (chr8_1) by deleting it using CRISPR.

Mutations in the second-most mutated regulatory element (chr5_2) result in overexpression of *TERT.*

References

Gradishar, W. J. et al. JNCCN 12, 542-590 (2014).

Weinhold, N., et al. Nature genetics, doi:10.1038/ng.3101 (2014).

Polak, P. et al. Nature 518, 360-364, doi:10.1038/nature14221 (2015).

Wilks, C. et al. Database 2014, doi:10.1093/database/-bau093 (2014).

Cibulskis, K. et al. Nature biotechnology 31, 213-219, doi:10.1038/nbt.2514 (2013).

Ding, L. et al. Nature 464, 999-1005, doi:10.1038/nature08989 (2010).

Neph, S. et al. Nature 489, 83-90, doi:10.1038/nature11212 (2012).

T. C. G. A. Research Network. Nature 490, 61-70, doi:nature11412 (2012).

Acknowledgements

This work was supported by NIH grants UL1TR000100 and P30CA023100.

We thank Aquila Fatima, Mahdieh Khosroheidari, Sharmeela Kaushal, Scott Vandenberg and Marika Smith for assistance.

We thank Richard Kolodner, Erin Smith and Olivier Harismendy for comments.

We thank Bing Ren, Haruhiko Ishii and Chloe M Rivera for providing the PHI74 and pBABE-Puro vectors.