

Evolutionarily conserved clusters of colon with lung cancer susceptibility loci, linked with most *DUSP* phosphatase genes, may help to dissect mechanisms of cancer susceptibility

Lei Quan¹, Miranda Lynch², Peter Demant³

¹Tianjin Key Laboratory of Exercise Physiology and Sport Medicine, Research Center for Exercise and Health Science, Tianjin University of Sport, Tianjin 310617, China

²Hauptman-Woodward Medical Research Institute, Buffalo, New York, U.S.A.

³Roswell Park Comprehensive Cancer Center, Buffalo, New York, U.S.A.

Abstract

We show evolutionarily conserved pairwise genetic linkage and clustering of majority of colon and lung cancer susceptibility QTLs in mice, rats and humans. The patterns of susceptibility or resistance to these two cancers in recombinant congenic mouse strains were concordant and the responsible susceptibility loci closely linked, in spite of completely different carcinogens and protocols used for induction of the two tumors. Most *DUSP* (Dual specificity phosphatase) genes are linked to these clusters. These data suggest that an important part of colon and lung cancer susceptibility is controlled by related and evolutionarily conserved processes.

Family studies and animal experiments have shown that whereas the familial cancers comprising less than 10% of cases are caused by germ-line mutations of high penetrance cancer genes, the common cancers are under control of multiple susceptibility genes^{1,2,3} with limited effects. Subsequent genome wide association studies (GWAS) in humans have indicated map positions of many of these genes⁴, but their identity and functions have not yet been sufficiently clarified. We show here in mice, rats, and humans that most colon and lung cancer susceptibility loci exhibit four unexpected genetic characteristics: a. They are mostly linked pairwise together in genome b. Their linkage and clustering in multiple short chromosomal regions are evolutionarily conserved for >70 million years c. Their effects on susceptibility to colon and lung tumors in recombinant congenic strains are highly correlated d. Most *DUSP* (*Dual specificity phosphatase*) genes are significantly and in many instances tightly linked with these colon-lung cancer susceptibility gene clusters as well as with susceptibility genes for other cancers. A comparative study of effects of these clustered genes in different species and organs may clarify the mechanisms of their action in tumorigenesis.

1. Co-mapping of colon and lung cancer susceptibility loci. Before the GWAS technologies became broadly used, we have mapped in mice numerous susceptibility genes for colon and lung cancer using recombinant congenic (RC) strains³. These linkage studies⁵ (Fig. 1A) showed that 12 of the 15 detected colon cancer susceptibility loci (mostly used locus symbol *Scc* - *Susceptibility to colon cancer*) and 22 of the published 41 lung cancer susceptibility loci (mostly used locus symbols *Sluc* – *Susceptibility to lung cancer* or *Pas* – *Pulmonary adenoma susceptibility*) are frequently pair-wise linked together⁵, forming 19 clusters (14 of them shorter than 2.5 cM) that included also 10 pre-GWAS human and 3 rat⁶ colon cancer susceptibility loci, each cluster containing one or more colon as well as one or more lung cancer susceptibility locus (Fig. 1B).

This genetic clustering was obvious in spite of different species, induction protocols, and in spite of mouse colon and lung tumors being analyzed in crosses of different strains, mostly BALB/c and STS/A for colon tumors, and mainly O20/A and C57BL/10 or A/J and C57BL/6 for lung tumors^{3,5}, and by different carcinogens: colon cancers by 1,2 dimethyl-hydrazine

(DMH), or azoxymethane (AOM)^{5,7,8} and lung cancers by N-ethyl-N-nitroso-urea (ENU), or ethylcarbamate (urethane)^{5,9} (Fig. 1C).

2. GWAS data confirmed the intra- and inter-species co-localization of colon and lung cancer susceptibility loci. New GWAS data based on the reliable standard p-value threshold of 5×10^{-8} revealed 31 novel colon cancer susceptibility loci in humans¹⁰. These data conformed with our previous report⁵ of evolutionary conservation of co-localization of colon and lung cancer susceptibility¹¹. We projected the map positions of these 31 human colon cancer susceptibility loci on their orthologous locations on mouse chromosomal map. Out of 24 new informative human colon cancer susceptibility loci 23 mapped to the vicinity of mouse *Sluc* loci, so with the previous data this revealed in total 27 clusters of colon-lung cancer susceptibility genes containing loci from human, mouse and rat¹¹. Nine of these 27 clusters were < 2.0 cM long.

Additional GWAS^{12,13} detected 33 new human colon cancer susceptibility loci. Combining GWAS SNPs with Oncotarget markers¹⁴ revealed 18 new human lung cancer susceptibility loci detected in populations of European origin, and 5 lung cancer susceptibility genes were detected in populations of Asian descent¹⁰, all at significance $< 5 \times 10^{-8}$. Remarkably, all these 23 new human lung cancer susceptibility loci^{10,14} were clustered with colon cancer susceptibility loci (Table 1). Moreover, the lung cancer susceptibility locus *VT11A*¹⁰ was independently identified as a colon cancer susceptibility locus¹⁵.

Collectively, this data comprised 159 chromosomal locations of colon and lung cancer susceptibility loci detected by linkage studies in humans, mouse, and rat and by GWAS in human. When their chromosomal locations were projected on the mouse genome, they formed 42 clusters (Table 1).

Inter-species clustering. 32 of the 42 clusters (76%) contained loci from at least two species (Table 1). This interspecies clustering likely reflects an extensive conservation of chromosomal positions of the colon and lung cancer susceptibility loci over >70 million years of separate evolution of mice, rats, and humans.

Inter-organ clustering. 37 of the 42 clusters (83.3%) contained both colon and lung cancer susceptibility loci (Table 1). Among the total of 159 analyzed loci, 77/95 colon cancer loci co-localized with at least one lung cancer locus: 22 of them within <1 cM and 18 within 1-

3 cM; 56 of 64 lung cancer loci co-localized with at least one colon cancer locus: 22 of them within <1 cM and 12 within 1-3 cM. In order to investigate whether there is a preponderance of lung cancer loci in some clusters and colon cancer loci in others, we tested correlation of numbers of the two types of loci in the clusters that were defined in Table 1. If the two types of loci would be preferentially present in different clusters, this correlation would be negative. However, this was not the case, as this correlation was weakly but significantly positive ($r=0.329$, $p=0.033$).

Statistical evaluation of colon-lung cancer susceptibility clustering. The most numerous clusters were the very tight ones: 54/159 loci were in clusters less than 1 cM long, which included 10 clusters shorter than 0.1 cM containing 21 loci, 8 clusters 0.1 – 0.5 cM long containing 18 loci, , and 7 clusters 0.5 - 1.0 cM long containing 15 loci. In contrast, the average distance of the 159 cancer susceptibility loci in the total autosomal mouse genome (1441 cM) was 9.06 cM. Only 16 of the 159 loci did not belong to any cluster and they were >10 cM distant from another colon or lung susceptibility locus.

The statistical analysis by Monte Carlo simulations examining the clustering of colon and lung cancer susceptibility loci using 100,000 simulated databases at window widths of 1, 2, 5, 10, 15, and 20 cM indicated that the p values of the deviation of the actual distribution of the colon and lung susceptibility loci from random distribution at these window widths were 0.00067, 0.093, 0.070, 0.056, 0.799, and 0.44, respectively. This agreed with the numerous short inter-specific clusters of colon and lung cancer susceptibility loci mentioned above and shown in Table 1.

3. Concordant RC strain susceptibility to colon and lung tumors. Due to different anatomical structure of colon and lung and different growth patterns of colon and lung tumors, the tumor parameters most affected by genetics are colon tumor numbers per mouse^{3,5,8,11} and lung tumor load per lung per mouse and lung tumor number per mouse^{3,5,9}. The reason for it is that while mouse colon tumors mostly keep growing as separate structures, growing lung tumors fuse with neighboring tumors and form large confluent masses. Previously we have shown in two independent lung tumor induction experiments using Wilcoxon rank sums two sample tests that the the colon tumor susceptible strain CcS-19 has significantly higher lung tumor load than the colon tumor resistant strain CcS-20 ($p<0.0001$) and the colon tumor susceptible strain CcS-11 has a

higher lung tumor load than the colon tumor resistant strain strain CcS-10 ($p=0.0012$) – see Fig. 2B in ref⁵, and the same difference caused by these strains in a separate experiment measuring lung tumor numbers (Support. Fig. 1 and Support. Table 1) in ref.⁵

To analyze comprehensively the RC strain differences in lung tumor susceptibility in relation to those in colon tumor susceptibility we compared the differences in lung tumor numbers caused by genotypes of the strains CcS-10, CcS-11, CcS-19 and CcS-20 by the t-distributed linear contrast test, that allows to test whether the strains CcS-10 and CcS-20 differ in lung tumor numbers from the strains CcS-11 and CcS-19.

Results from the linear contrast analysis of the lung tumor count data Data consist of tumor counts as outcome variable, with listing of strain (CcS-10, CcS-11, CcS-19, and CcS-20) and induced tumor type (lung).

The analysis of interest was to examine the specific contrast of testing for a difference between Strains (CcS-10, and CcS-020) versus Strains (CcS-11 and CcS-19), among the lung cancer induced mice. The results from testing this contrast are given here:

```
> summary(G, test=adjusted('single-step'))

Simultaneous Tests for General Linear Hypotheses

Multiple Comparisons of Means: User-defined Contrasts

Fit: lm(formula = TumorCts ~ Strain, data = lungdat)

Linear Hypotheses:

c10c20vsc11c19 == 0, Estimate = -6.470, Standard error
1.718, t value= -3.765, Pr(>|t|) 0.000491 *** (44 degrees of
freedom)
---
(Adjusted p values reported -- single-step method)
```

These results indicate that there is a significant value for the contrast, which would be interpreted as demonstrating that, among the mice treated to induce lung tumors, the two strains CcS-10 and CcS-20 differ significantly in tumor counts compared to the two strains CcS-11 and CcS-19.

These aren't pairwise comparisons, but rather using the contrast formulation to aggregate the strains of interest to examine for significant differences in tumor counts.

The independent genetic mapping of lung tumor susceptibility in a cross between strains highly susceptible and highly resistant to colon tumors revealed *Sluc/Pas* loci that co-localized with *Scs* loci. We tested whether the concordant resistance of CcS-10 to colon and lung tumors and concordant susceptibility of CcS-19 to these two tumors are caused by similar sets of loci. CcS-10 and CcS-19 are RC strains each containing a different set of 12.5% genes of the strain STS (highly susceptible to colon tumors) and 87.5% of BALB/c genes^{3,5,16}. To determine whether the lung tumor susceptibility's concordance with colon tumor susceptibility in these two strains is caused by related loci, we performed an independent linkage study of lung tumor susceptibility in F2 hybrids between the lung and colon tumor resistant strain CcS-10 and the lung and colon tumor susceptible strain CcS-19. We induced lung tumors in 226 (CcS-10 x CcS-19)F2 hybrids and by their genotyping we detected and mapped 7 *Sluc* loci with significant main effect on lung tumor size or load in these mice. Every detected *Sluc* locus was closely linked with a colon cancer susceptibility locus, either of mouse (Chr. 1 *Sluc-5b* – *Scs3*, Chr. 2 *Sluc31* – *Scs1*, Chr.10 *Sluc29* - *Scs14*, Chr. 18 *Sluc14* – *Scs5*), or of human (Chr. 4 *Pas9* – 9p21.3, Chr. 5 *Sluc33* – 12q24.21, Chr. 8 *Sluc9* – 16q24.1) (Fig. 1 and Table 1). No other significant loci were detected in this cross. This strongly suggests that in CcS-10 and CcS-19 the susceptibility to the two tumors is influenced by the effects of the paired colon and lung cancer susceptibility loci^{5,11}.

4. Linkage of *DUSP* family of protein phosphatase genes with colon and lung cancer susceptibility loci. A surprising feature of the clusters of colon and lung cancer susceptibility genes reported here is their frequent and tight linkage with *DUSP* (*Dual specificity phosphatase*) genes (Table 2). Many DUSPs are closely linked to these clusters or located within them. DUSP enzymes remove phosphate groups from phosphorylated serine, threonine, and tyrosine from various proteins, thereby significantly modifying their functions¹⁷. *DUSP* genes encode six types of dual specificity phosphatases¹⁷ and the close linkage with colon and lung cancer susceptibility genes applies to all of them. Table 2 shows

their chromosomal location in the mouse and their distance in cM from the projected location of the nearest mouse, rat, or human colon or lung cancer susceptibility locus (as listed in Table 1). *DUSP* enzymes have manifold effects in development and progression of cancer^{18,19,20} that include also modulation of effects of mitogen induced kinases. Out of 45 *DUSP* genes, 13 (29%) are within less than 1 cM of a colon or lung cancer susceptibility locus, 20 (44%) are within less than 2 cM of such locus, and 33 (73%) within less than 5 cM of such locus. Several of them are located inside the clusters of colon-lung cancer susceptibility loci. This very close proximity of many *DUSP* genes with colon or lung cancer susceptibility genes suggests that some of them may be candidates for these susceptibility genes. In fact, some *DUSP* genes have been directly implicated as cancer susceptibility genes. *DUSP10* has been identified in GWAS as a human colon cancer susceptibility gene and it is 1.2 – 1.4 cM distant from the mouse colon and lung cancer susceptibility loci *Scs3* and *Sluc5b*²¹, respectively. *DUSP4* has also been suggested in an early GWAS²² as a potential candidate gene for human colon cancer susceptibility gene at 8p12, and although falling short of the standard significance threshold, it co-localizes also with a human lung cancer susceptibility gene at 8p12 in cluster 19 (Table 1). The *DUSP* family member *PTEN* has tumor suppressor effects and it is the second most frequently deleted gene in cancers¹⁷.

We evaluated statistically the tendency of *DUSP* genes to occur in proximity to the colon/lung susceptibility loci via Monte Carlo Simulation methods. Specifically, we used simulation to investigate whether the actual position of *DUSP* genes vis-à-vis the susceptibility loci were in closer proximity than would be expected if the same *DUSP* loci occurred in random positions in the genome. We used the arithmetic mean distances of simulated random *DUSP* positions to the nearest colon-lung susceptibility locus and compared them to the true mean distances of *DUSP* genes to the nearest colon-lung susceptibility locus. We found that the true *DUSP* positions are on average closer to colon-lung susceptibility loci than would be expected by random chance, with empirical p-value of 0.02, based on 100,000 simulated datasets. In summary, these analyses show that *DUSP* genes are significantly more likely to occur in close proximity to colon or lung loci than would be expected by chance, and that the true mean values of distances of *DUSP* genes to susceptibility loci for colon and lung cancer are substantially below those means in a collection of randomly distributed *DUSP* positions.

As 7 out of 45 *DUSP* genes (2 of them located on X chromosome) were not linked to any of the colon or lung susceptibility loci, we investigated whether these and other *DUSP* genes were linked to susceptibility genes for cancers of other organs using their map positions reported in ref⁴. As shown in the right-hand half of the Table 2, with the exception of the two X-linked loci *DUSP9* and *DUSP21* and *CDC14A* (on human chromosome 1), every *DUSP* gene is linked to at least one cancer susceptibility gene. A possible corollary of this observation is that many of the susceptibility genes for cancers in other organs map close to the colon-lung cancer susceptibility clusters.

Discussion. We highlighted here four previously non-reported aspects of genetics of colon and lung cancer susceptibility: significant pairwise clustering of susceptibility loci for colon and lung cancer, its conservation in three species for >70 million years, correlated direction of phenotypic effect of the paired loci on both colon and lung tumorigenesis in recombinant congenic strains, and frequent association of *DUSP* genes with colon and lung cancer susceptibility genes. These data offer a possibility to extend the understanding of individual genetic susceptibility to colon and lung cancers, its molecular and functional basis, and the responsible genes, which at the present are not sufficiently defined.

The data from three species that strongly indicate an evolutionary conserved co-localization of most colon and lung cancer susceptibility genes and the concordant direction of effects of the paired loci is compatible either with a single susceptibility gene or shared conserved regulatory elements, although more complex explanations will likely emerge too.

The precision of map positions of genes in Table 1 may be negatively affected by mapping the same gene in different studies and different species using different but close DNA markers and by intra- and inter-species indel polymorphisms. Moreover, many SNPs most significantly associated with QTLs are close to genomic regulatory elements and it is conceivable that different elements may control the same gene in different organs (or species). Finally, some cancer QTLs represent complex of interacting genes and or regulatory elements spanning several megabases of genome (reference⁴, Figure 3). These complications, however, do not invalidate the evidence for clustering, as the length of inter-cluster intervals far exceeds the potential mapping imprecisions. To the contrary, this suggests that many members of a cluster may be more tightly linked than the Table 1 shows.

Sud et al.⁴⁴ reported that a sizeable proportion of SNPs is associated with multiple cancers and we showed here that the pair-wise association of colon and lung cancer susceptibility loci is evolutionary conserved and hence likely based on shared functional pathways, most of which are not yet known. We did not investigate systematically whether a similarly strong clustering of susceptibility loci for other tumor types, but data in Table 2 suggest that it may be detected later. A shared heritability of six solid cancers was indicated by meta-analysis of GWAS data²³, but no information about the loci involved was presented. Our experimental design in mice^{3,5} permits in the future to assay separately the effects of different alleles of individual loci on different aspects of colon and lung cancer susceptibility in terms of quantitative and qualitative tumor pathology, their transcriptional and epigenetic effects in the target organs and tumors (for example^{24,25}), somatic mutations²⁶ and inflammatory response²⁷. Results of such studies can be further expanded by targeted manipulation of candidate susceptibility genes. Analyzing effects of single susceptibility genes in mice could help to disentangle the control of individual colon and lung cancer susceptibility in humans as well. In addition, close association of some of them with *DUSP* genes can help to understand the mechanisms of their effects.

The inter-species conservation of these “early” cancer QTLs suggests that they likely operate in important biological pathways. Their high replicability makes them potentially useful for understanding the mechanisms of individual genetic risk of these two frequent cancers and for development of novel preventative and therapeutic methods. Recent high-powered GWAS meta-analyses in cancer and other diseases identified very large numbers of novel QTLs. It is possible that with numbers of QTLs for each tumor type exceeding several hundreds, their ensuing high density in genome can make it difficult to detect statistically their systemic colocalization. However, the future contribution of these new QTLs to personal medicine requires that they exhibit genetic stratification²⁸, but until now the predominant contribution to genetic risk has been provided mainly by the early detected QTLs, because they tend to have larger phenotypic effects^{29,30}. A study of effect-size distribution of susceptibility QTLs of 14 cancers suggests that in spite of considerable quantitative differences between cancer types, the cumulative effects of stronger and thus “earlier” QTLs will likely determine the largest part of individual genetic susceptibility³¹.

The present study is based on such “medium strength” QTLs that were discovered in the early GWAS involving a relatively limited number of cases and controls^{10,12,13,14} compared to the recent meta-analyses. All GWAS-detected human cancer susceptibility loci analyzed here were described in the early genome-wide association studies^{10,12,13,14} in which they were identified with the required significance level of 5×10^{-8} . No loci from other GWAS were included. Such early detected loci appear to be exceptionally replicable, because they were not only detected in the early GWAS, but also were over-represented among the loci detected in all subsequent GWAS due to their repeated detection²⁹. Moreover, generally the overall disease predictive capacity of GWAS loci, measured as area under the ROC curve (AUC), tended to increase with the number of loci detected in the early GWAS, but addition of larger numbers of loci first detected in later GWAS is predicted to result in a much smaller improvement of overall disease predictive capacity^{29,30}. This implies that the human paired colon-lung susceptibility loci analyzed in the present study, which were all detected in the early GWAS studies, will also likely remain responsible for a relatively large proportion of genetic risk for these two cancers.

Ethics Statement

All animal experiments were approved by the IACUC committee at Roswell Park Cancer Institute (IACUC protocol 905M).

Mice

Mice received acidified drinking water (pH 2.5–3.0) and a standard laboratory diet (LM-485, Harlan Teklad, U.S.) ad libitum. Recombinant congenic (RC) strains are sets of 20 inbred strains derived from the same parental strains. Each CcS RC strain (CcS-1 through CcS-20) has 87.5% of the genome from the “background” BALB/c strain and 12.5% of the genome from the “donor” STS strain^{3,5}. Each OcB RC strain (OcB-1 through OcB-20) has 87.5% of the genome from the “background” O20 strain and 12.5% of the genome from the “donor” B10.O20 strain.

Tumor Induction and Histological Analysis

For lung tumor induction, pregnant female mice were given an intraperitoneal (i.p.) injection of 30 mg/kg body weight of the carcinogen N-ethyl-N-nitrosourea (ENU) dissolved

in phosphate-buffered citric acid (pH 5.8) at day 17 of gestation. The offspring of carcinogen-injected females were thus exposed to ENU transplacentally and their whole lungs were examined for tumors at the age of four months. Colon tumors were induced in adult mice with eight weekly subcutaneous injection of 15 mg/kg body weight of azoxymethane (AOM) or 1,2-dimethylhydrazine (DMH). Mice were euthanized after four months since the last injection and their colons were examined. All tissues were removed, fixed in 10% neutral buffered formalin and embedded in histowax. The embedded lungs and colons were sectioned semi-serially (5-mm sections at 100-mm intervals), stained with haematoxylin-eosin and examined microscopically at 50x and 400x magnifications. To distinguish unequivocally individual tumors, position of a tumor in sequential sections, its shape and size, positional relation to bronchi or blood vessels, and other characteristics of tumor cells have been used.

Genotyping

More than 90 % of the genetic material from the ‘donor’ strain in a RC strain is concentrated in 9 to 13 discrete contiguous chromosomal regions with intermediate length (5–25 cM) that are usually located on 7 to 11 different chromosomes³. The positions and lengths of the majority of the ‘donor’ strain-derived chromosomal regions in CcS and OcB RC strains have been determined with more than 500,000 microsatellite or single nucleotide polymorphism (SNP) markers across the whole genome. Based on such information, each known segregating chromosomal region is represented by at least one genetic marker. More markers have been tested in the longer donor chromosomal regions and the maximal distance between two markers was less than 10 cM.

Definition of Clusters

Susceptibility loci for colon and lung cancers were retrieved from literature and the NHGRI GWAS catalog (for human GWAS loci only). Only loci with a genome-wide significance were retained, i.e. $p < 5 \times 10^{-8}$ for GWAS loci in human or $P < 0.05$ after Bonferroni correction of multiple testing for loci in rat and mouse. Entries for which a genomic region and/or association statistic could not be unambiguously assigned were excluded. Loci identified in human and rat were projected to the mouse genome based on positions of their

orthologous regions in mice, acquired from the major genome browsers, i.e. the Jackson Laboratory (<https://informatics.jax.org>), the Rat Genome Database (<https://rgd.mcw.edu>), Ensemble (<https://www.ensembl.org>) and the UCSC genome browser (<https://www.genome.ucsc.edu>). Two or more loci are defined as a cluster when their distance is less than 10 cM. Although we set 10cM as the upper limit for the distance between paired loci within a cluster, the actual distances between most loci in a cluster are much shorter (Table 1).

Statistical analysis

Concordance of strain differences in colon and lung tumor susceptibility. Due to different anatomical structure of colon and lung and different growth patterns of colon and lung tumors, the tumor parameters most affected by genetics are colon tumor numbers per mouse^{3,5,8,11} and lung tumor load per lung per mouse^{3,5,9}. The reason for it is that while mouse colon tumors mostly keep growing as separate structures, growing lung tumors fuse with neighboring tumors and form large masses. To compare the quantitative variation of the two phenotypes in the strains CcS-10, CcS-11, CcS-19 and CcS-20, we ranked all AOM-treated individual mice for number of colon tumors and all individual ENU treated progeny for lung tumor load. These rank-normalized values in different strains were compared by the t-distributed linear contrast test at 44 degrees of freedom.

Assessment of clustering. Monte Carlo multi-window analysis was used to assess the distribution of positions of the mouse colon and lung cancer susceptibility loci in mouse autosomal genome and the orthologous positions of human and rat colon and lung cancer susceptibility loci projected on the mouse autosomal genome as listed in Table 1. We accounted for chromosome length, and only included within-chromosome distances. The test involved five window widths (1, 2, 5, 10, and 20 cM) and 100,000 simulations for each of them.

Similarly, Monte Carlo analysis of closeness of *DUSP* loci to colon/lung susceptibility loci. We took the true positions of the susceptibility loci, and the true positions of the *DUSP* loci, and computed the set of pairwise distances of a *DUSP* locus with its nearest susceptibility locus, and computed the arithmetic means of this collection of true proximity values. We then generated a random set of the same number of *DUSP* positions across the genome and computed the set of pairwise distances of a random *DUSP* position with its nearest

susceptibility locus, and computed means of this set of distances. We accounted for chromosome length, and only included within-chromosome distances in determining nearest locus to a *DUSP* position. We repeated the random set generation 100,000 times and computed the number of these random datasets that had mean value below the true means, using both arithmetic and geometric means, which provided empirical p-values for comparing true mean *DUSP* proximities to an equivalent mean of random positions.

Acknowledgements

We thank Dr. Alan Hutson, Chairperson of Department of Biostatistics and Bioinformatics of the Roswell Park Comprehensive Cancer Center for generous support, discussions and advice.

This work had been supported by the NCI grant R01 CA 115158 and an institutional grant from the Roswell Park Comprehensive Cancer Center to P. Demant, and by NCI grant P30CA016056 to the Roswell Park Comprehensive Cancer Center supporting the use of its Biostatistics and Genomic Shared Resources and by the grant 16JCYBJC from the Natural Science Foundation from the Tianjin City and Tianjin Thousand Young Talents Plan to L. Quan.

References

1. Pharoah, P. D. *et al.* Polygenic susceptibility to breast cancer and implications for prevention. *Nat Genet* **31**, 33-36, doi:10.1038/ng853 (2002).
2. Balmain, A., Gray, J. & Ponder, B. The genetics and genomics of cancer. *Nat Genet* **33 Suppl**, 238-244, doi:10.1038/ng1107 (2003).
3. Demant, P. Cancer susceptibility in the mouse: genetics, biology and implications 021521for human cancer. *Nat Rev Genet* **4**, 721-734, doi:10.1038/nrg1157 (2003).
4. Sud, A., Kinnersley, B. & Houlston, R. S. Genome-wide association studies of cancer: current insights and future perspectives. *Nat Rev Cancer* **17**, 692-704, doi:10.1038/nrc.2017.82 (2017).
5. Quan, L. *et al.* Most lung and colon cancer susceptibility genes are pair-wise linked in mice, humans and rats. *PLoS One* **6**, e14727, doi:10.1371/journal.pone.0014727 (2011).
6. De Miglio, M. R. *et al.* Identification and chromosome mapping of loci predisposing to colorectal cancer that control Wnt/beta-catenin pathway and progression of early lesions in the rat. *Carcinogenesis* **28**, 2367-2374, doi:10.1093/carcin/bgm119 (2007).
7. Bissahoyo, A. C. *et al.* A New Polygenic Model for Nonfamilial Colorectal Cancer Inheritance Based on the Genetic Architecture of the Azoxymethane-Induced Mouse Model. *Genetics* **214**, 691-702, doi:10.1534/genetics.119.302833 (2020).
8. Chen, J. & Huang, X. F. The signal pathways in azoxymethane-induced colon cancer and preventive implications. *Cancer Biol Ther* **8**, 1313-1317, doi:10.4161/cbt.8.14.8983 (2009).
9. Festing, M. F. *et al.* At least four loci and gender are associated with susceptibility to the chemical induction of lung adenomas in A/J x BALB/c mice. *Genomics* **53**, 129-136, doi:10.1006/geno.1998.5450 (1998).
10. MacArthur, J. *et al.* The new NHGRI-EBI Catalog of published genome-wide association studies (GWAS Catalog). *Nucleic Acids Res* **45**, D896-D901, doi:10.1093/nar/gkw1133 (2017).
11. Quan, L., Hutson, A. & Demant, P. Cross-Cancer Analysis Reveals Novel Pleiotropic Associations-Letter. *Cancer Res* **77**, 6042-6044, doi:10.1158/0008-5472.CAN-16-3262 (2017).
12. Huyghe, J. R. *et al.* Discovery of common and rare genetic risk variants for colorectal cancer. *Nat Genet* **51**, 76-87, doi:10.1038/s41588-018-0286-6 (2019).
13. Law, P. J. *et al.* Association analyses identify 31 new risk loci for colorectal cancer susceptibility. *Nat Commun* **10**, 2154, doi:10.1038/s41467-019-09775-w (2019).
14. McKay, J. D. *et al.* Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nat Genet* **49**, 1126-1132, doi:10.1038/ng.3892 (2017).
15. Wang, H. *et al.* Trans-ethnic genome-wide association study of colorectal cancer identifies a new susceptibility locus in VTI1A. *Nat Commun* **5**, 4613, doi:10.1038/ncomms5613 (2014).

16. Moen CJ, *et al.* The recombinant congenic strains--a novel genetic tool applied to the study of colon tumor development in the mouse. *Mamm Genome* **1**, 217-27. (1991)
17. Patterson, K. I. et al. Dual-specificity phosphatases: critical regulators with diverse cellular targets. *Biochem. J.* **418**, 475-489, doi: 10.1042/bj20082234 (2009).
18. Keyse, S. M. Dual-specificity MAP kinase phosphatases (MKPs) and cancer. *Cancer Metastasis Rev* **27**, 253-261, doi:10.1007/s10555-008-9123-1 (2008).
19. Meeusen, B. & Janssens, V. Tumor suppressive protein phosphatases in human cancer: Emerging targets for therapeutic intervention and tumor stratification. *Int J Biochem Cell Biol* **96**, 98-134, doi:10.1016/j.biocel.2017.10.002 (2018).
20. Ruvolo, P. P. Role of protein phosphatases in the cancer microenvironment. *Biochim Biophys Acta Mol Cell Res* **1866**, 144-152, doi:10.1016/j.bbamcr.2018.07.006 (2019).
21. Png, C. W. et al. DUSP10 regulates intestinal epithelial cell growth and colorectal tumorigenesis. *Oncogene* **35**, 206-217, doi:10.1038/onc.2015.74 (2016).
22. Fernandez-Rozadilla, C. et al. A colorectal cancer genome-wide association study in a Spanish cohort identifies two variants associated with colorectal cancer risk at 1p33 and 8p12. *BMC Genomics* **14**, 55, doi:10.1186/1471-2164-14-55 (2013).
23. Jiang, X. et al. Shared heritability and functional enrichment across six solid cancers. *Nat Commun* **10**, 431, doi:10.1038/s41467-018-08054-4 (2019).
24. Dampier, C.H. et al. Oncogenic Features in Histologically Normal Mucosa: Novel Insights Into Field Effect From a Mega-Analysis of Colorectal Transcriptomes. *Clin Transl Gastroenterol* **11**, e00210 (2020).
25. Diez-Obrero, V. et al. Genetic Effects on Transcriptome Profiles in Colon Epithelium Provide Functional Insights for Genetic Risk Loci. *Cell Mol Gastroenterol Hepatol* **12**, 181-197 (2021).
26. Rogers, Z.N. et al. Mapping the in vivo fitness landscape of lung adenocarcinoma tumor suppression in mice. *Nat Genet* **50**, 483-486 (2018). Zhang, Y. D. et al. Assessment of polygenic architecture and risk prediction based on common variants across fourteen cancers. *Nat Commun* **11**, 3353, doi:10.1038/s41467-020-16483-3 (2020).
27. R. Lang, M. Hammer, J. Mages, DUSP meet immunology: dual specificity MAPK phosphatases in control of the inflammatory response. *J Immunol* **177**, 7497-7504 (2006).
28. Wray, N. R., Wijmenga, C., Sullivan, P. F., Yang, J. & Visscher, P. M. Common Disease Is More Complex Than Implied by the Core Gene Omnigenic Model. *Cell* **173**, 1573-1580, doi:10.1016/j.cell.2018.05.051 (2018).
29. Marigorta, U. M., Rodriguez, J. A., Gibson, G. & Navarro, A. Replicability and Prediction: Lessons and Challenges from GWAS. *Trends Genet* **34**, 504-517, doi:10.1016/j.tig.2018.03.005 (2018).
30. Turnbull C., Sud A., Houlston R. C. Cancer genetics, precision prevention and a call for action. *Nature Genet.* **50**: 1212-1218: doi:10.1038/s41588-018-0202-0 (2018).
31. Zhang Y. D. et al. Assessment of polygenic architecture and risk prediction based on common variants across common cancers. *Nat Commun* **11**, 3353, doi: 10.1038/s41467-020-16483-3 (2020).

Figure 1. Concordant susceptibility of recombinant congenic (RC) strains to colon and lung tumors and co-localization of the responsible loci. A schematic representation of the derivation of recombinant congenic strains by random backcrossing of STS and BALB/c mice and subsequent inbreeding. The resulting RC strains carried each a different set of $\pm 12.5\%$ STS-derived genes on BALB/c genetic background. These sets overlapped only minimally with the STS-derived sets in other RC strains. Two RC strains, CcS-10 and CcS-20, were highly resistant to both colon and lung tumors, whereas strains CcS-11 and CcS-19 were highly susceptible to both tumors. Mapping of lung tumor susceptibility loci controlling difference between CcS-10 and CcS-19 in (CcS-10 x CcS-19)F₂ hybrids revealed 7 loci, all of which were closely linked to chromosomal locations of mouse or human colon tumor susceptibility loci.

Figure 1

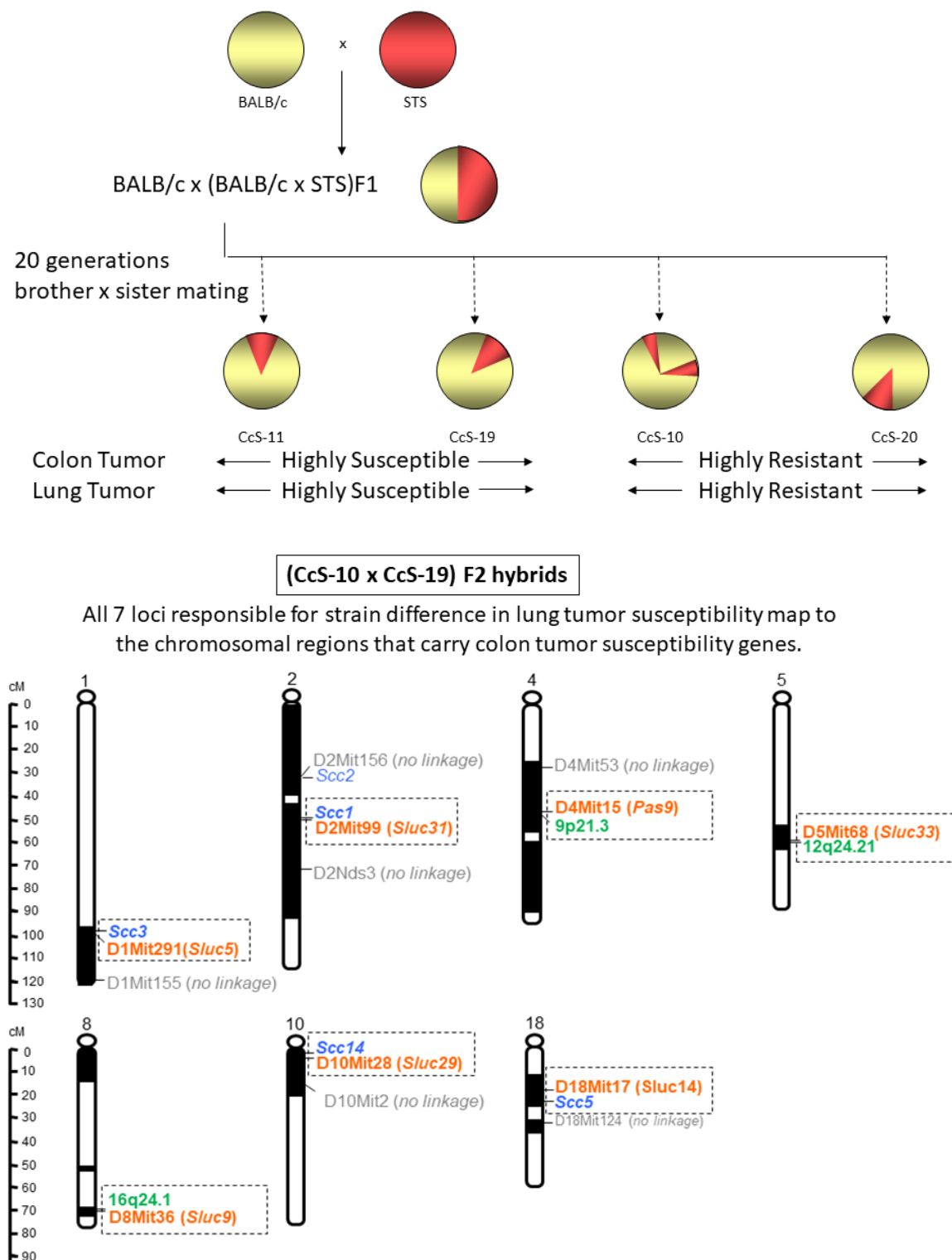


Table 1.Co-localization of colon and lung cancer susceptibility loci identified in human,mouse and rat. (6 pages)

Table 1. Co-localization of colon and lung cancer susceptibility loci identified in human, mouse and rat.							
Cluster	Cancer Susceptibility Loci			Reference	Position in Mouse ^a		
	Locus	Loci Type	Gene/Marker		Chr	cM	Mb
Chromosome 1 Total length 98.5cM							
1	Sluc15	Mouse Lung (sol ^b)	D1Mit170	PMID: 11584065	1	15.2	36.0
	2q32.2	Human Colon	NABP1	PMID: 23266556	1	26.6	51.4
	2q33.1	Human Colon	SATB2	PMID: 30510241	1	28.7	56.7
	1q25.3	Human Colon (sol)	LAMC1	PMID: 23266556, 24737748	1	65.3	153.2
2	Sluc5a	Mouse Lung (sol)	D1Mit36	PMID: 9809979	1	76.7	169.2
	1q41	Human Colon	DUSP10	PMID: 20972440	1	88.2	184.0
	Sluc5b	Mouse Lung	D1Mit221	PMID: 9809979	1	89.46	185.2
	Scc3	Mouse Colon	D1Mit208	PMID: 8944029	1	89.54	185.4
Chromosome 2 Total length 103.9cM							
3	10p14	Human Colon	GATA3	PMID:18372905	2	6.7	9.8
	Sluc16	Mouse Lung	D2Mit5	PMID: 11584065	2	6.9	9.5
4A ^c	10p13	Human Colon	CUBN	PMID: 25990418	2	9.9	13.2
	Scc2	Mouse Colon	D2Mit156	PMID: 8944029	2	31.7	57.1
4B	2q24.2	Human Colon	MARCH7 / TANC1	PMID: 30510241	2	33.9	60.2
	Sluc2	Mouse Lung	D2Mit56	PMID: 8944028	2	42	70.7
	Scc1	Mouse Colon	Ptprj	PMID: 12089527	2	50.2	90.4
	Sluc31	Mouse Lung	D2Mit99	PMID: 21390212	2	55.1	105.1
	15q13.3	Human Colon	SCG5 / GREM1 / FMN1	PMID: 18084292	2	57.4	113.7
	15q21.1	Human Lung	SEMA6D	PMID: 28604730	2	60.7	124.0
5	20p12.3	Human Colon	HAO1 / BMP2 / PLCB1	PMID: 19011631, 23263487	2	65.7	134.4
	20q13.12	Human Colon	TOX2 / HNF4A	PMID: 29471430, 30510241	2	84.1	163.2
6	20q13.13	Human Colon	PREX1	PMID: 26151821, 31089142	2	87.0	167.9
	Sluc17	Mouse Lung	D2Mit200	PMID: 11584065	2	102.3	179.7
	20q13.33	Human Colon	LAMA5	PMID: 20972440	2	102.7	180.1
	20q13.33	Human Colon	TNFRSF6B / RTEL1	PMID: 30510241	2	103.6	181.3
	20q13.33	Human Lung	RTEL1	PMID: 28604730	2	103.6	181.3
	Chromosome 3 Total length 82.7cM						
7	13q13.3	Human Colon	SMAD9	PMID: 30510241	3	25.6	54.7
	4q32.2	Human Colon	FSTL5	PMID: 25023989	3	34	76.0
8	4q24	Human Colon (sol)	TET2	PMID: 30510241, 31089142	3	61.8	133.4
	1p31.1	Human Lung	FUBP1	PMID: 28604730	3	77.1	152.2
	Scc7	Mouse Colon	D3Mit163	PMID: 10485458	3	80.7	156.6

Chromosome 4			Total length 88.6cM				
9	<i>Sluc18</i>	Mouse Lung	D4Mit4	PMID: 11584065	4	20.1	39.3
	9q22.33	Human Colon	<i>GALNT12 / TGFBRI</i>	PMID: 30510241	4	26	47.0
	9q31.3	Human Colon	<i>LPAR1</i>	PMID: 30510241	4	32.2	58.4
10A	9p21.3	Human Lung	<i>CDKN2A</i>	PMID: 28604730, 26732429	4	42.2	89.2
	9p21.3	Human Colon	<i>ANRIL / CDKN2A / CDKN2B</i>	PMID: 30510241, 31089142	4	42.2	89.2
	<i>Pas9</i>	Mouse Lung	<i>D4Mit77 / p16</i>	PMID: 9790761	4	42.2	89.2
10B	1p32.3	Human Colon	<i>TTC22 / PCSK9</i>	PMID: 30510241, 31089142	4	49.7	106.6
	<i>Scc11</i>	Mouse Colon	D4Mit11	PMID: 14562056	4	57.4	123.2
	1p34.3	Human Colon	<i>FHL3</i>	PMID: 30510241, 31089142	4	57.9	124.7
11	<i>Sluc21</i>	Mouse Lung	D4Mit70	PMID: 11584065	4	67	134.6
	1p36.2	Human Colon	<i>CDC42 / WNT4</i>	PMID: 25990418	4	69.8	137.8
	<i>Sluc6</i>	Mouse Lung	D4Mit158	PMID: 9809979	4	70	137.9
	<i>rCcs1</i>	Rat Colon	D5Rat74	PMID: 17510081	4	78	149.2
Chromosome 5			Total length 90.2cM				
12	22q12.1	Human Lung	<i>CHEK2</i>	PMID: 24880342	5	53.7	110.8
	12q24.21	Human Colon	<i>TBX3 / SH2B3 / MED13L</i>	PMID: 23266556, 26151821, 29917119	5	60.3	119.6
	<i>Sluc33</i>	Mouse Lung	D5Mit68	PMID: 21390212	5	60.4	119.8
	13q13.1	Human Lung (sol)	<i>BRCA2</i>	PMID: 24880342	5	89.2	150.5
Chromosome 6			Total length 79.0cM				
13	<i>Sluc7</i>	Mouse Lung (sol)	D6Mit158	PMID: 9809979	6	11.9	26.2
	<i>Scc13</i>	Mouse Colon	D6Mit122	PMID: 14562056	6	28.2	59.1
	4q22.2	Human Colon	<i>ATOH1 / SMARCA1</i>	PMID: 29917119	6	30.1	64.7
14	<i>Pas1c</i>	Mouse Lung	D6Mit177	PMID: 9790761	6	42.2	93.5
	3p14.1	Human Colon	<i>LRIG1</i>	PMID: 26151821	6	43.2	94.6
15	12p13.31	Human Colon	<i>CD9 / PLEKHG6 / TNFRSF1A</i>	PMID: 24836286	6	59.3	125.4
	<i>Sluc3</i>	Mouse Lung	D6Mit218	PMID: 8944028	6	61.9	127.0
	12p13.32	Human Colon	<i>CCND2</i>	PMID: 23263487, 24737748	6	61.9	127.1
16	<i>Pas1a</i>	Mouse Lung	D6Mit14	PMID: 8499946	6	77.6	145.6
	<i>Scc25</i>	Mouse Colon	<i>RASSF8</i>	PMID: 31879319	6	77.7	145.7
Chromosome 7			Total length 89.1cM				
17	19q13.43	Human Colon	<i>TRIM28</i>	PMID: 30510241	7	7.7	13.0
	<i>Sluc30</i>	Mouse Lung	D7Mit57	PMID: 11584065	7	9.9	19.6
	19q13.2	Human Colon	<i>TGFB1</i>	PMID: 24836286	7	14	25.6
	19q13.2	Human Lung	<i>CYP2A6</i>	PMID: 28604730	7	14.4	26.3

Table 1. page 2 of 6

18	11q13.4	Human Colon (sol)	POLD3	PMID: 22634755, 30510241	7	54.3	100.0
	<i>Scc12</i>	Mouse Colon	D7Mit67	PMID: 14562056	7	70.2	128.3
	<i>Sluc19</i>	Mouse Lung	D7Mit105	PMID: 11584065	7	70.3	128.5
	<i>Sluc8</i>	Mouse Lung (sol)	D7Nds4	PMID: 9809979	7	88.9	144.8
Chromosome 8 <i>Total length 76.2cM</i>							
19A	13q34	Human Colon	<i>COL4A2 / COL4A1 / RAB20</i>	PMID: 30510241, 31089142	8	5.7	11.4
	<i>Scc8</i>	Mouse Colon	D8Mit17	PMID: 10485458	8	5.9	13.3
19B	<i>Sluc20</i>	Mouse Lung	D8Mit3	PMID: 11584065	8	12.9	24.6
	8p12	Human Lung	<i>NRG1</i>	PMID: 28604730	8	18.8	31.8
	8p12	Human Colon	<i>DUSP4</i>	PMID: 23350875	8	21.2	34.8
20	19p13.11	Human Colon	<i>KLF2</i>	PMID: 30510241, 31089142	8	35	72.3
	4q31.21	Human Colon	<i>HHIP</i>	PMID: 30510241, 31089142	8	37.9	79.9
21	16q22.1	Human Colon	<i>CDH1</i>	PMID: 19011631	8	53.2	106.6
	16q23.2	Human Colon	<i>MAF</i>	PMID: 30510241, 31089142	8	62.6	115.7
	<i>Sluc9</i>	Mouse Lung	D8Mit35	PMID: 9809979	8	62.8	116.1
	16q24.1	Human Colon	<i>LOC146513 / FOXL1</i>	PMID: 29917119	8	70.3	121.1
Chromosome 9 <i>Total length 75.1cM</i>							
22	11q22.1	Human Colon (sol)	<i>YAP1</i>	PMID: 30510241	9	2.45	7.9
	<i>Sluc10</i>	Mouse Lung	D9mit2	PMID: 9809979	9	20.8	37.3
	11q23.3	Human Lung	<i>AMICA1</i>	PMID: 28604730	9	24.8	45.0
	11q23	Human Colon	<i>POU2AF1 / C11orf53</i>	PMID: 18372901	9	28	51.2
	15q25.1	Human Lung	<i>CHRNA5</i>	PMID: 18385738, 19654303, 28604730	9	29.8	54.9
	15q22.33	Human Colon	<i>SMAD3</i>	PMID: 30510241	9	34.2	63.6
	15q22.31	Human Colon	<i>SMAD6</i>	PMID: 31089142	9	34.4	64.0
23	6p12.1	Human Colon	<i>BMP5</i>	PMID: 29917119	9	42.3	75.7
	<i>Sluc11</i>	Mouse Lung	D9Mit12	PMID: 9809979	9	52.2	99.9
	3q21-q24	Human Colon	<i>EPHB1</i>	PMID: 16923799	9	54.2	101.9
	3q22.2	Human Colon	<i>SLC02A1</i>	PMID: 30510241	9	54.7	103.0
	3p22.1	Human Colon (sol)	<i>CTNNB1</i>	PMID: 26151821	9	72.2	120.9
Chromosome 10 <i>Total length 77.9cM</i>							
24	6q23-25	Human Lung	<i>ESR1, MYCT1</i>	PMID: 15272417	10	2	4.6
	<i>Scc14</i>	Mouse Colon	D10Mit75	PMID: 14562056	10	2	6.9
	<i>Sluc29</i>	Mouse Lung	D10Mit28	PMID: 11584065	10	3	9.4
25	6q22.2	Human Lung (sol)	<i>TRNA-UGC / ROS1 / DCBLD1</i>	PMID: 23143601	10	26.2	52.0
	<i>Sluc22</i>	Mouse Lung	D10Mit122	PMID: 11584065	10	54.7	103.8
	<i>Scc9</i>	Mouse Colon	D10Mit46	PMID: 10485458	10	64.7	116.3

Table 1. page 3 of 6

	12q13.3	Human Colon	<i>STAT6/LRP1/NAB2</i>	PMID: 20972440	10	74.6	127.6
Chromosome 11			<i>Total length 88.0cM</i>				
26	22q12.2	Human Lung	<i>MTMR3/HORMAD2/LIF</i>	PMID: 21725308	11	2.5	4.4
	7p13	Human Colon	<i>MYO1G/SNHG15/CCM2</i>	PMID: 30510241	11	4.3	6.5
	Scc6	Mouse Colon	D11Mit2	PMID: 10485458	11	7.3	12.3
27	Scc15	Mouse Colon	D11Mit26	PMID: 14562056	11	37.8	60.5
	17p12	Human Colon	<i>LINC00675</i>	PMID: 30510241, 31089142	11	40.6	66.9
	<i>Sluc4</i>	Mouse Lung	D11Mit15	PMID: 8944028	11	42.9	69.9
	17p13.3	Human Colon	<i>NXN</i>	PMID: 24836286	11	45.9	76.2
	<i>Par1</i>	Mouse Lung (sol)	D11Mit54	PMID: 9230199	11	59.8	96.2
28	17q24.3	Human Lung	<i>BPTF</i>	PMID: 22797724	11	70.1	107.0
	17q24.3	Human Colon	<i>LINC00673</i>	PMID: 30510241	11	77.3	113.5
	17q25.3	Human Colon	<i>RAB40B/METRNL</i>	PMID: 30510241	11	85.3	121.3
Chromosome 12			<i>Total length 63.9cM</i>				
29	14q23.1	Human Colon	<i>DACT1</i>	PMID: 26404086, 30510241	12	29.4	71.3
	<i>Ccs1</i>	Mouse Colon	D12Mit5	PMID: 7806225	12	37.2	80.9
	<i>Par3</i>	Mouse Lung	<i>Fos</i>	PMID: 9230199	12	39.7	85.4
	<i>Sluc12</i>	Mouse Lung (sol)	D12Nds2	PMID: 9809979	12	62.2	113.8
Chromosome 13			<i>Total length 67.3cM</i>				
30	9q22.2-31.2	Human Colon	FGD3	PMID: 14566058	13	25.2	49.2
	<i>Sluc23</i>	Mouse Lung	D13Mit139	PMID: 11584065	13	26.3	51.7
	5q31.1	Human Colon	<i>PITX1</i>	PMID: 23263487	13	30.1	55.8
	<i>rCcr7</i>	Rat Colon	D17Rat53	PMID: 17510081	13	32.7	62.9
	5p15.33	Human Lung	<i>TERT/CLPTMIL</i>	PMID: 19836008, 18978787, 19654303	13	40.1	73.6
Chromosome 14			<i>Total length 66.4cM</i>				
31A	10q22.3	Human Colon	<i>ZMIZ1</i>	PMID: 24836286, 3108914	14	15	25.4
	<i>Sluc13</i>	Mouse Lung	D14Mit120	PMID: 9809979	14	20.9	36.1
	14q22.2	Human Colon	<i>BMP4</i>	PMID: 19011631	14	24	46.3
	<i>rCcr2</i>	Rat Colon	D17Rat31	PMID: 17510081	14	25.4	49.1
31B	13q12.12	Human Lung	<i>MIPEP/TNFRSF19</i>	PMID: 21725308	14	32	60.7
	8p21.1	Human Lung	<i>EPHX2/CHRNA2</i>	PMID: 28604730	14	34.4	66.1
	13q22.1	Human Colon (sol)	<i>KLF5</i>	PMID: 30510241, 31089142	14	49.4	99.2
Chromosome 15			<i>Total length 59.0cM</i>				
32	<i>Sluc24</i>	Mouse Lung	D15Mit13	PMID: 11584065	15	1.8	3.4
	5p13.1	Human Colon	<i>LINC00603/PTGER4</i>	PMID: 29917119	15	2	5.2
33	8q23.3	Human Colon	<i>EIF3H</i>	PMID: 18372905, 26151821	15	19.4	51.7

Table 1. page 4 of 6

34A	8q24	Human Colon	<i>POU5F1A1 / MYC</i>	PMID: 18372901, 17618283, 24836286	15	26.2	61.9
	<i>Sluc25</i>	Mouse Lung	D15Nds3/D15Mit17	PMID: 11584065	15	26.2	61.9
	<i>Sluc26</i>	Mouse Lung	D15Mit96	PMID: 11584065	15	42	87.4
34B	12q12	Human Colon	<i>PRICKLE1 / YAF2</i>	PMID: 30510241	15	47.4	93.2
	12q13.13	Human Colon	<i>LARP4 / DIP2B / ATF1</i>	PMID: 20972440	15	56.2	100.0
	12q13.13	Human Lung	<i>ACVR1B</i>	PMID: 26732429	15	56.5	101.2
	<i>Sluc34</i>	Mouse Lung	D15Mit16	PMID: 21390212	15	58.1	102.9
Chromosome 16 <i>Total length 57.8cM</i>							
35	3q28	Human Lung	<i>TP63</i>	PMID: 20871597, 21725308, 22797724	16	17.4	25.6
	3q29	Human Lung	<i>C3ORF21</i>	PMID: 20876614	16	21.3	30.9
	3q13.2	Human Colon	<i>BOC</i>	PMID: 30510241	16	28.8	44.4
	<i>Sluc27</i>	Mouse Lung (sol)	D16Mit19	PMID: 11584065	16	45.4	78.8
Chromosome 17 <i>Total length 61.3cM</i>							
36A	6q27	Human Lung	<i>RNASET2</i>	PMID: 28604730	17	4.9	8.1
	6q26-q27	Human Colon	<i>SLC22A3</i>	PMID: 21242260	17	8.5	12.4
	5q21.1	Human Colon	<i>RGMB / CHD1</i>	PMID: 30510241	17	8.9	15.7
36B	6p21	Human Colon	<i>CDKN1A</i>	PMID: 22634755	17	15.2	29.0
	6p21.32	Human Colon	<i>HLA-DRB1 / HLA-DQA1 / FKBP5</i>	PMID: 30510241, 31089142, 29917119	17	18	34.2
	6p21.1-33	Human Lung	<i>BAG6 / BAT3 / MSH5</i>	PMID: 18978787, 19654303, 26732429	17	18.6	35.1
37	<i>Scc4</i>	Mouse Colon	D17Mit72	PMID: 8944029	17	49.8	78.9
	<i>Sluc32</i>	Mouse Lung	D17Mit72	PMID: 21390212	17	49.8	78.9
Chromosome 18 <i>Total length 59.4cM</i>							
38	<i>rCcr3</i>	Rat Colon	D18Rat38	PMID: 17510081	18	17.4	30.2
	<i>Sluc14</i>	Mouse Lung	D18Mit17	PMID: 9809979	18	21.1	39.4
	<i>Scc5</i>	Mouse Colon	D18Mit57	PMID: 8944029	18	25.5	48.8
39A	<i>Par2</i>	Mouse Lung	D18Mit9	PMID: 9230199	18	42.6	68.5
	<i>Pas7</i>	Mouse Lung	D18Mit188	PMID: 9790761	18	45.9	72.6
39B	18q21	Human Colon	<i>SMAD7</i>	PMID: 17934461, 18372901	18	51.1	75.3
	<i>Sluc28</i>	Mouse Lung	D18Mit7	PMID: 11584065	18	51.9	76.8
Chromosome 19 <i>Total length 56.9cM</i>							
40	11q12.2	Human Colon (sol)	<i>MYRF / FEN1 / FADS1</i>	PMID: 24836286	19	6.5	10.1
	10q11.23	Human Colon (sol)	<i>AICF</i>	PMID: 29917119	19	26.6	31.8
	10q24.2	Human Colon	<i>ABCC2 / MRP2</i>	PMID: 24737748	19	36.7	43.7
41	<i>Sluc1</i>	Mouse Lung	D19Mit9	PMID: 8944028	19	39	46.7
	10q24.3	Human Lung	<i>OBFC1</i>	PMID: 28604730	19	39.6	47.5
	10q25.2	Human Colon	<i>TCF7L2 / VTI1A</i>	PMID: 24836286, 25105248, 3108914	19	51.6	55.7

Table 1. page 5 of 6

10q25.2	Human Lung	<i>VTIIA</i>	PMID: 23143601	19	51.6	55.3
10q26.12	Human Colon	<i>HSPA12A</i>	PMID: 23263487	19	54.5	58.7
<i>Sluc35</i>	Mouse Lung	D19Mit6	PMID: 21390212	19	54.6	61.0
<p>^a Positions of mouse loci or mouse orthologues of human/rat loci are listed. Although we set 10cM as the upper limit for the distance between paired loci within a cluster, the actual distances between many loci are much shorter. Locus positions within 1cM or 3cM of a paired locus are highlighted using a yellow color with or without brown boxes, respectively.</p> <p>^b Solitary loci that do not belong to any cluster are also listed with no special color indication. Sol: solitary.</p> <p>^c Some long clusters are divided into sub-clusters when there are 2 parts that are at least 5cM away from each other and each part contains at least 2 loci.</p>						
TUMOR TYPE AND SPECIES			DISTANCE OF LOCI			
COLON	Human		0 - 1 CM			
	Mouse					
	Rat		1.1 - 3.0 cM			
LUNG	Human					

Table 2. Dusp gene family and cancer susceptibility genes. (2 pages)

Table 2 Dusp gene family and cancer susceptibility genes

Symbol	Mouse Genome					Human Genome				
	Chr.	Position (Mb/cM)	Co-localizing Loci C&L ^a	Type ^b	Distance (Mb/cM)	Chr.	Position (Mb)	Co-localizing Other Cancer Loci (Pos in Mb) ^c	Type of Cancer	Distance (Mb)
MAP Kinase Phosphatase										
Dusp1	17	26.5 / 13.3	6p21 (CDKN1A)	HC	2.5 / 1.9	5	172.8	NA	NA	NA
Dusp2	2	127.3 / 61.9	15q21.1 (SEMA6D)	HL	3.3 / 1.2	2	96.1	NA	NA	NA
Dusp4^d	8	34.8 / 21.2	8p12 (DUSP4)	HC	0 / 0	8	29.3	8p21.2 (26.0), 8p21.1 (27.5), 8p12 (29.6-32.6)	Breast, Prostate, Thyroid	0
Dusp5	19	53.5 / 48.0	10q25.2 (TCF7L2/VT11A)	HC&HL	1.8 / 3.6	10	110.5	10q25.1 (105.8), 10q25.2 (112.5-114.5)	Melanoma, Breast, Glioma	2
Dusp6	10	99.3 / 51.1	Sluc22 (D10Mit122)	ML	4.5 / 3.6	12	89.3	12q21.3 (88.5)	Testicle	0.8
Dusp7	9	106.4 / 57.5	3q22.2 (SLCO2A1)	HC	3.4 / 2.8	3	52.0	NA	NA	NA
Dusp8	7	142.8 / 87.6	Sluc8 (D7Nds4)	ML	2.0 / 1.3	11	1.5	11p15.5 (1.8-2.9)	Bladder, Breast, Prostate, Leukemia	0.3
Dusp9	X	73.6 / 37.4	NA	NA	NA	X	153.6	NA	NA	NA
Dusp10	1	184.0 / 88.2	1q41 (DUSP10)	HC	0 / 0	1	221.7	1q41 (221.8), 1q42.13 (228.9)	Non-melanoma skin cancer	7.2
Dusp16	6	134.7 / 65.8	12p13.2 (CCND2)	HC	7.6 / 3.7	12	12.4	12p13.1 (14.2)	Breast, Testicle	1.8
Dusp24/Styx11	5	135.7 / 75.4	NA	NA	NA	7	75.9	NA	NA	NA
Atypical Dusp										
Dusp29/Dupd1	14	21.7 / 11.8	10q22.3 (ZMIZ1)	HC	3.7 / 3.2	10	75.0	10q22.3 (79.1)	Breast	4.1
Dusp3	11	102.0 / 65.5	Par1 (D11Mit54)	ML	4.8 / 5.7	17	43.7	17q21.2(41.7), 17q21.3(45.4-48.3)	Esophagus, Ovary, Prostate	0.6
Dusp11	6	85.9 / 37.5	Pas1c (D6Mit177)	ML	7.6 / 4.7	2	73.7	2p14 (67.4)	Pancrea	6.3
Dusp12	1	170.9 / 77.3	Sluc5a (D1Mit36)	ML	1.7 / 0.6	1	161.7	1q24.1(165.9)	Testicle	4.2
Dusp13	14	21.7 / 11.9	10q22.3 (ZMIZ1)	HC	3.7 / 3.1	10	75.1	10q22.3 (79.1)	Breast	4
Dusp14	11	84.0 / 51.3	17p13.3 (NXN)	HC	7.8 / 5.4	17	37.4	17q12 (37.7), 17q21.1 (39.9)	Ovary, Prostate, Endometria, Cervix	0.3
Dusp15	2	152.9 / 75.4	NA	NA	NA	20	31.8	20q11.22 (34.0-34.6)	Breast, Skin, Melanoma	2.2
Dusp18	11	3.9 / 2.7	22q12.2 (MTMR3)	HL	0.5 / 0.2	22	30.6	22q12.1(28.7-28.9), 22q12.2 (29.2-29.9)	Esophagus, Pancrea, Breast	0.7
Dusp19	2	80.6 / 48.3	Scc1 (Ptprj)	MC	9.8 / 1.9	2	183.1	2q31.1 (176.2), 2q32.2 (191.1)	Ovary, Liver	6.9
Dusp21	X	18.1 / 13.4	NA	NA	NA	X	44.8	NA	NA	NA
Dusp22	13	30.7 / 13.3	NA	NA	NA	6	0.3	6p25.3 (0.4-1.2), 6p21 (2.9)	Skin, CLL, Breast	0.1
Dusp23	1	172.6 / 80.1	Sluc5a (D1Mit36)	ML	3.4 / 3.4	1	159.7	1q22 (156.2)	Testicle, Gastric	3.5
Dusp26	8	31.1 / 17.9	8p12 (NRG1)	HL	0.7 / 0.9	8	33.6	8p12 (29.6-32.6)	Breast, Thyroid	1
Dusp27	1	166.1 / 73.5	Sluc5a (D1Mit36)	ML	3.1 / 3.2	1	167.1	1q24.1 (165.9)	Testicle	1.2
Dusp28	1	92.9 / 46.6	NA	NA	NA	2	240.5	2q37.3 (241.4)	Prostate	0.9
Epm2a	10	11.3 / 3.7	Sluc29 (D10Mit28)	ML	1.9 / 0.7	6	145.6	6q25.1 (149.3-151.6)	Breast	3.7
Ptpmt1	2	90.9 / 50.4	Scc1 (Ptprj)	MC	0.5 / 0.2	11	47.6	11q11.2 (56.4)	ALL	8.8
Rngtt	4	33.3 / 14.8	Sluc18 (D4Mit4)	ML	6 / 5.3	6	89.3	6q14.1 (81.5)	Breast	7.8
Styx	14	45.3 / 22.9	14q22.2 (BMP4)	HC	1.0 / 1.1	14	53.2	14q22.1 (50.9-52.9)	AML, Prostate	0.3
Slingshot Protein Phosphatase										
Ssh1	5	113.9 / 55.8	22q12.1 (CHEK2)	HL	3.1 / 2.1	12	109.2	12q23.3 (106.4), 12q24.12 (111.4), 12q24.21(114.7)	Glioma, Esphagus, Prostate, Breast	2.2
Ssh2	11	77.2 / 46.4	17p13.3 (NXN)	HC	1.0 / 0.5	17	28	17q11.1 (27.2)	CLL	0.8
Ssh3	19	4.3 / 4.0	11q12.2 (FEN1)	HC	5.8 / 2.5	11	67.1	11q12.2(61.8), 11q13.1 (65.8), 11q13.3 (69.2-69.6)	Breast, Prostate, Melanoma	1.3
Protein Tyrosine Phosphatase										
Ptp4a1	1	30.9 / 11.5	Sluc15 (D1Mit170)	ML	5.1 / 3.7	6	64.2	NA	NA	NA
Ptp4a2	4	129.8 / 63.4	Sluc21 (D4Mit70)	ML	5.1 / 3.6	1	32.4	1p34.3 (37.6-38)	Ovarian	5.2
Ptp4a3	15	73.7 / 34.1	NA	NA	NA	8	142.4	8q24.3 (142.7)	Bladder	0.3
Cdc14 Phosphatase										
Cdc14a	3	116.3 / 50.2	NA	NA	NA	1	100.8	NA	NA	NA
Cdc14b	13	64.2 / 33.3	rCcr7 (D17Rat53)	RC	1.3 / 0.6	9	99.3	9q22.33 (97.8-99.0)	Thyroid	0.3
Cdkn3	14	46.8 / 24.3	14q22.2 (BMP4)	HC	0.5 / 0.3	14	54.9	14q22.1 (52.9), 14q23.1 (58.6)	ALL, Prostate	1

Ptpdc1	13	48.6 / 24.9	9q22.2-31.2 (FGD3)	HC	0.6 / 0.3	9	96.8	9q22.2 (93.0), 9q22.33 (97.8-99.0)	Thyroid, Ovarian	1
Pten Protein Phosphatase										
Pten	19	32.8 / 28.1	10q11.23 (A1CF)	HC	1.0 / 1.5	10	89.6	10q23.31 (89.0)	CLL	0.6
Tns1	1	73.9 / 38.1	NA	NA	NA	2	218.7	2q35 (217.0-217.4)	Breast, Thyroid	1.3
Tns2	15	102.1 / 57.3	Sluc34 (D15Mit16)	ML	0.8 / 0.8	4	79.3	4q13.3 (73.0-73.6)	Prostate	5.7
Tpte	8	22.3 / 11.2	Sluc20 (D8Mit3)	ML	2.3 / 1.7	21	10.9	21q21.1 (15.1)	Breast	4.2
Tpte2	5	109.6 / 53.2	22q12.1 (CHEK2)	HL	1.2 / 0.5	13	20.1	13q12.2 (27.9)	Pancrea	6.9

a. Colon or lung cancer susceptibility genes co-localizing with Dusp genes. All cancer susceptibility loci were identified in genome-wide linkage or association studies and projected to the mouse genome.

b. Type of susceptibility loci co-localizing with Dusp genes: HC, human colon cancer; HL, human lung cancer; ML, mouse lung cancer; RC, rat colon cancer.

c. Co-localization between DUSP genes and cancer susceptibility genes, other than those for colon and lung cancers, identified in human GWAS.

d. Dusp genes that showed co-localization in both **a** and **c** were highlighted in bold.