Unique Somatic Mutation Genotype Profile in Kentucky Patients with Colon Adenocarcinoma

Megan Harper, Daheng He, Kaylin Batey, Emily Cassim, Miranda Lin, Reema Patel, Michael Cavnar, Prakash Pandalai, Chi Wang, Erin Burke, Joseph Kim

Background: Kentucky has the highest nationwide rates of cancer, including colon adenocarcinoma (COAC). It remains unclear why Kentuckians, in particular Appalachians, have higher incidence and worse outcomes despite COAC screening frequencies similar to national averages. We sought to identify whether genetic differences in Kentucky COAC patients contribute to these disparities.

Methods: To demonstrate and corroborate genomic findings, we compared COAC cases with somatic genomic analyses from The Cancer Genome Atlas (TCGA, N=338) and Genomics Evidence Neoplasia Information Exchange (GENIE, N=3185) to data from our institutional Kentucky Cancer Registry (2005-2019, N=72). Rectal cancers were excluded given different pathological behaviors and response to treatments compared to COAC. GENIE sites included all North American sites (US and Canada) and data as of April 2021. We evaluated 32 established COAC oncogenes. We included genes related to mismatch repair (MMR) and the highest/lowest mutation frequencies in each database. MSI status was also evaluated but was unavailable for GENIE. Frequencies between the genes in the 3 datasets were compared using Fischer's exact test. P-values were adjusted using the false discovery rate (FDR).

Results: The most frequent COAC mutations in TCGA/GENIE (i.e., APC, KRAS, TP53, SMAD4, PIK3CA, BRAF, and PTEN) were consistent with prior reports, while 7 of 12 most frequent mutations in Kentuckians (i.e., ARFRP1, AURKA, BCL2L1, CDK8, FLT3, SRC, and ZNF217) were unique and did not include SMAD4, BRAF, or PTEN. Interestingly, the APC mutation was found in 91.2% in Kentucky Appalachians, while it was only seen in 63.9% of Kentucky non-Appalachians, 71.9% in TCGA, and 69.2% in GENIE (all $p<0.05$). Kentucky mutation frequencies were significantly higher for ARFRP1, AURKA, BCL2L1, CDK8, FLT3, and SRC and lower for KMT2C/MLL3 and KMT2D/MLL2 compared to both TCGA and GENIE. Interestingly, 10 of the 12 most common mutations in Kentuckians were in oncogenes while more than half of the mutations in TCGA/GENIE were in tumor suppressor genes (TSGs). Furthermore, the functional pathways in mutated Kentucky genes affected DNA repair (but not MMR), apoptosis, and mTOR, which are typical of mutations induced by tobacco products and coal mining, both of which are prevalent in Kentucky. This contrasts with TCGA/GENIE, which more frequently demonstrated mutations affecting Wnt/β-catenin pathways and epigenetic processes. Although there was no difference in frequency of MMR mutations between the groups, Kentuckians had MSI-High in 4 of 60 patients tested (6.7%), while TCGA showed MSI-High in 58 out of 336 patients (17.3%) ($p=0.035$).

Conclusions: Here we report a unique COAC mutation profile in Kentuckians, which has both therapeutic and cancer prevention implications. Despite previously reported higher rates of TSG mutations in various cancers, here Kentuckians displayed a propensity for mutations in oncogenes, particularly those associated with tobacco use and coal mining. Altogether these findings provide foundation for future studies exploring potential social, and environmental factors that may influence the incidence and survival discrepancies in Kentucky.

Gene Name	TSG/ OG	KY	TCGA			GENIE		
		% patients mutated	% patients mutated	KY OR	FDR *p*-value	% patients mutated	KY OR	FDR *p*-value
ARFRP1	OG	16.7%*	7.7%	**2.39**	**0.030**	3.5%	**5.46**	**<0.001**
AURKA	OG	16.7%*	7.4%	**2.50**	**0.029**	3.6%	**5.34**	**<0.001**
BCL2L1	OG	20.8%*	9.8%	**2.43**	**0.024**	4.2%	**6.01**	**<0.001**
CDK8	OG	16.7%*	5.9%	**3.17**	**0.023**	4.0%	**4.80**	**<0.001**
FLT3	OG	16.7%*	6.8%	**2.73**	**0.023**	5.6%	**3.35**	**0.001**
KMT2C/ MLL3	TSG	1.4%	12.1%*	**0.10**	**0.023**	8.5%	**0.15**	**0.034**
KMT2D/ MLL2	TSG	0.0%	11.8%	**0.00**	**0.009**	15.5%*	**0.00**	**<0.001**
SRC	OG	16.7%*	6.8%	**2.73**	**0.023**	3.7%	**5.27**	**<0.001**

Table 1. Comparison of Colon Adenocarcinoma Somatic Mutations in Kentucky, TCGA, and GENIE. Kentucky patients present with tumor mutations more commonly in ARFRP1, AURKA, BCL2L1, CDK8, FLT3, and SRC, and less frequently in KMT2C/MLL3 and KMT2D/MLL2 compared to national databases, as denoted by KY Odds Ratio (OR). Additionally, mutations in Kentuckians were more common in oncogenes (OG) compared to TCGS/GENIE. GENIE sites included all North American sites and data as of April 2021. Bolded cells are significant (false discovery rate (FDR) $p<0.05$). *included in top 12 most frequently mutated genes for