

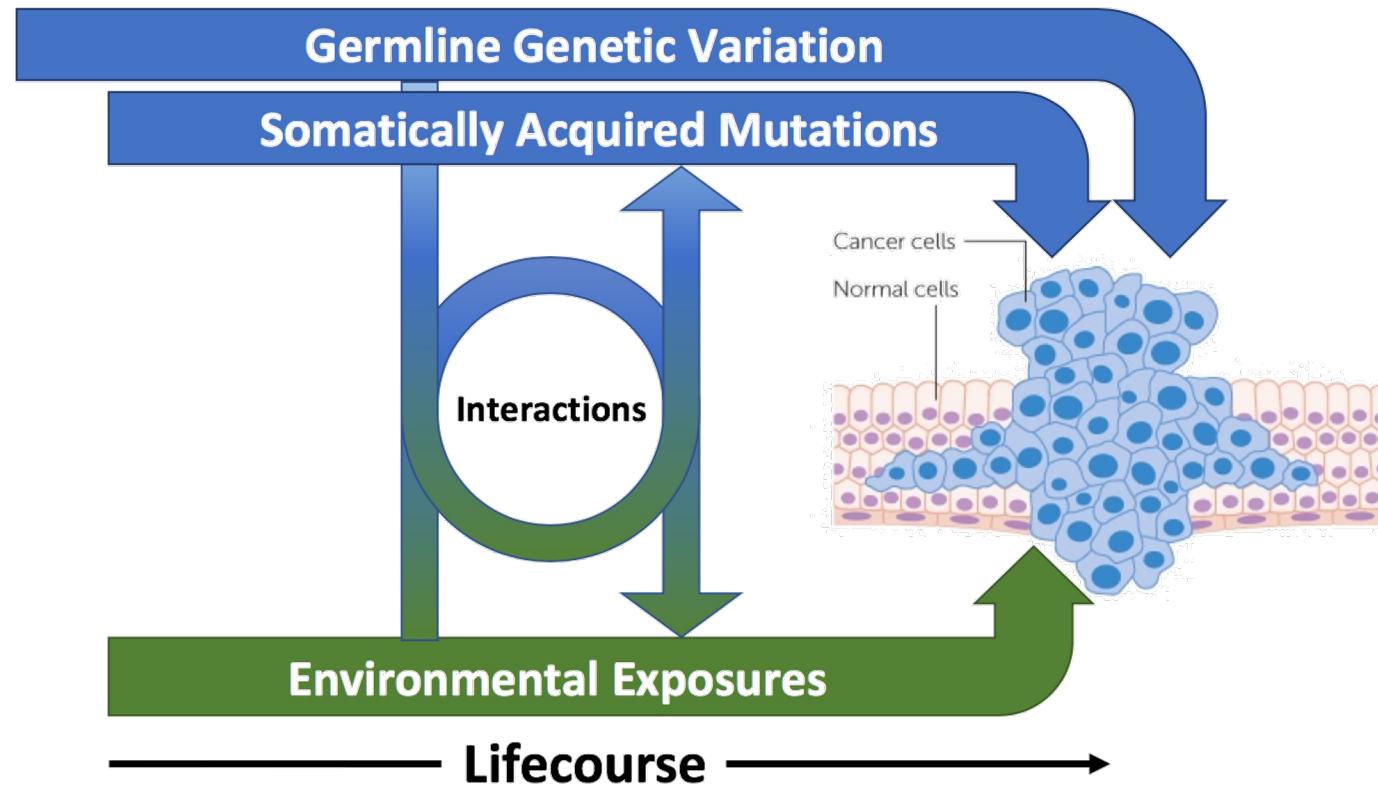
Genetic Mosaicism and Clonal Hematopoiesis

*Mitchell J. Machiela, ScD MPH
Earl Stadtman Investigator*

*Weiyin Zhou, MS
Bioinformatics Analyst IV*

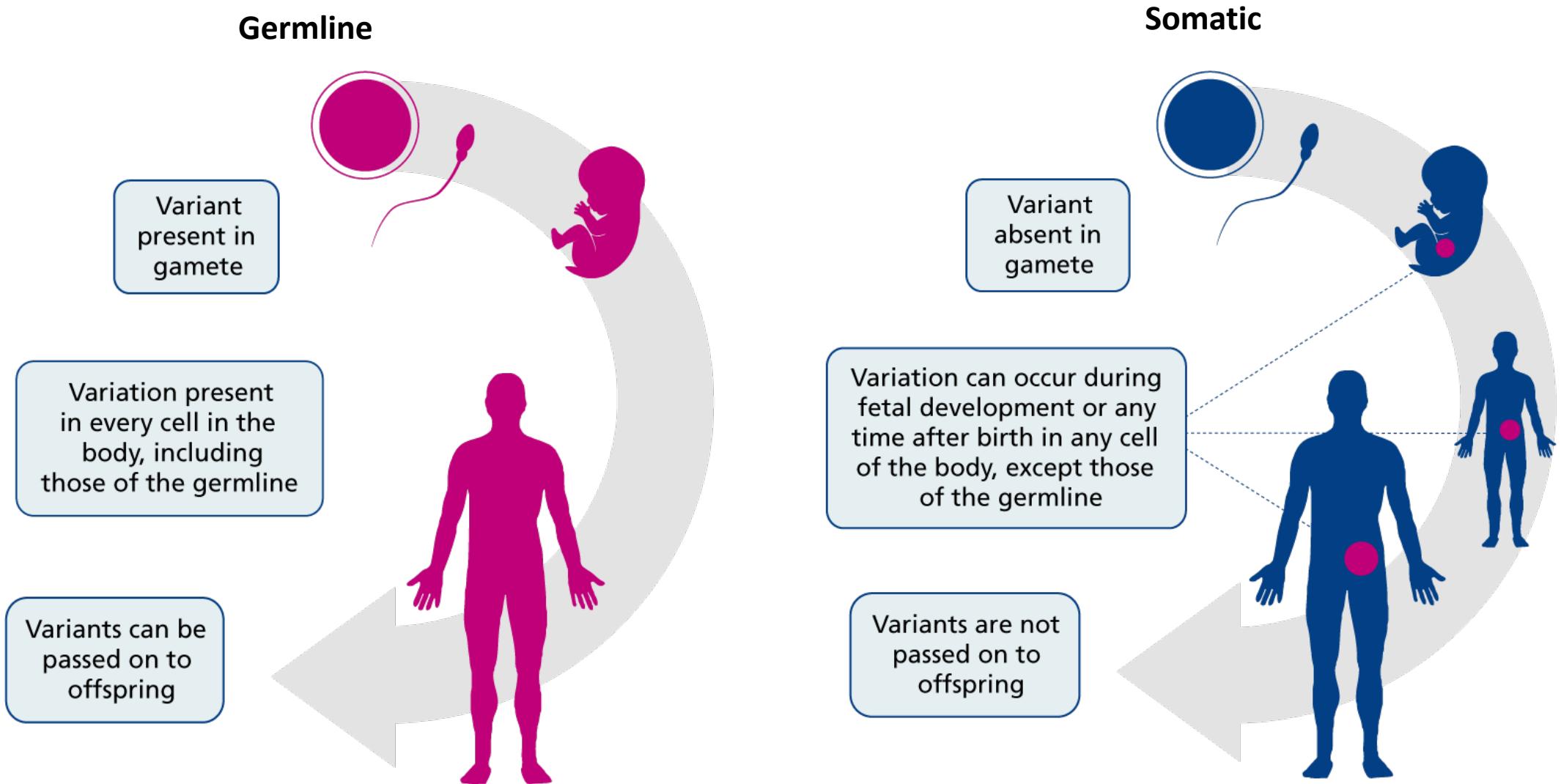
Cancer is Complex

An integrative approach is essential to disentangle genetic etiology



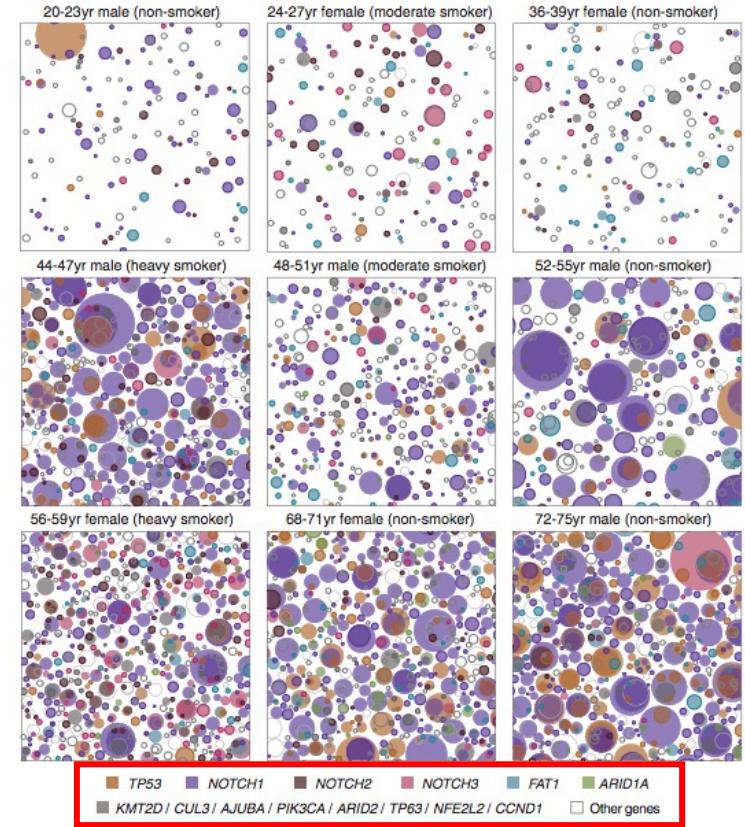
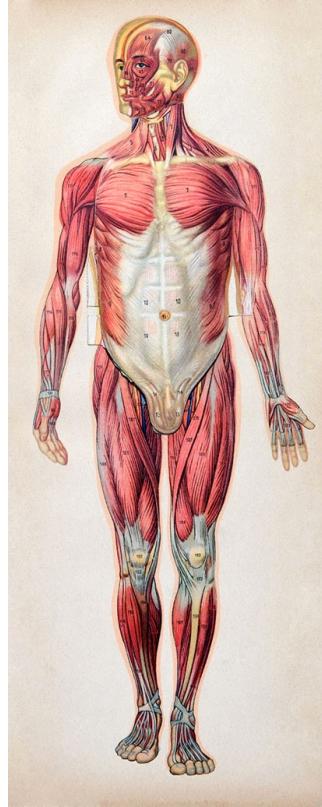
Studies of early molecular changes in normal tissues (before tumor formation) have tremendous potential to improve understanding of cancer etiology

Genetic Variation



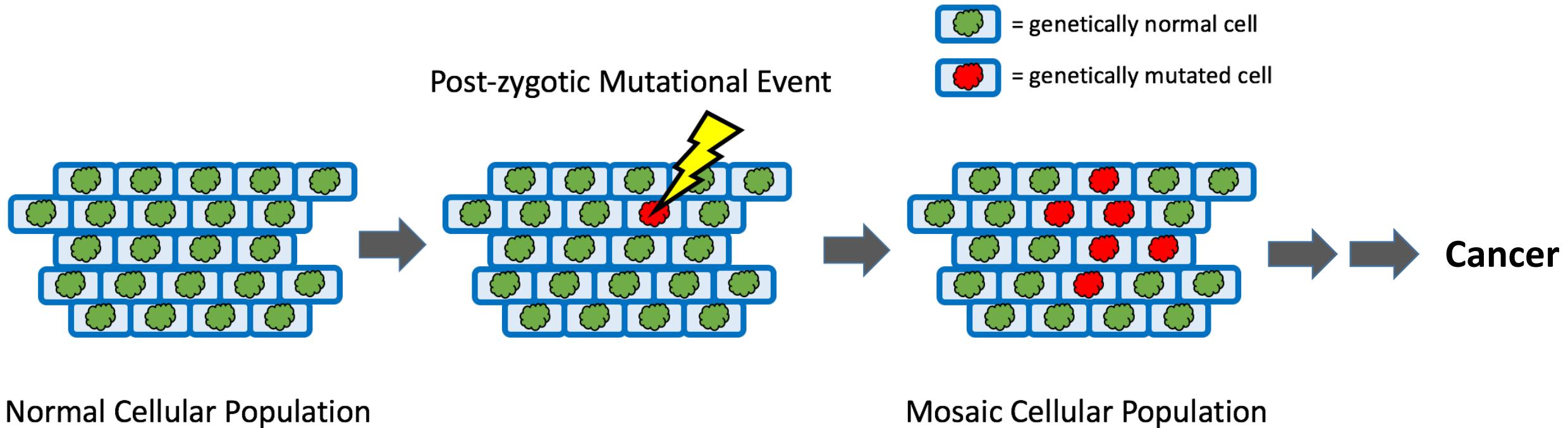
Normal Tissues Harbor Multitudes of Mutations

- Improved technologies → improved detection
- Age-related mutations are abundant in normal tissues
- All precursors or clones do not progress to cancer (even clones with mutations in cancer predisposition genes)
 - additional mutations/selective forces required for transformation
- Blood ideal tissue to study
 - easily accessible tissue
 - large existing well-characterized collections with genomic data
 - most mosaicism studies in blood DNA



Clonal Mosaicism

- The presence of an acquired somatic mutation(s) in a clonal subset of cells that differs from the inherited germline genome



Clonal Mosaicism

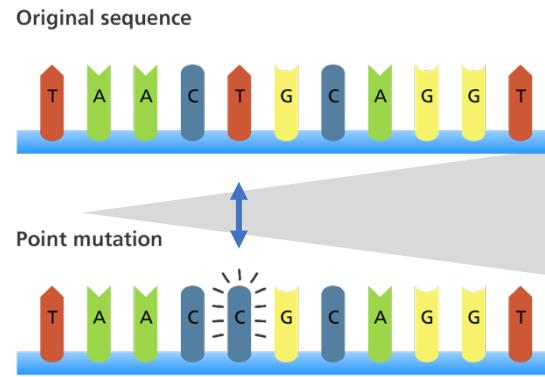
- The presence of an acquired somatic mutation(s) in a clonal subset of cells that differs from the inherited germline genome
 - ***Detectable*** clonal mosaicism requires two fundamental features:
 1. the acquisition of a somatic mutation that goes unrepaired
 - balance between mutational processes and repair processes
 2. survival and subsequent clonal expansion transmitting acquired mutation on to daughter cells

Clonal Mosaicism

- The presence of an acquired somatic mutation(s) in a clonal subset of cells that differs from the inherited germline genome
 - autosomes (1-22), sex chromosomes (female XX and male XY)
 - studies primarily focus on circulating leukocytes (small fraction of buccal)
 - not chimerism
 - the presence of genomic material from two or more individuals within a single individual
 - ex: circulating fetal cells in the mother, transplant

Size Spectrum of Mosaic Events

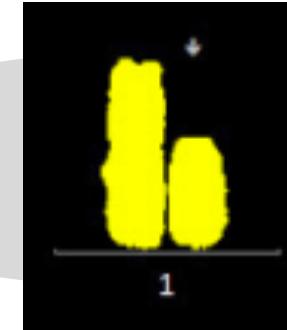
Mosaic Single Nucleotide Variants (SNVs)



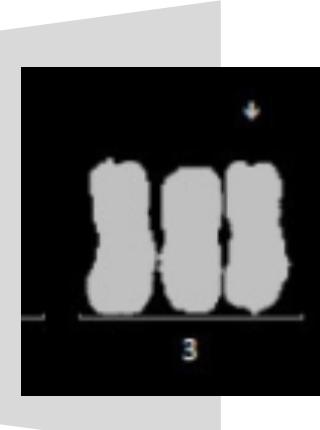
Next generation sequencing

Mosaic Chromosomal Alterations (mCAs)

Partial Chromosomal Events

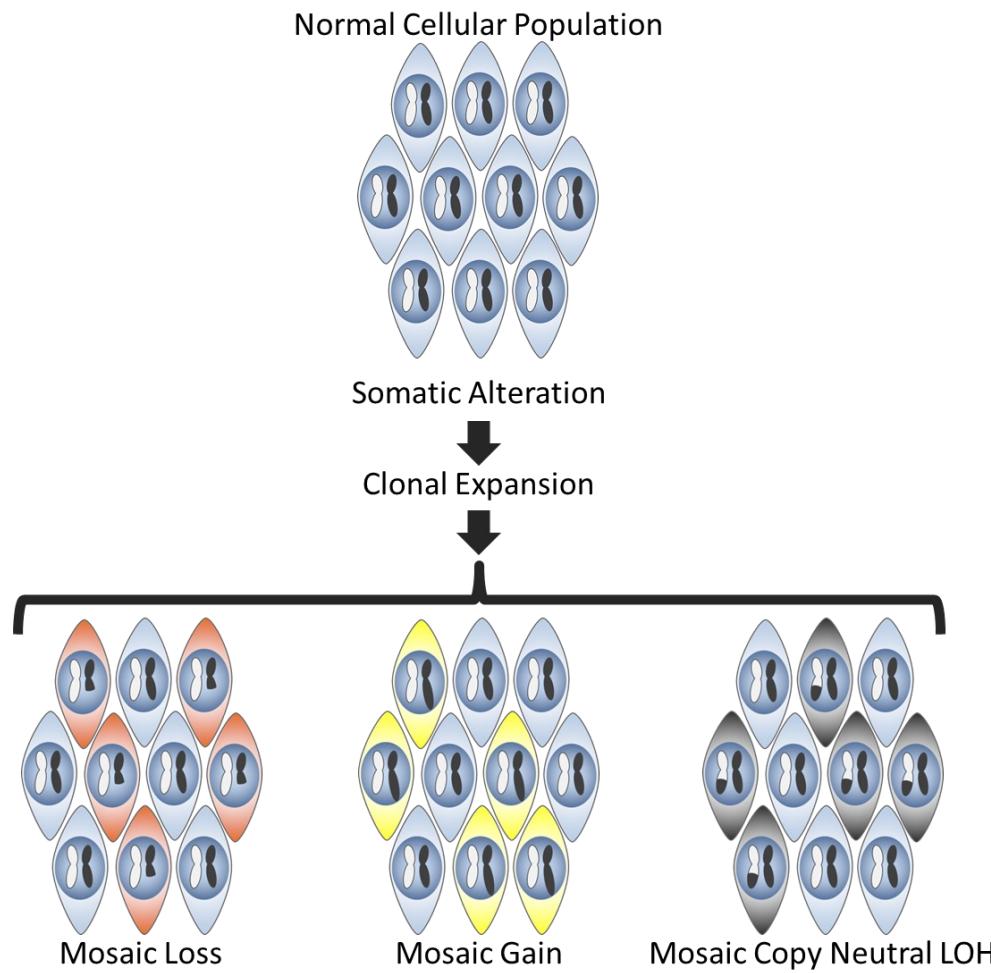


Whole Chromosome



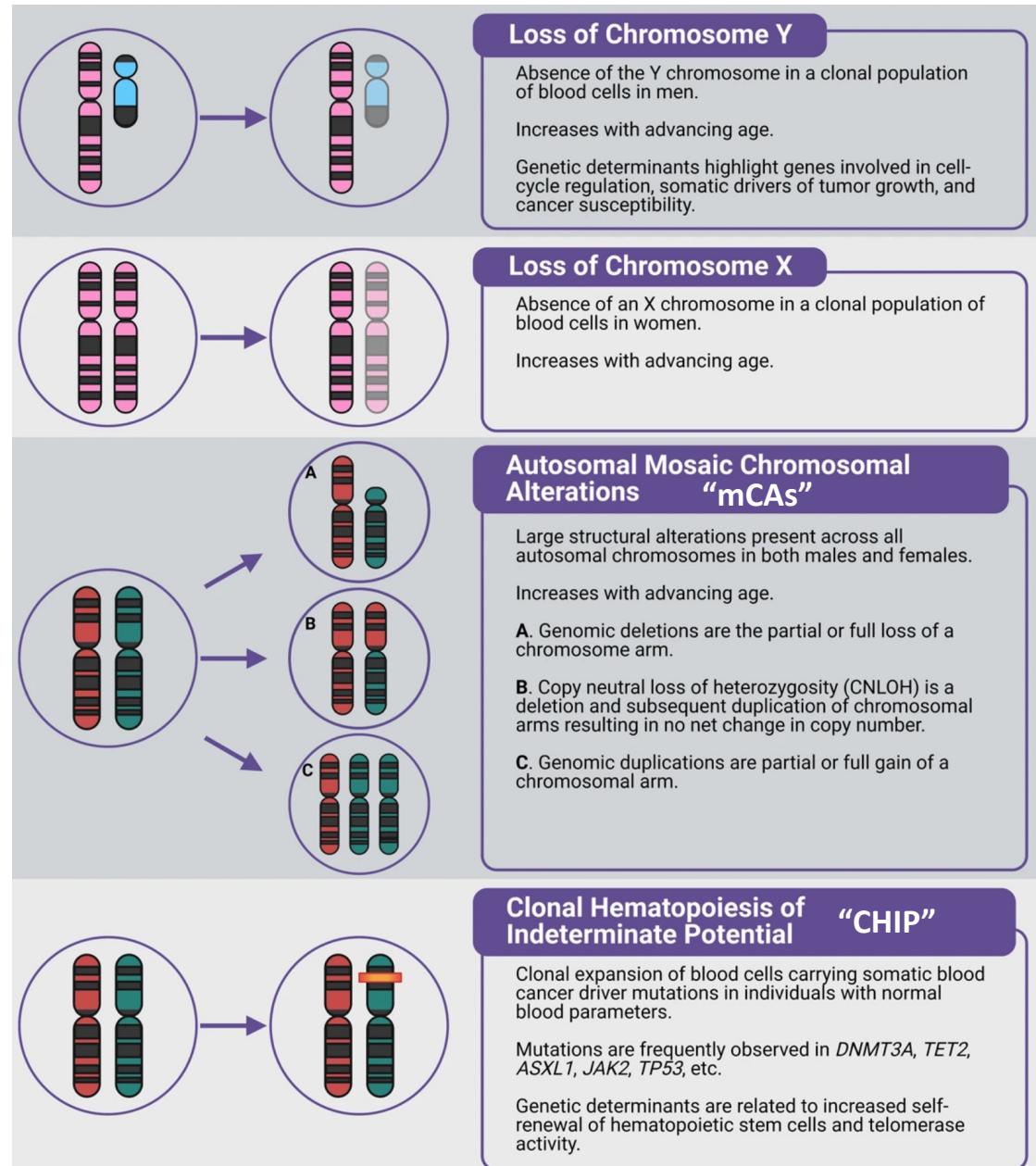
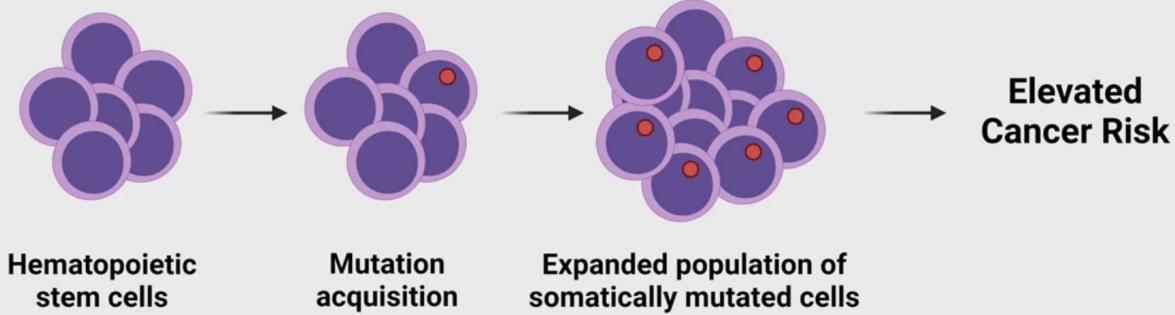
Commercially available genotyping arrays

Types of Mosaic Chromosomal Alterations

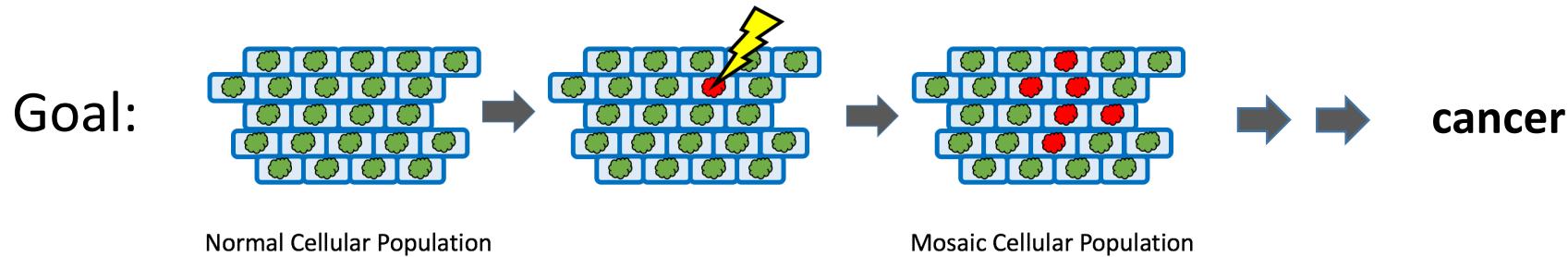


Clonal Hematopoiesis = Clonal Mosaicism in Blood

TYPES OF CLONAL HEMATOPOIESIS

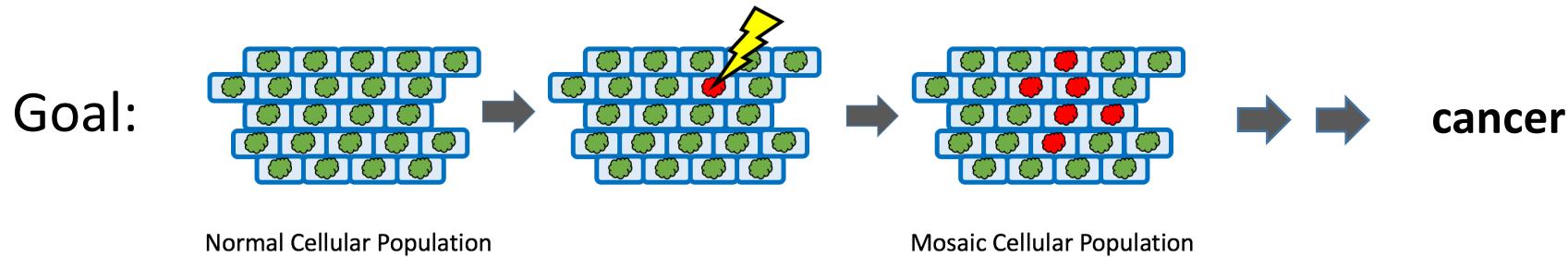


Outline



1. How are mosaic chromosomal alterations detected?
2. What risk factors are associated with the development and clonal expansion of mosaic chromosomal alterations?
3. Do mosaic chromosomal alterations alter disease risk (e.g., cancer)?

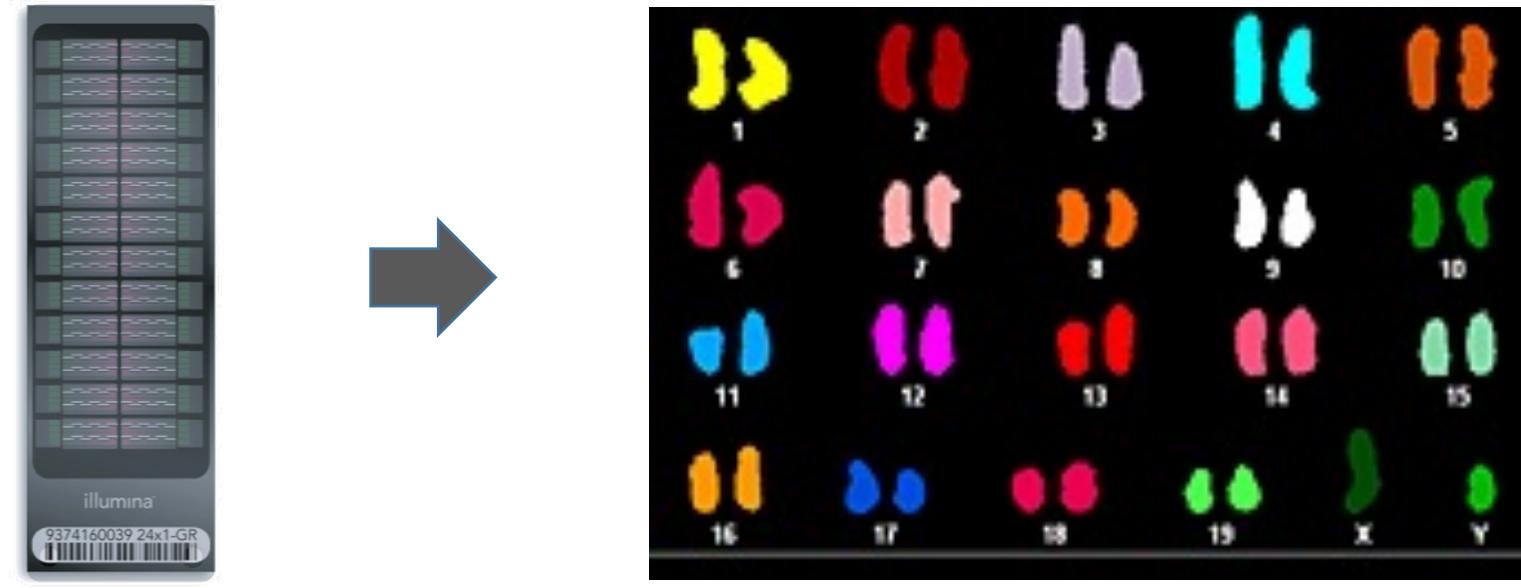
Outline



1. How are mosaic chromosomal alterations detected?
2. What risk factors are associated with the development and clonal expansion of mosaic chromosomal alterations?
3. Do mosaic chromosomal alterations alter disease risk (e.g., cancer)?

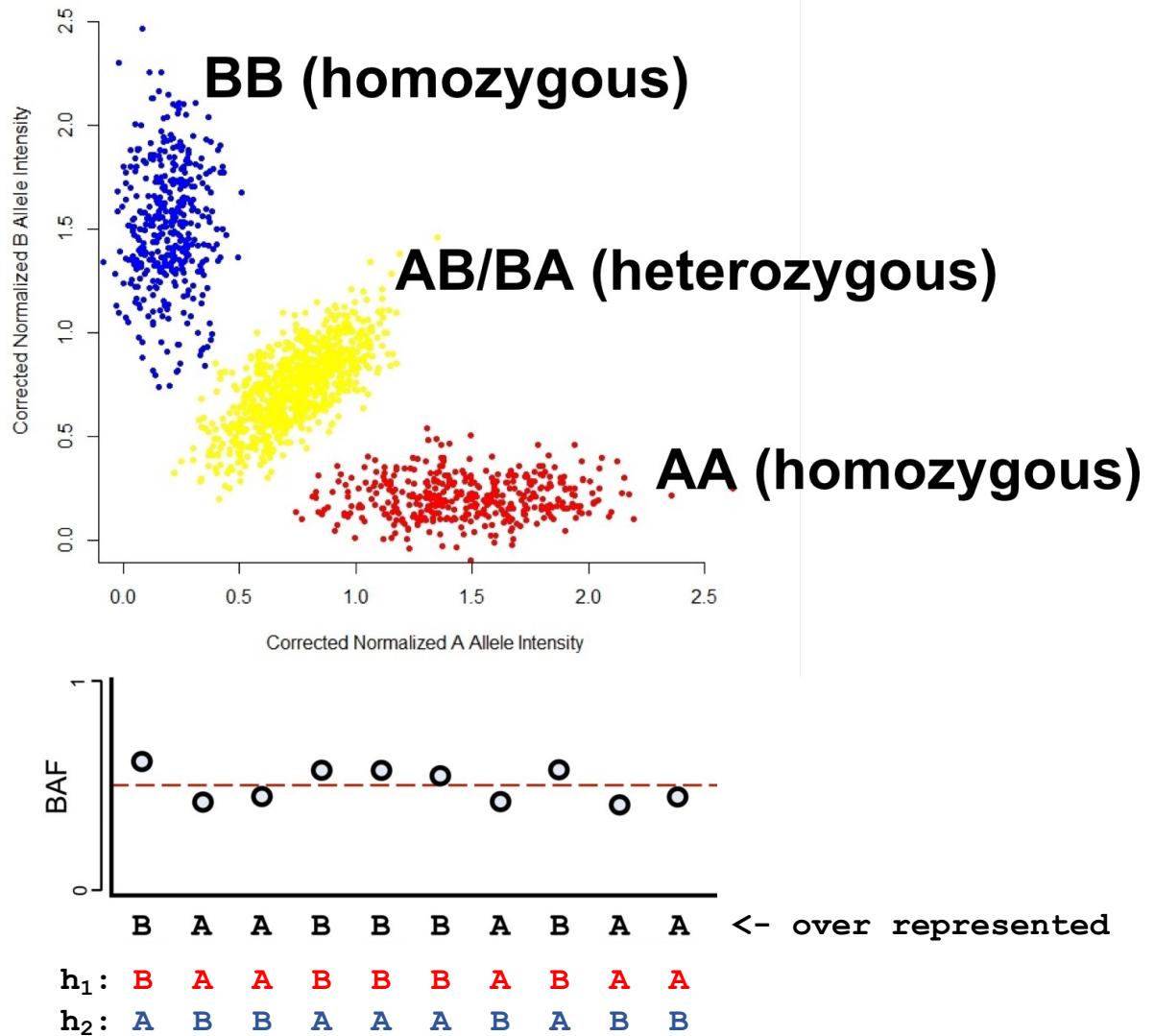
Virtual Karyotypes from Existing Genotyping Data

Create “virtual karyotypes” from existing SNP genotype array intensity data



Metrics for Detecting Mosaicism

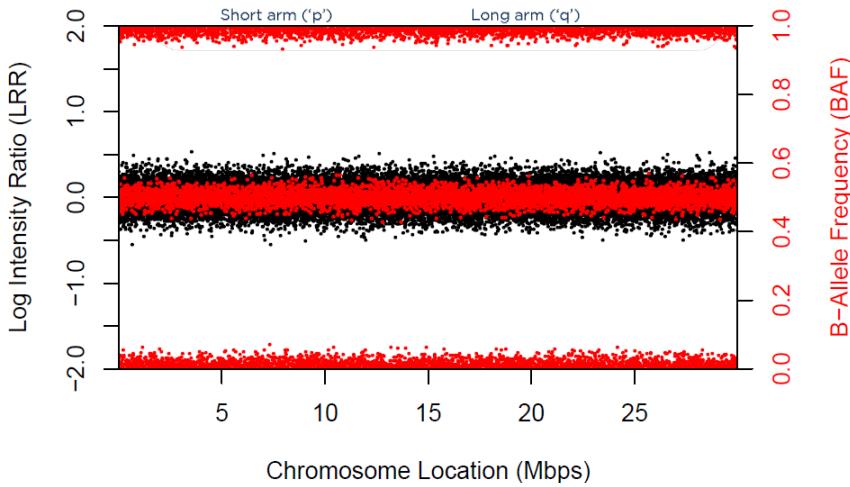
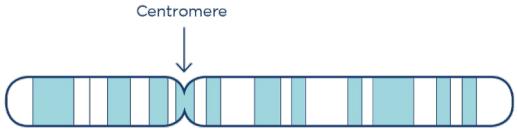
- B Allele Frequency (BAF)
 - allelic imbalances
- Log₂ R ratio (LRR)
 - abnormalities in intensity magnitude
- Haplotype phase
 - over represented haplotypes



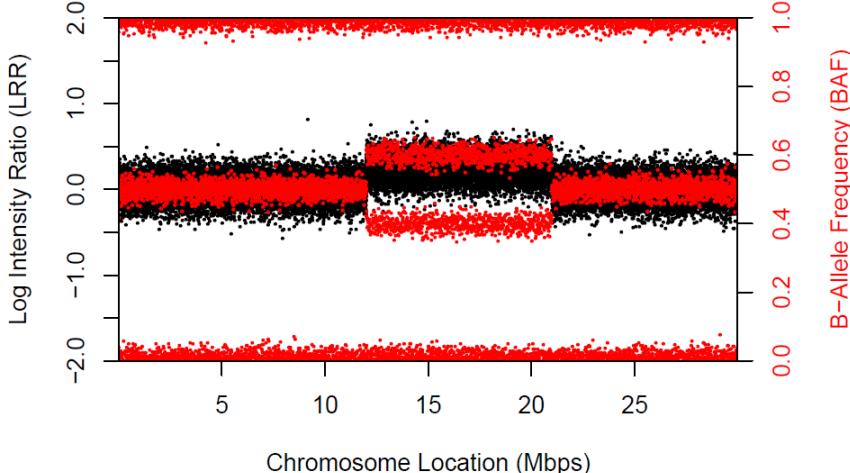
Detecting Mosaic Chromosomal Alterations (mCAs)

two points per marker

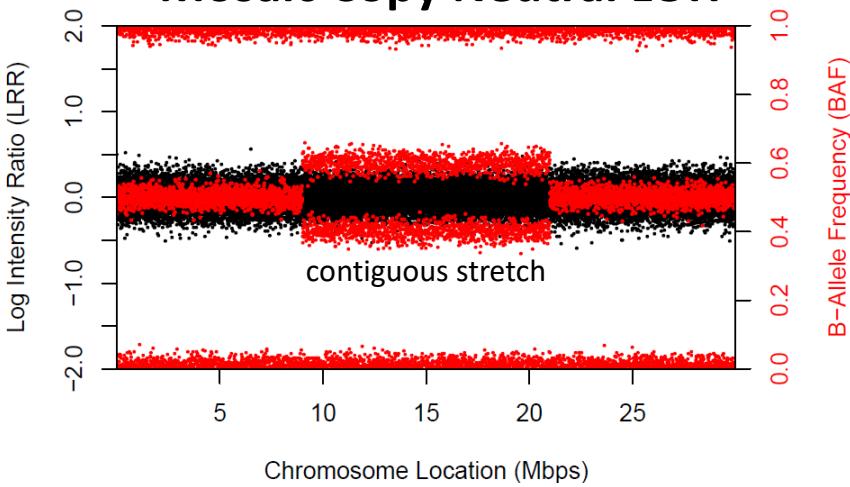
- **BAF**
- **LRR**



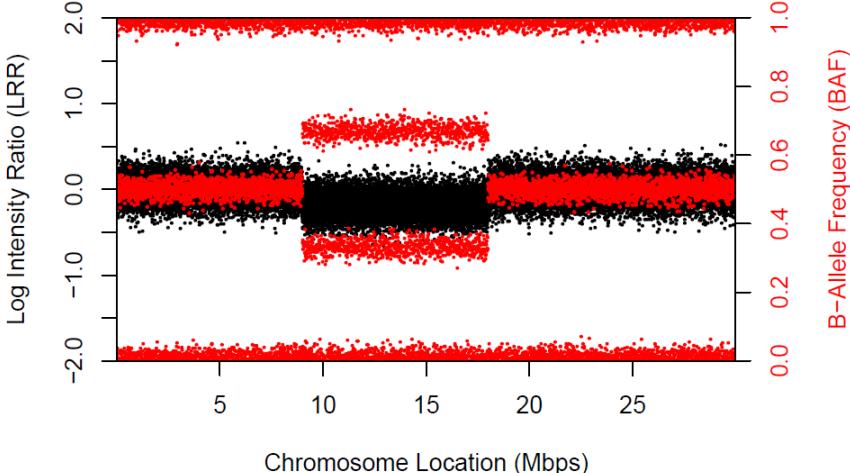
Mosaic Gain



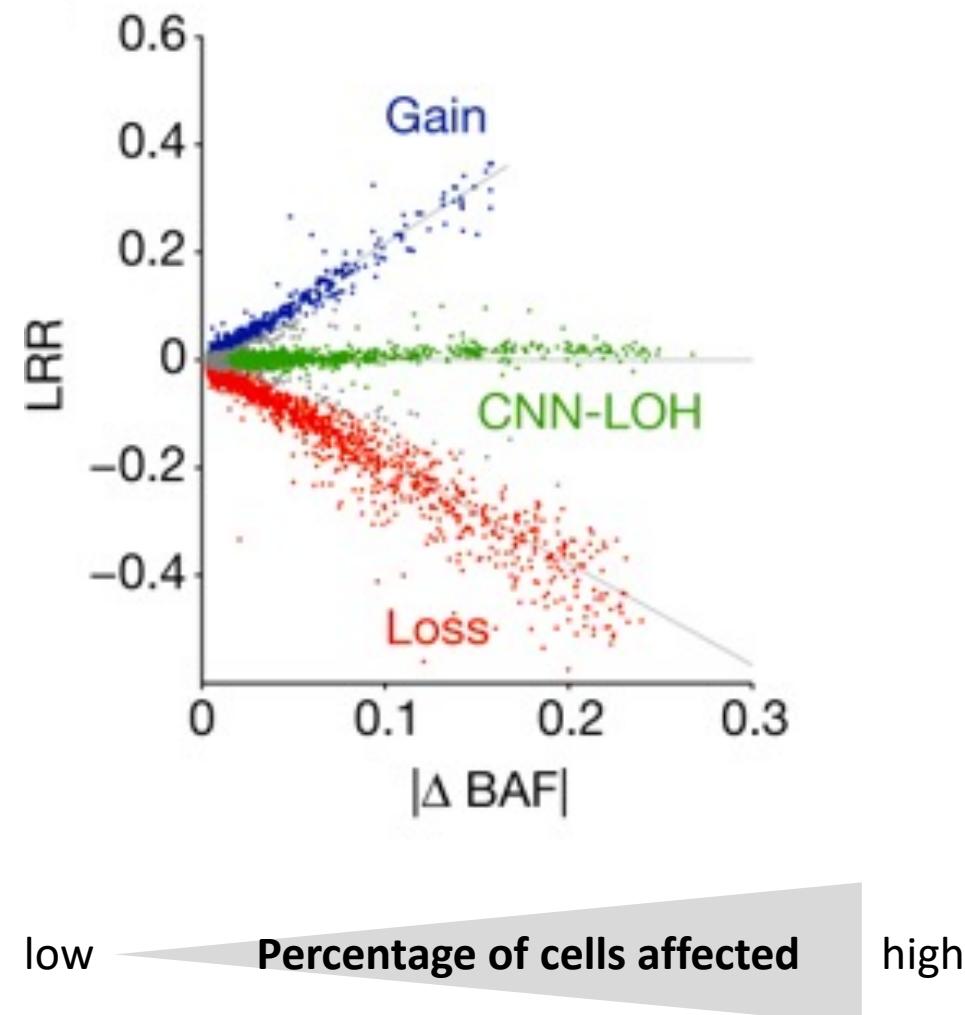
Mosaic Copy Neutral LOH



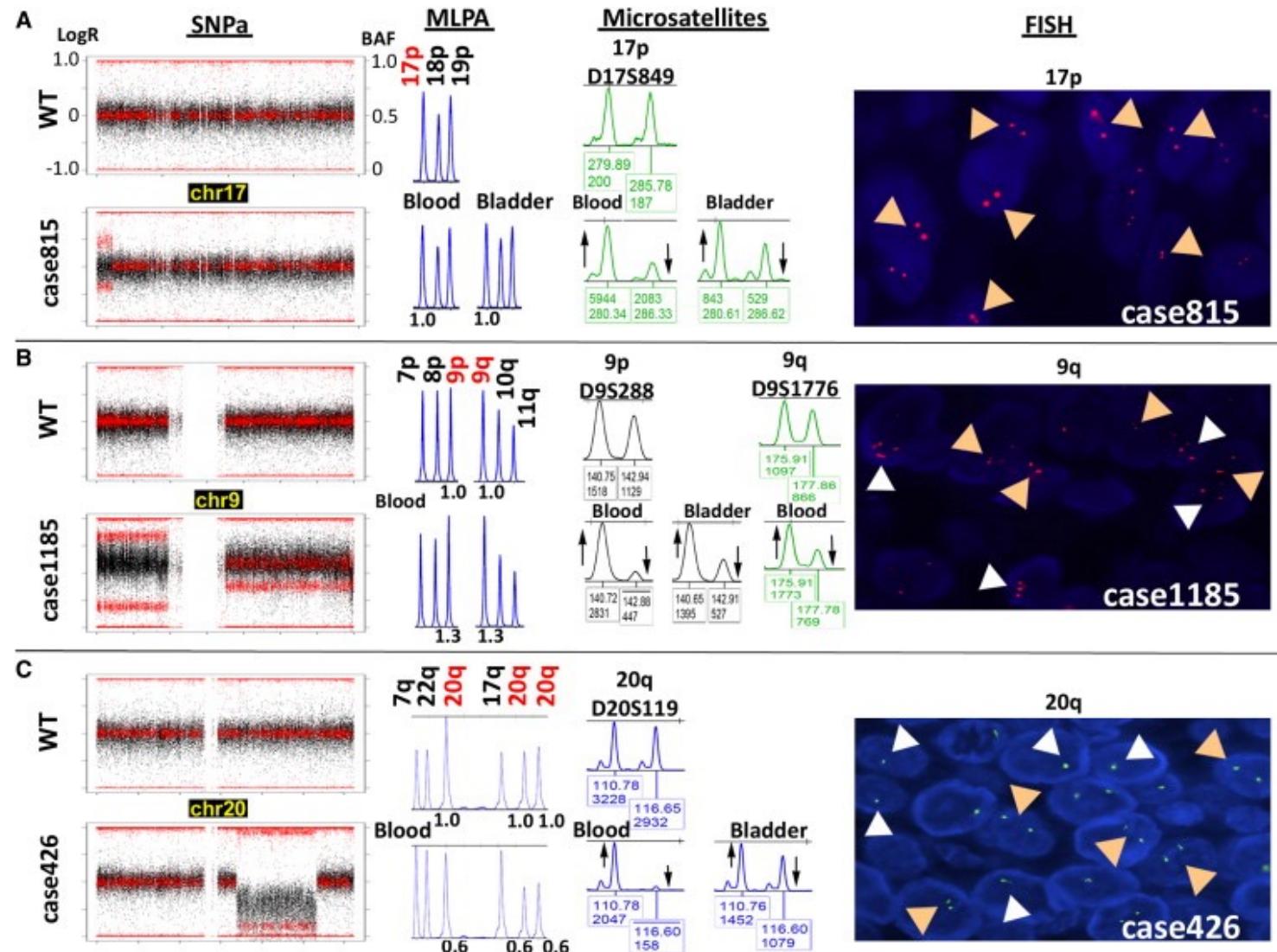
Mosaic Loss



Combining BAF and LRR from Multiple Detected mCAs



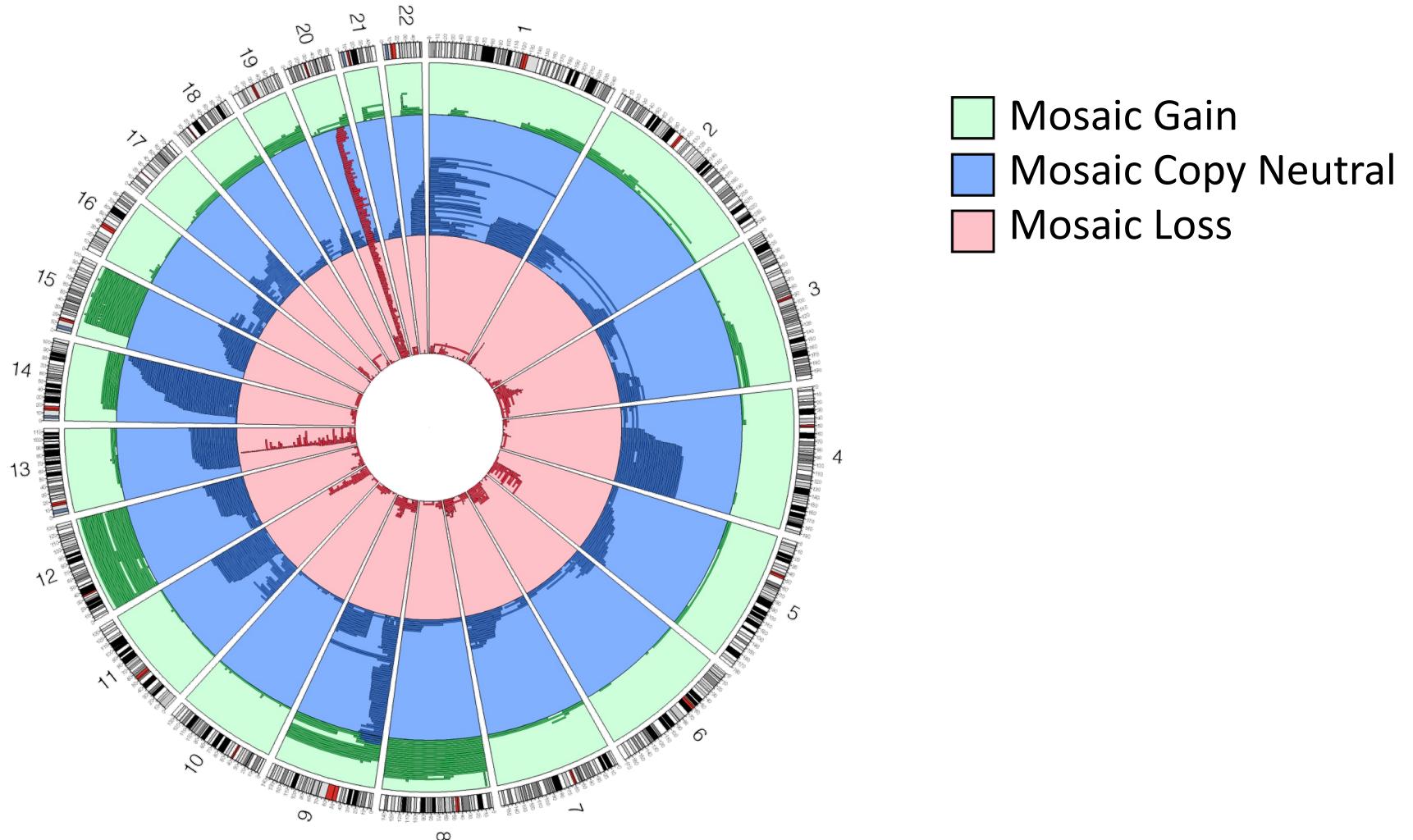
Experimental Validation



Distribution of Detectable Mosaic Chromosomal Alterations

Large survey of primarily blood derived DNA of 127,179 individuals:

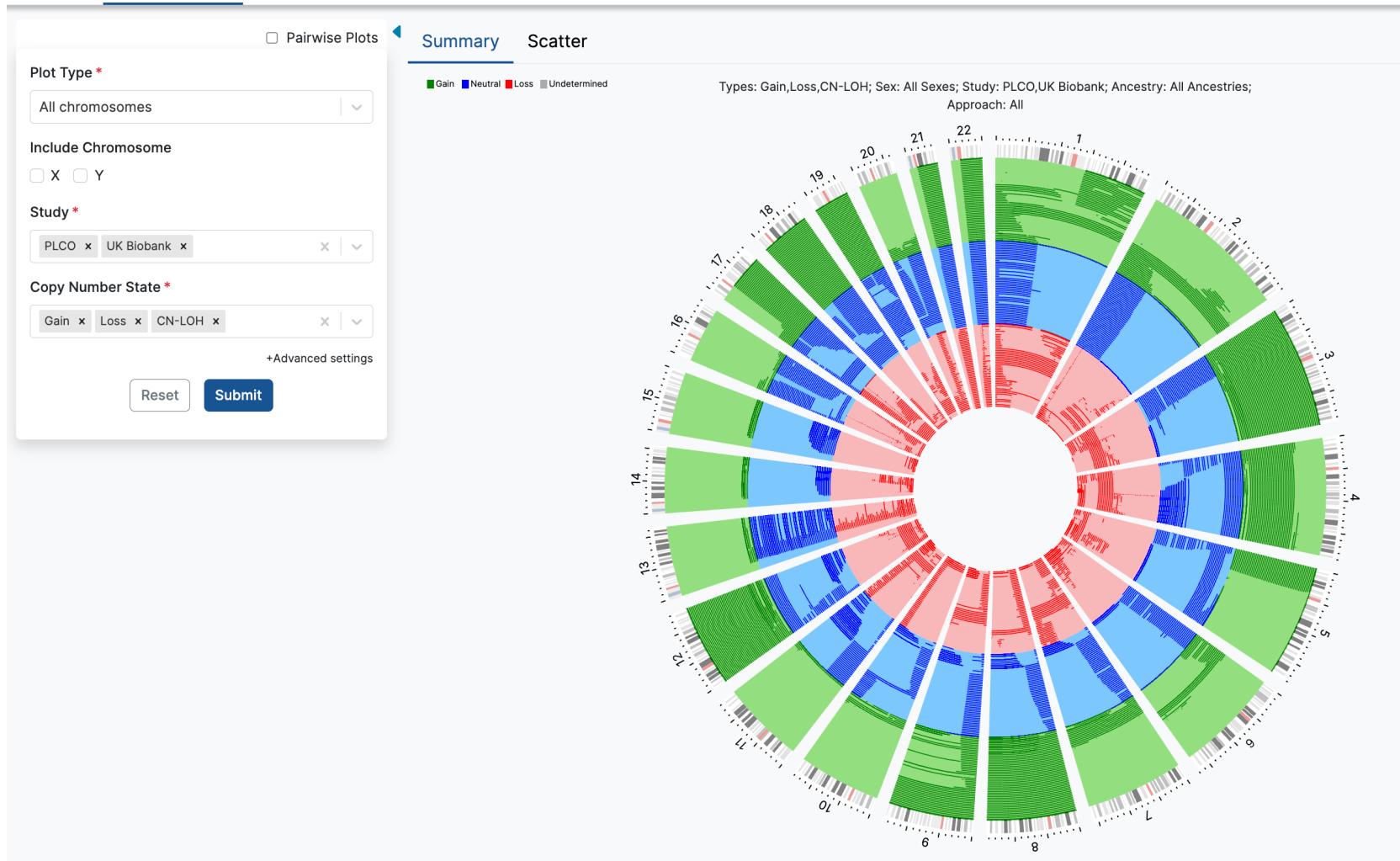
- 1,315 events >2 Mb
- 925 individuals (0.7%)



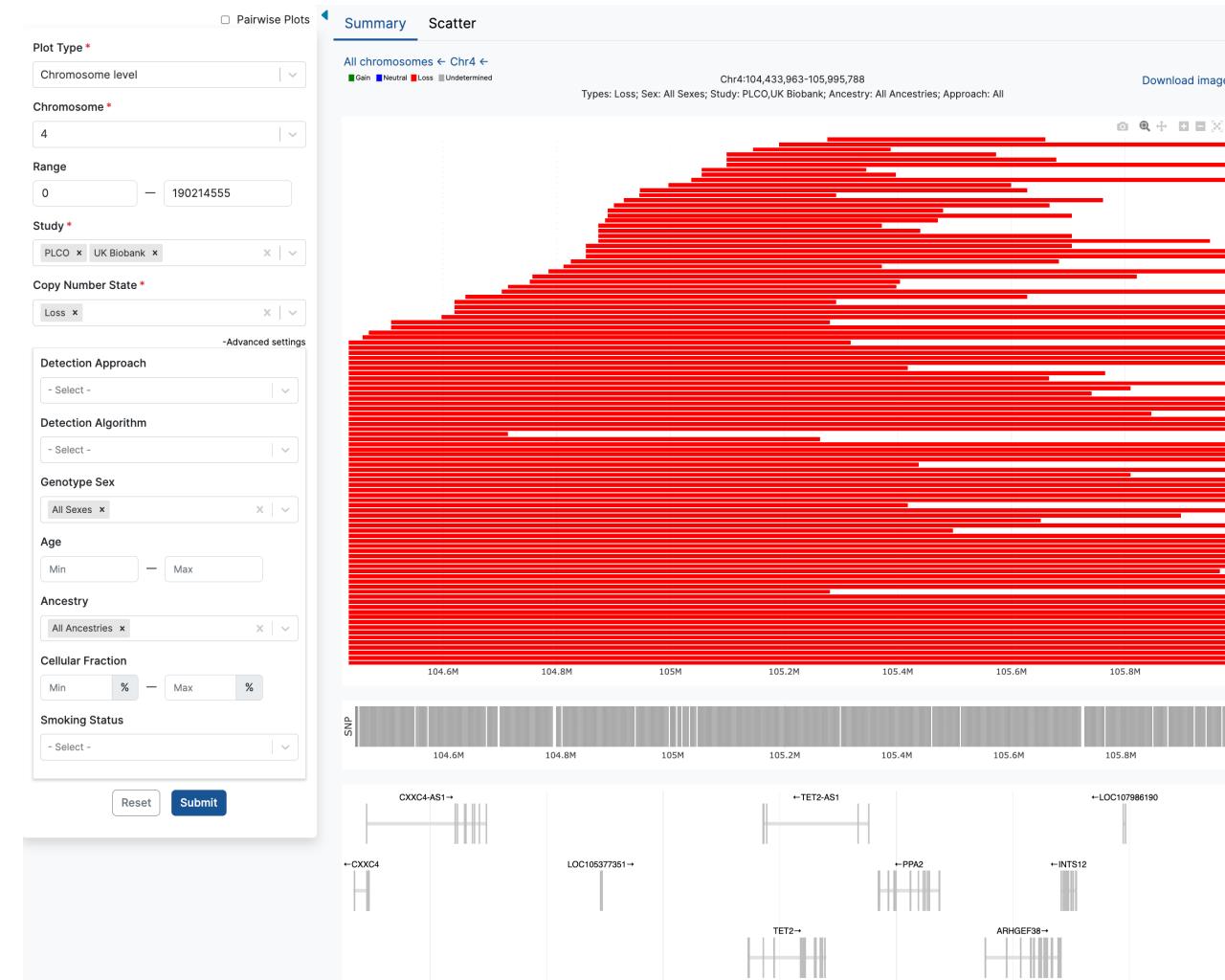
mCA Explorer – An Interactive Catalog of mCAs



Home mCA Explorer API Access About



mCA Explorer – Visualize Location of mCAs



mCA Explorer – Compare mCAs among groups

[Back to circos compare ← Chr4 ←](#)

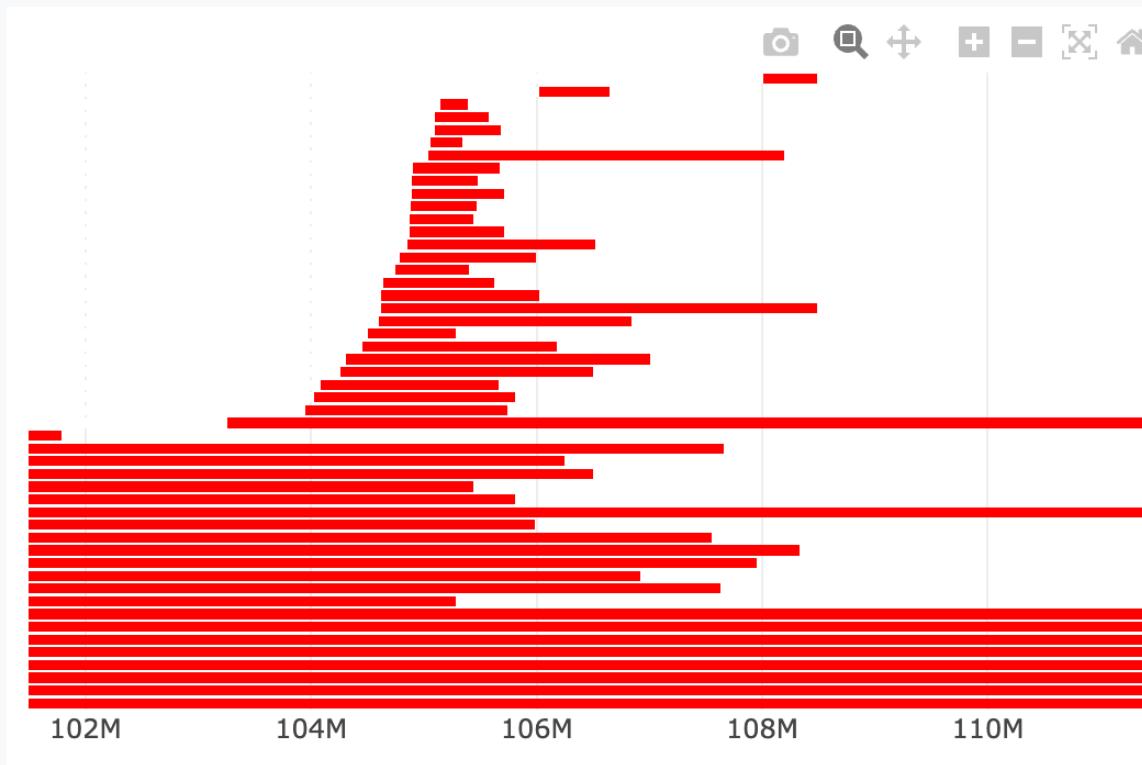
Gain Neutral Loss Undetermined

Chr4:101,495,153-111,368,222

Study: PLCOUK Biobank; Types: Loss

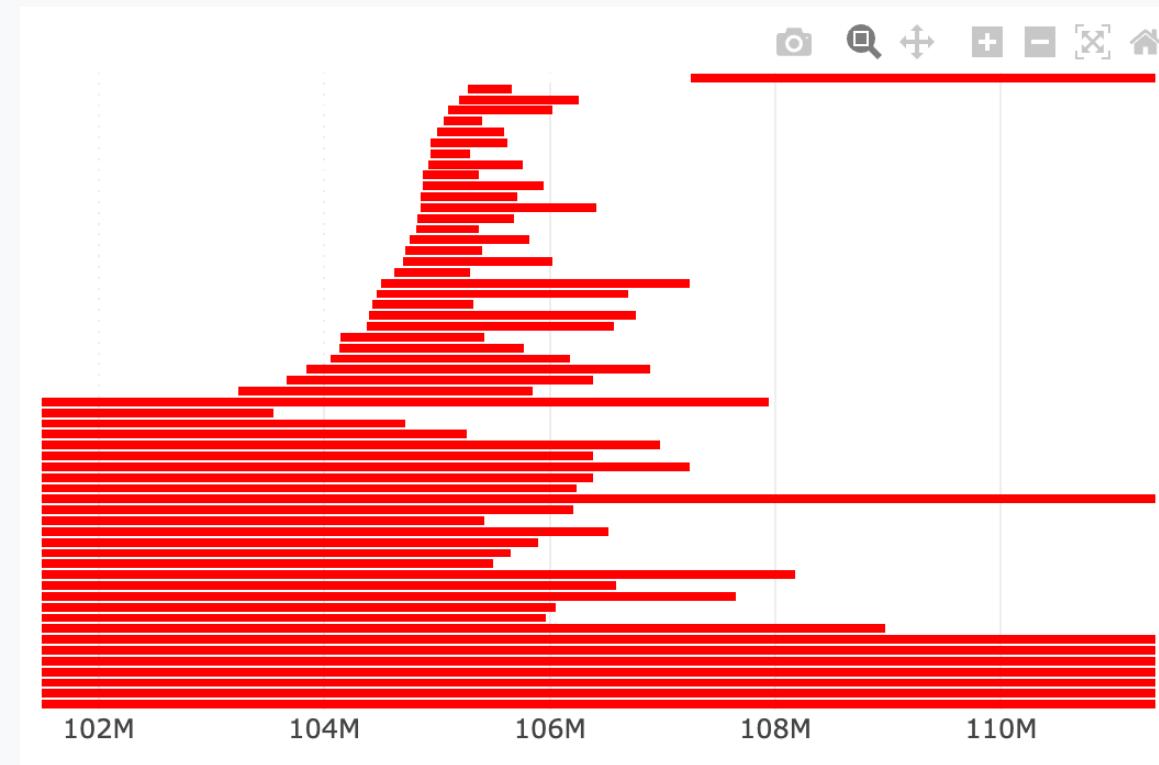
[Download comparison images](#)

Sex: Male



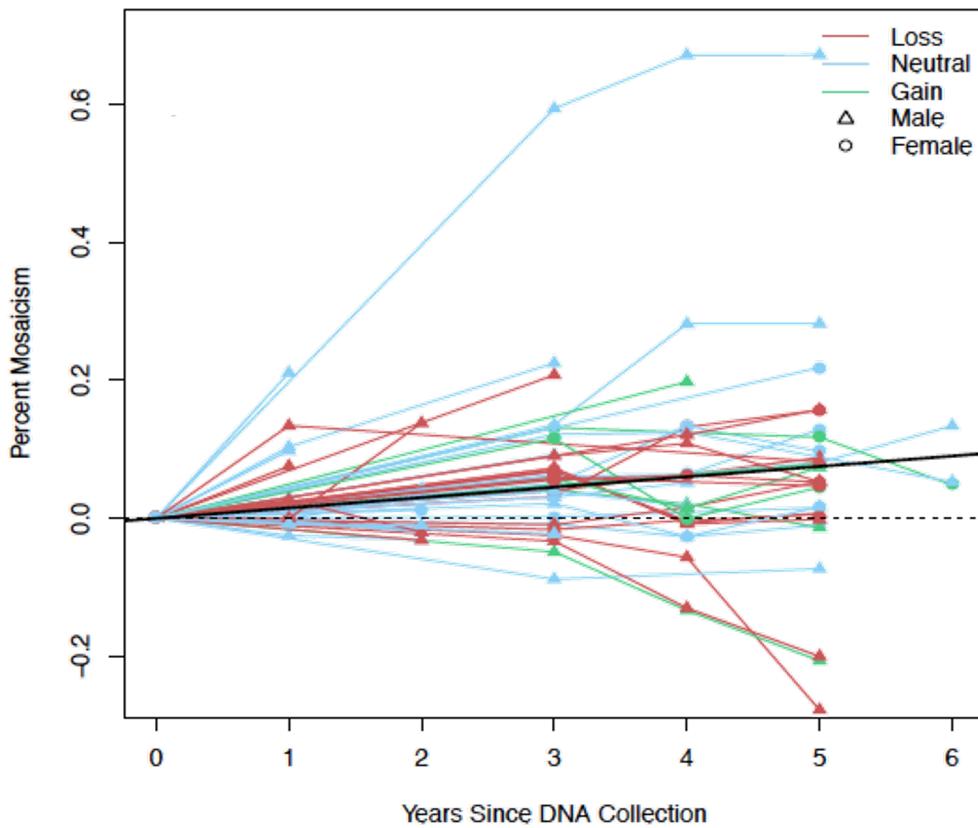
Chr4:101,495,153-111,368,222

Sex: Female



Chr4:101,495,153-111,368,222

Mosaicism is Dynamic Changes Over Time

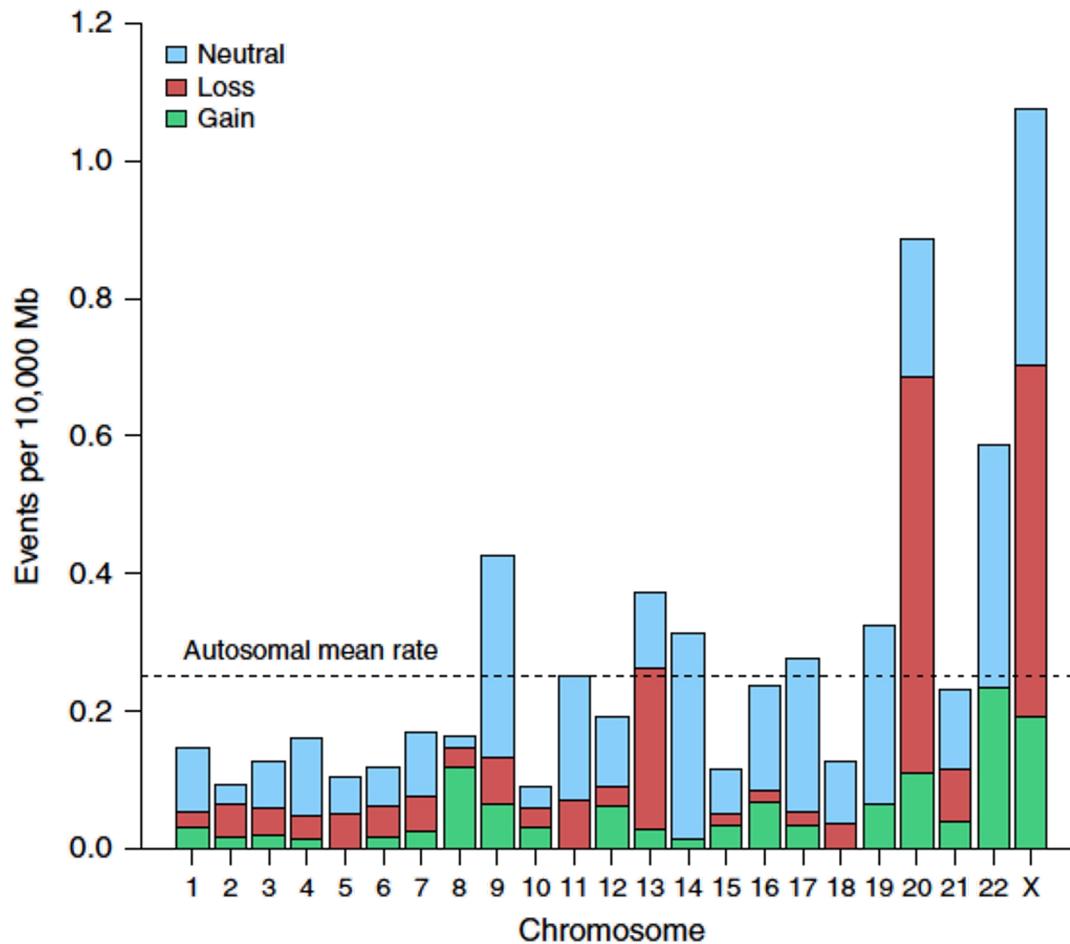


CME	Time ^a	Chr.	Start position	End position	Size of CME (kb)	CME type	Abnormal cells (%)
1	T ₀	7	138342	158779565	158,641	Neutral	76
	T ₆	7	138342	158779565	158,641	Neutral	81
2	T ₀	11	192856	37154718	36,962	Neutral	12
	T ₆	11	192856	40000000	39,807	Neutral	14
3	T ₀	13	32383476	114092980	81,710	Neutral	9
	T ₆	13	28815864	114092980	85,277	Neutral	10
4	T ₀	15	20000000	100206215	80,206	Gain	29
	T ₆	15	18842870	100206215	81,363	Gain	39
5	T ₀	17	52467	21000000	20,948	Loss	22
	T ₆	17	52467	17601072	17,549	Loss	26
6	T ₀	17	21000000	78644427	57,644	Gain	23
	T ₆	17	17621285	78644427	61,023	Gain	26
7	T ₀	20	30933238	44192573	13,259	Loss	64
	T ₆	20	28000000	44321540	16,322	Loss	72

Chr., chromosome.

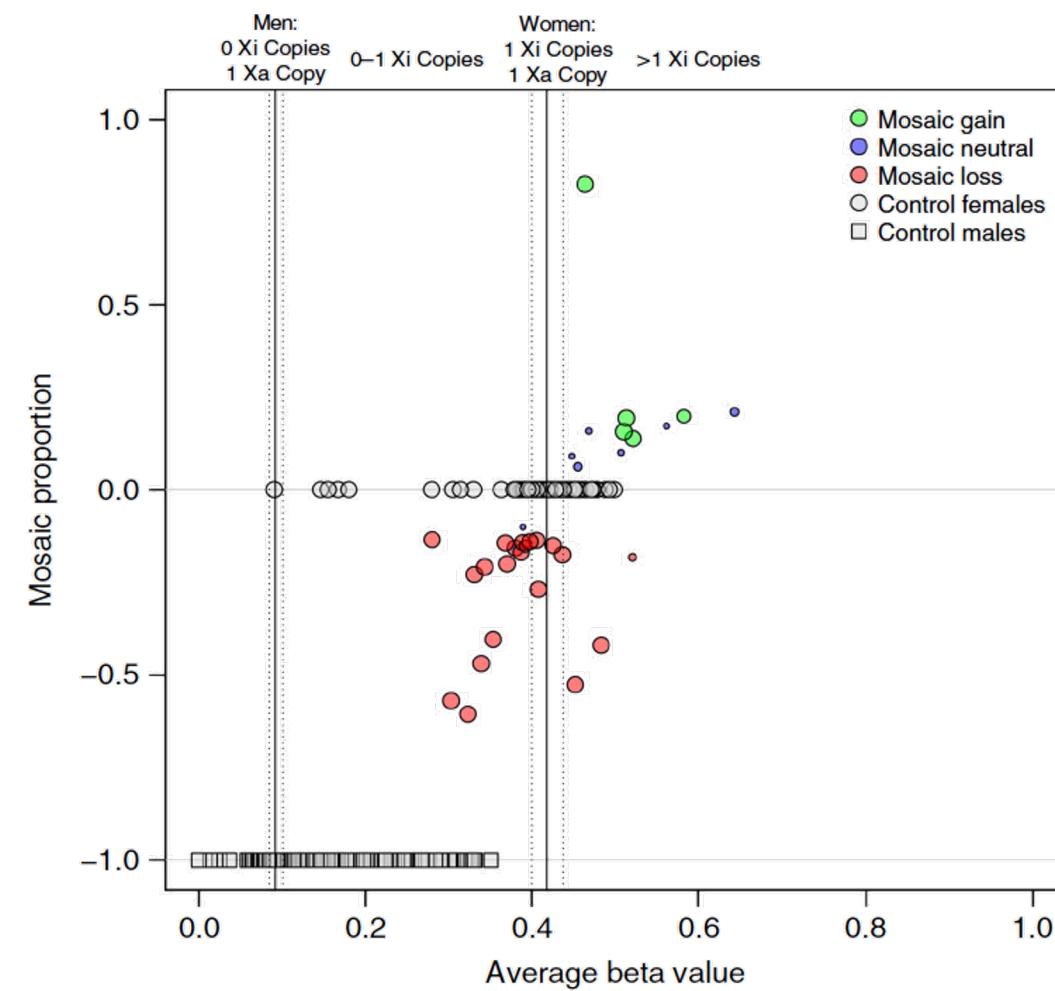
^aSamples were collected at time 0 (T₀) or 6 years later (T₆).

Elevated Rate on the Female X Chromosome

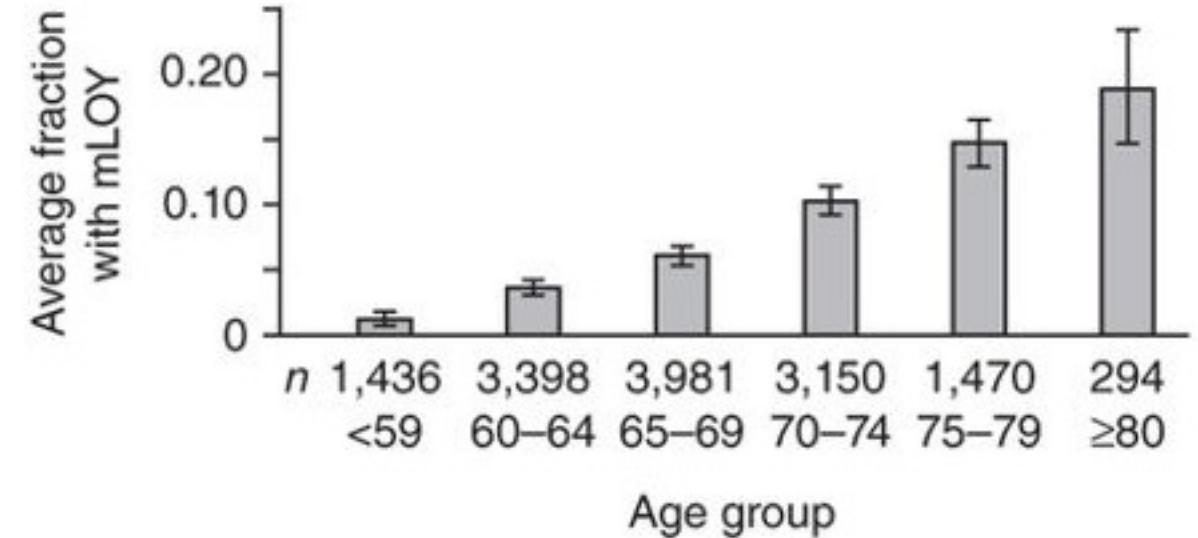
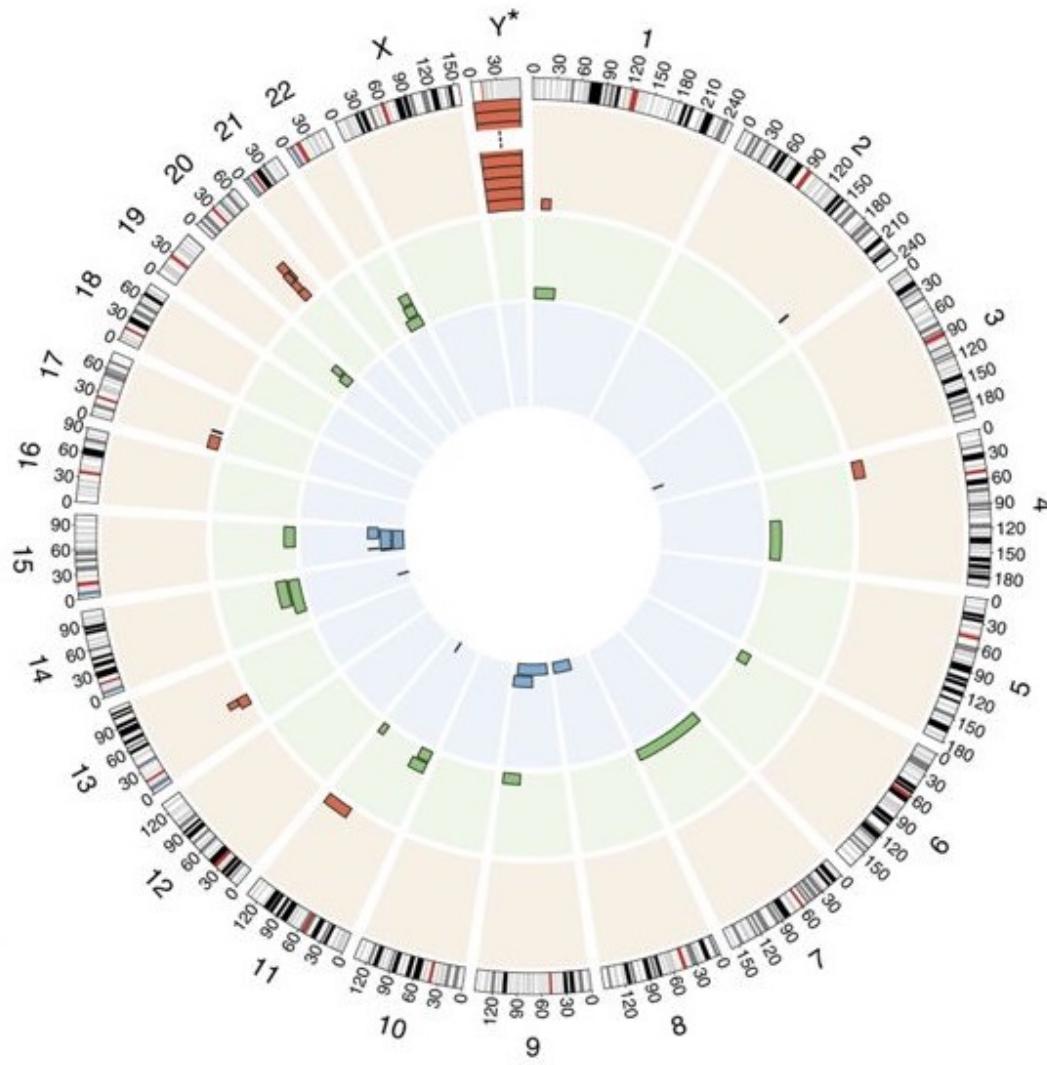


Chromosome	N_{loss}	$N_{CNN-LOH}$	N_{gain}	$N_{undetermined}$	N_{total}
chr1	29	318	17	134	498
chr2	66	56	10	48	180
chr3	18	53	41	63	175
chr4	47	64	8	41	160
chr5	49	40	24	38	151
chr6	32	68	6	64	170
chr7	70	43	5	40	158
chr8	22	35	42	44	143
chr9	19	210	38	78	345
chr10	70	29	5	31	135
chr11	98	257	1	105	461
chr12	28	67	156	95	346
chr13	177	111	0	73	361
chr14	51*	223	38	135	447
chr15	14	121	59	93	287
chr16	43	142	2	53	240
chr17	66	112	37	89	304
chr18	14	20	57	40	131
chr19	6	90	17	75	188
chr20	140	55	3	29	227
chr21	20	35	31	67	153
chr22	39*	88	62	113	302
All autosomes	1118	2237	659	1548	5562
Female chrX	1862	28	24	866	2780

Whole Chromosome Losses/Gains of Female X



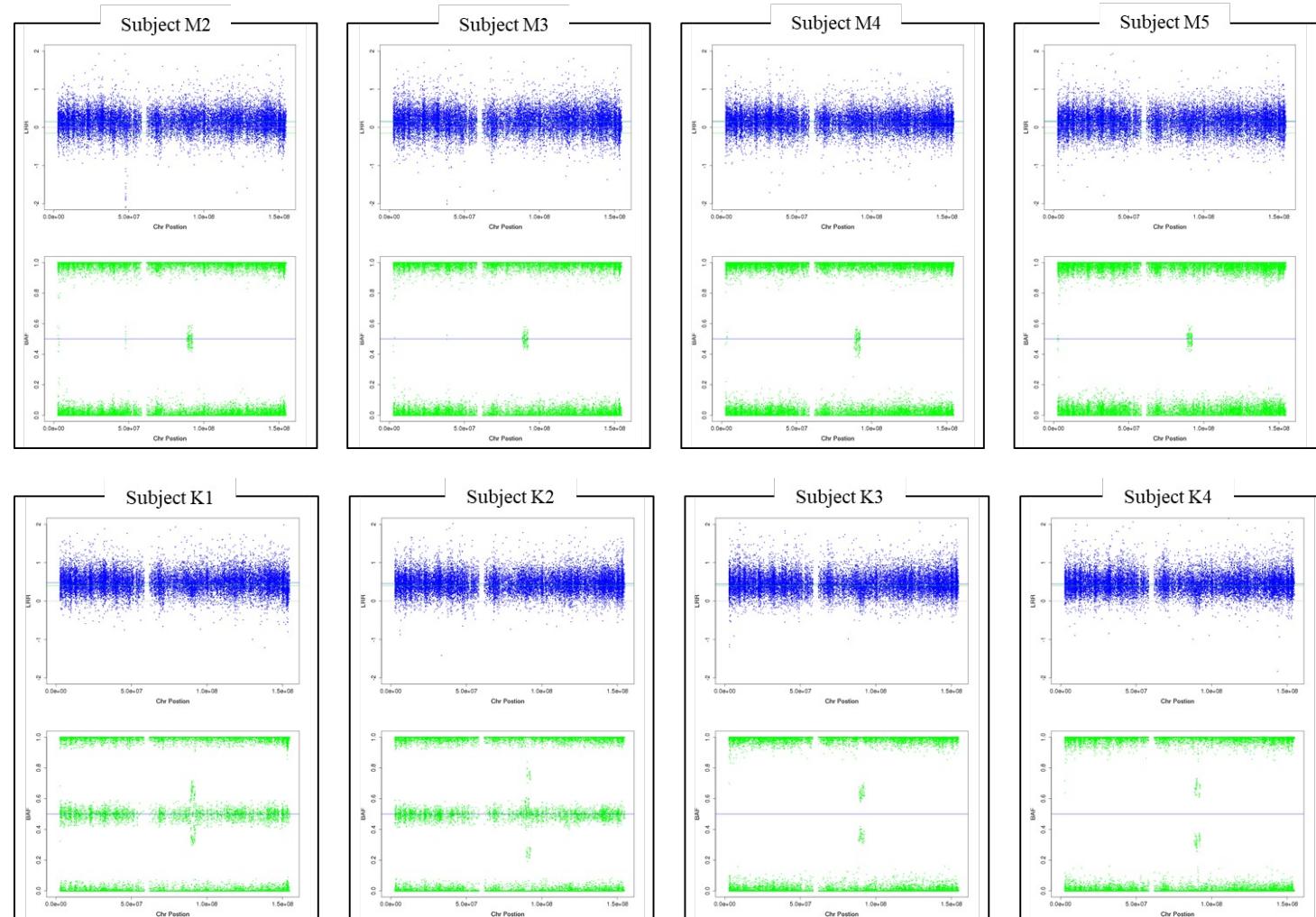
Frequency of Mosaic Y Loss



The most common structural chromosomal alteration seen in men

Frequency of X Mosaicism in Males

- exceptionally rare
- 12 of 196,219 men (0.0006%)
- no cases of X loss
- reflects the deleterious impact on cell survival or clonal fitness



Co-occurring mCAs

Positive association

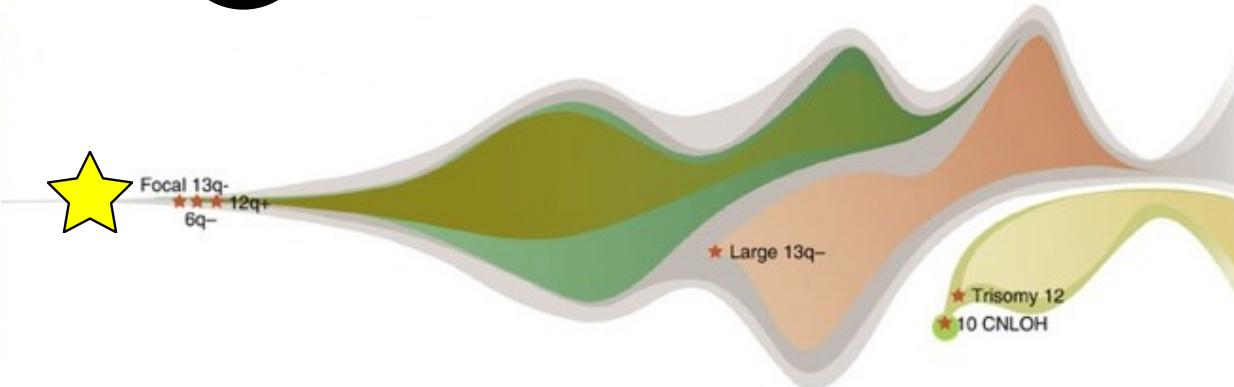
Negative association

* p-value at or below Bonferroni significance threshold of 8.7E-05

What Are We Detecting?

①

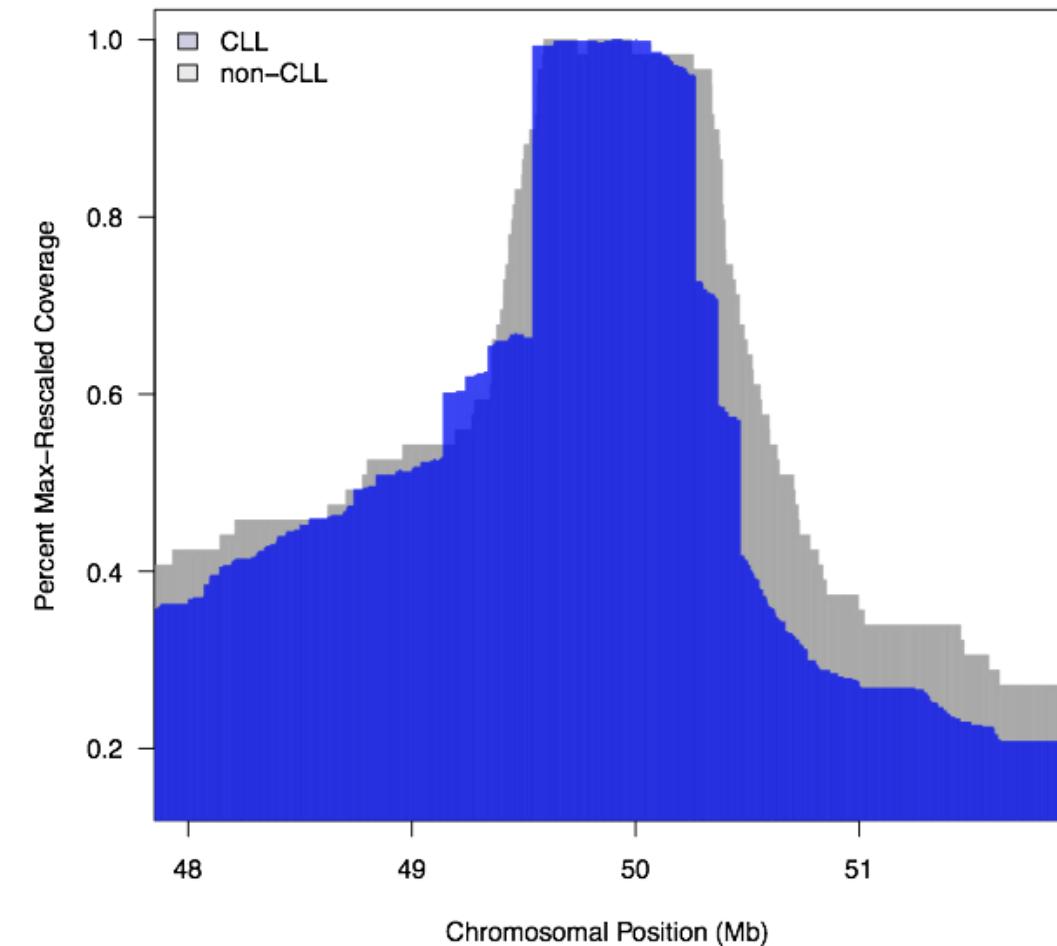
Early Pre-dysplastic Clones



Clones with one of many early events with potential predisposition for future hematologic cancer risk

Early Pre-Leukemic Clones

- Mosaic 20q deletions cover same genomic footprint as del(20q) in myeloid malignancies
- Mosaic 13q14 deletions span the same genomic region as del(13q14) in MBL and CLL
- Frequencies of 20q and 13q14 deletions are higher than population rates of the respective disease
 - Not all individuals with mosaicism progress to disease
 - Mosaicism in blood DNA increases overall hematologic cancer risk (OR>10)
 - Observed up to 15 years prior to diagnosis

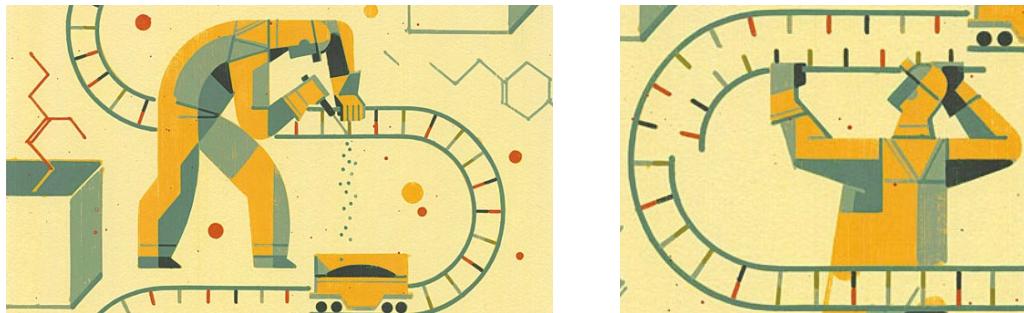


What Are We Detecting?

②

Marker of Genomic Integrity

Interaction of Exposures and Capacity to Maintain/Repair Damage



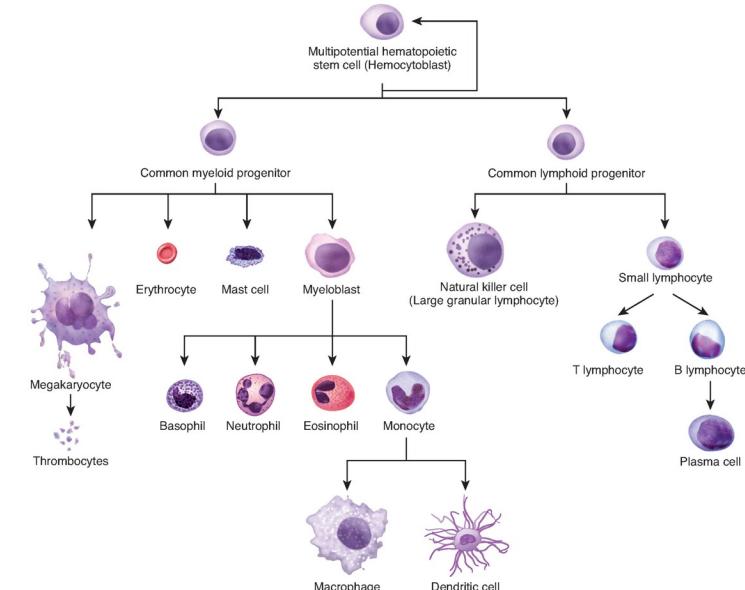
Molecular barometer measuring balance between:

- Acquisition of mutations
- Inborn DNA Repair
- Clonal Selection

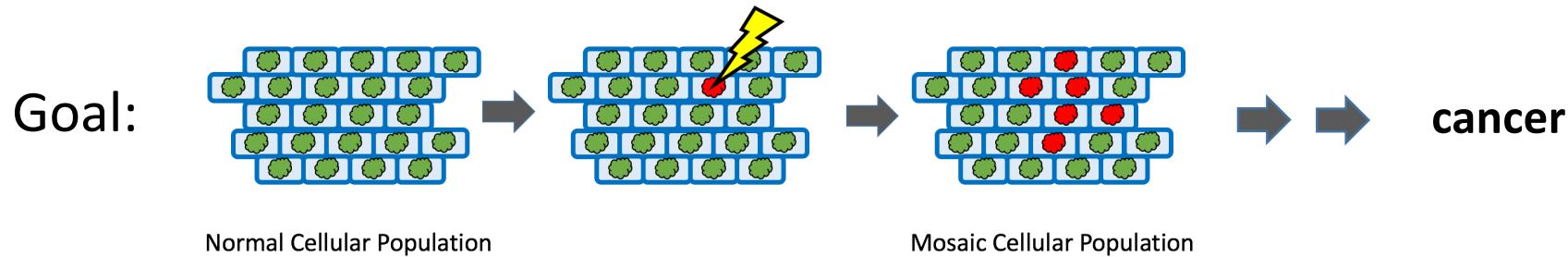
③

Marker of Immune System Function

Hematopoietic stem cell production and differentiation



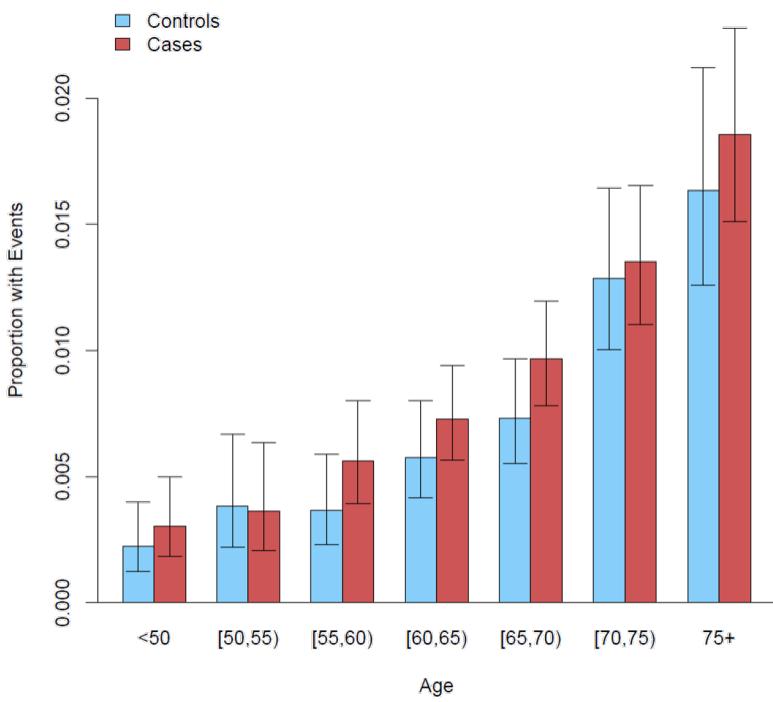
Outline



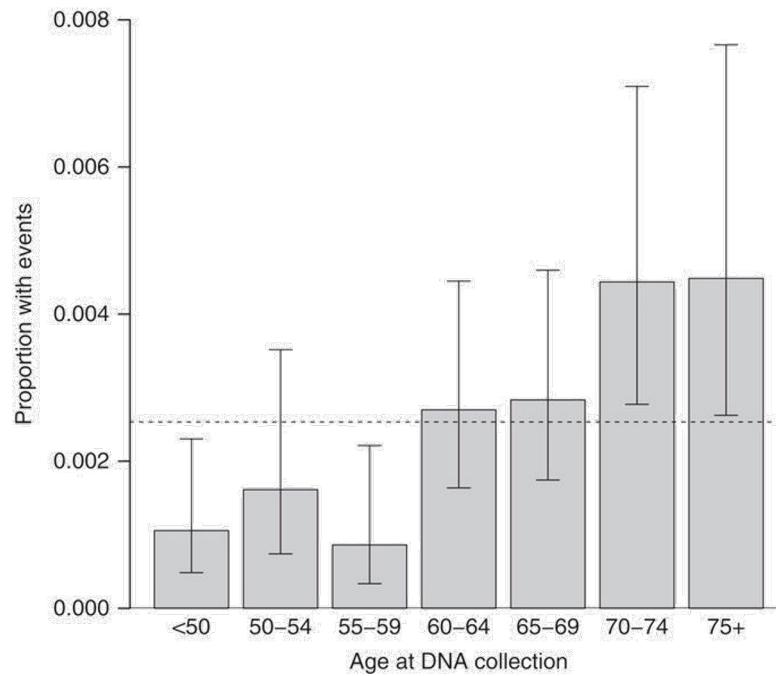
1. How are mosaic chromosomal alterations detected?
2. What risk factors are associated with the development and clonal expansion of mosaic chromosomal alterations?
3. Do mosaic chromosomal alterations alter disease risk (e.g., cancer)?

Robust Association with Increasing Age

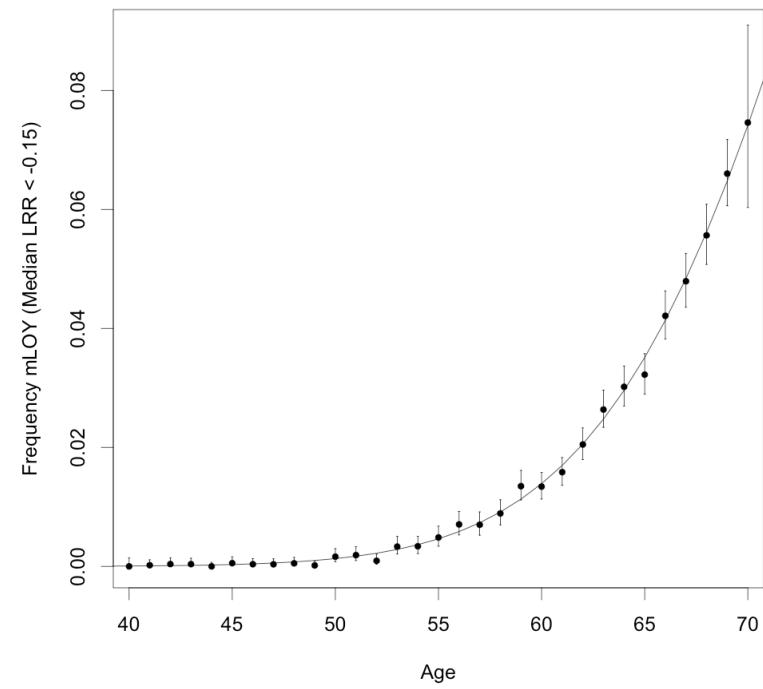
Autosomes



Female X



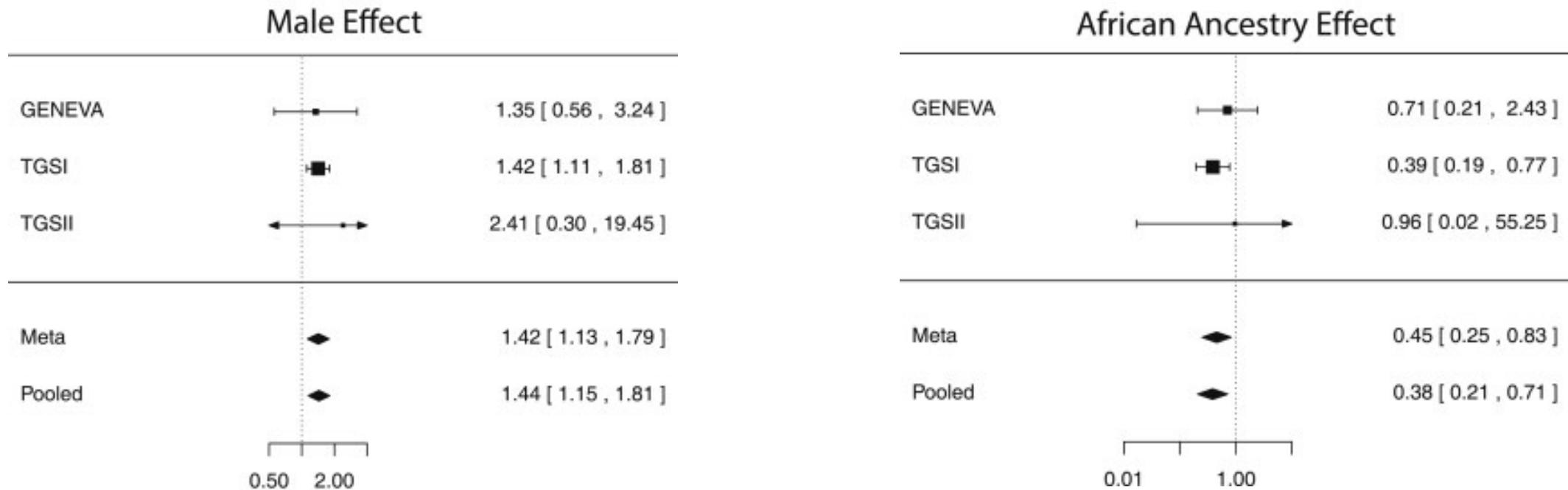
Male Y



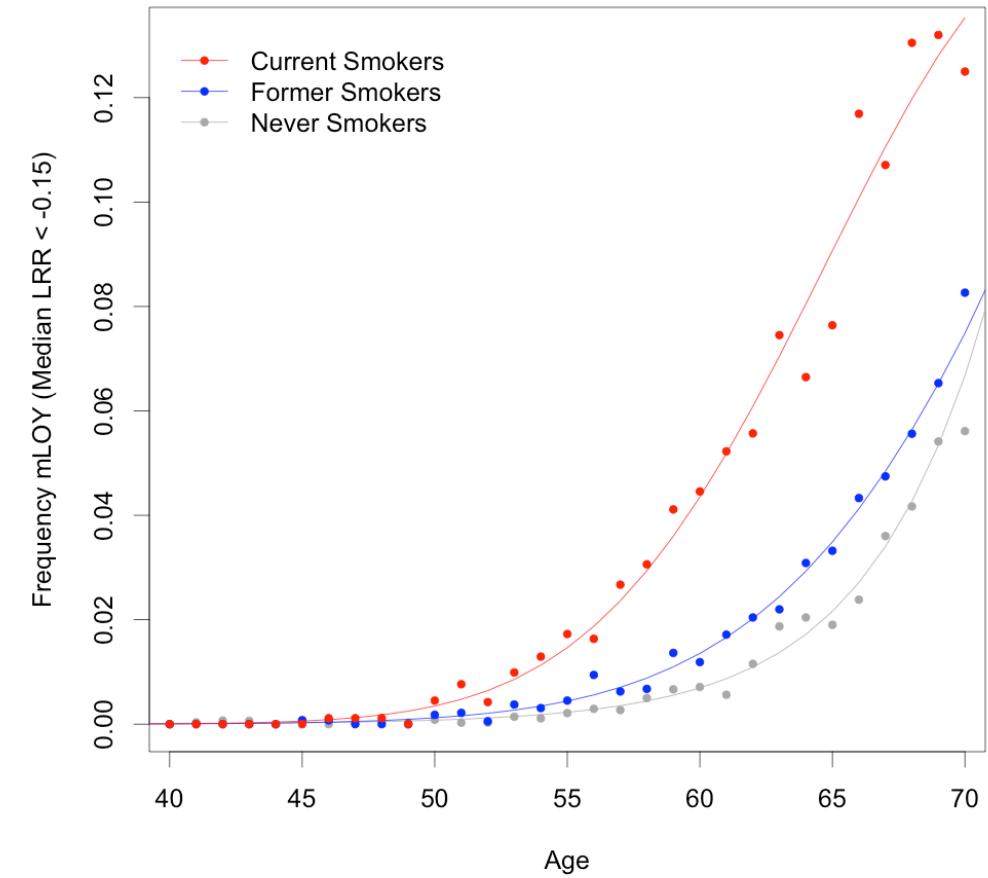
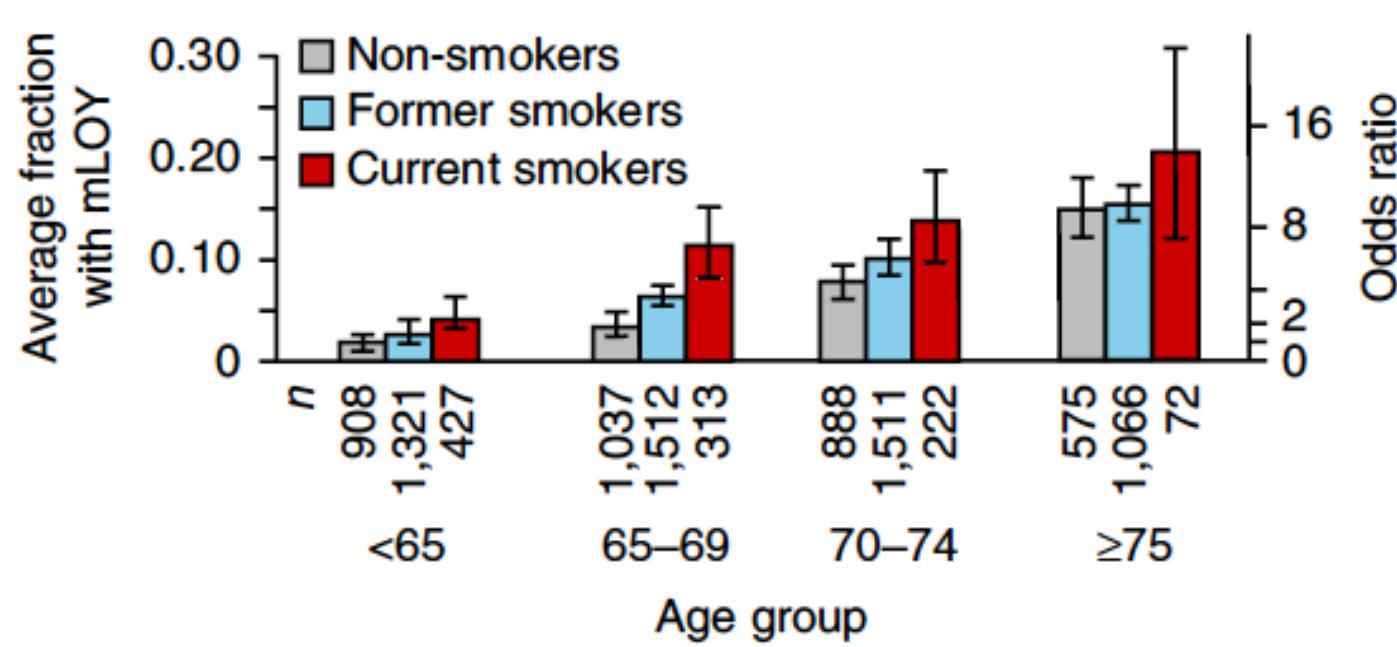
non-linear increase

Association with Sex and Ancestry

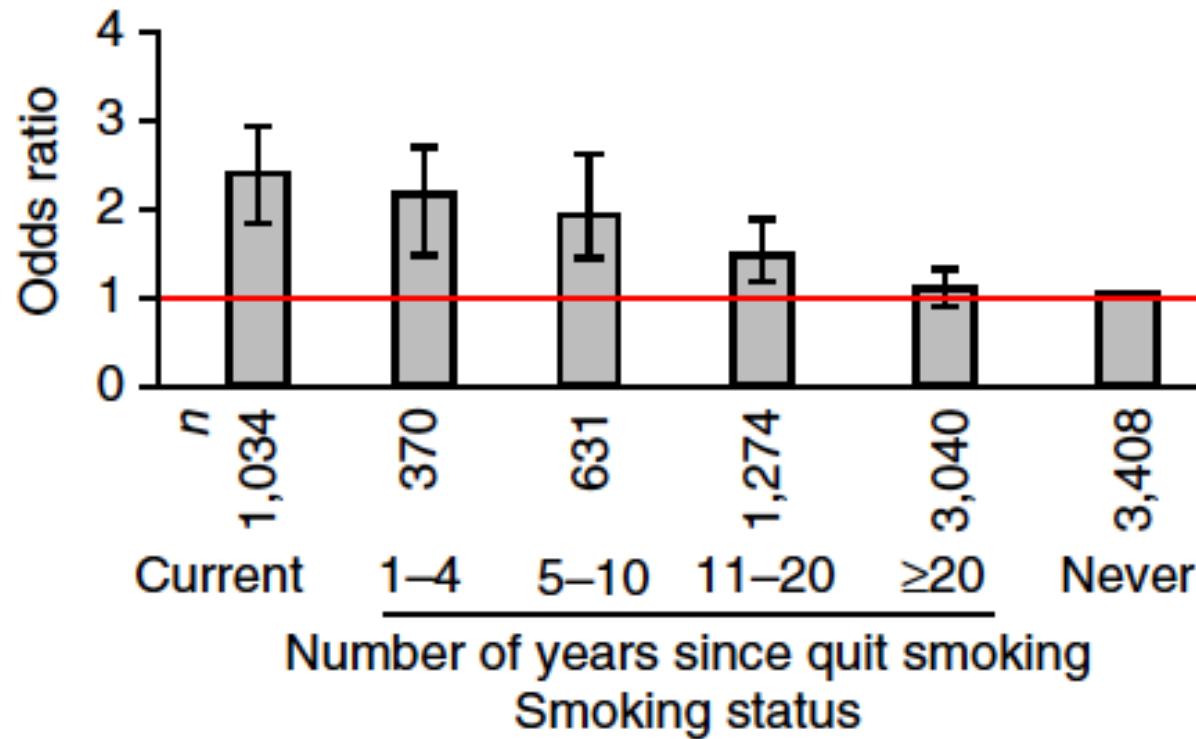
Autosomal Mosaicism



Mosaic Y Loss and Smoking

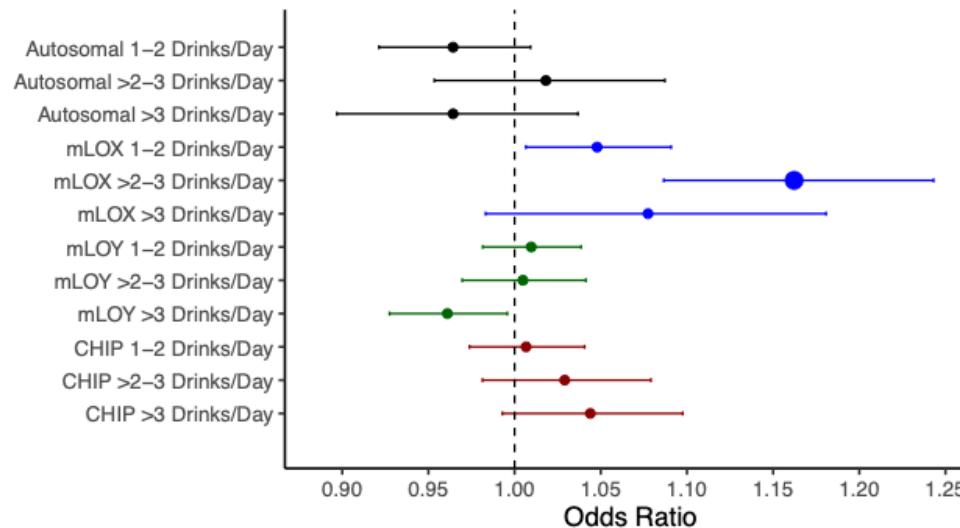


Mosaic Y Loss and Smoking Cessation

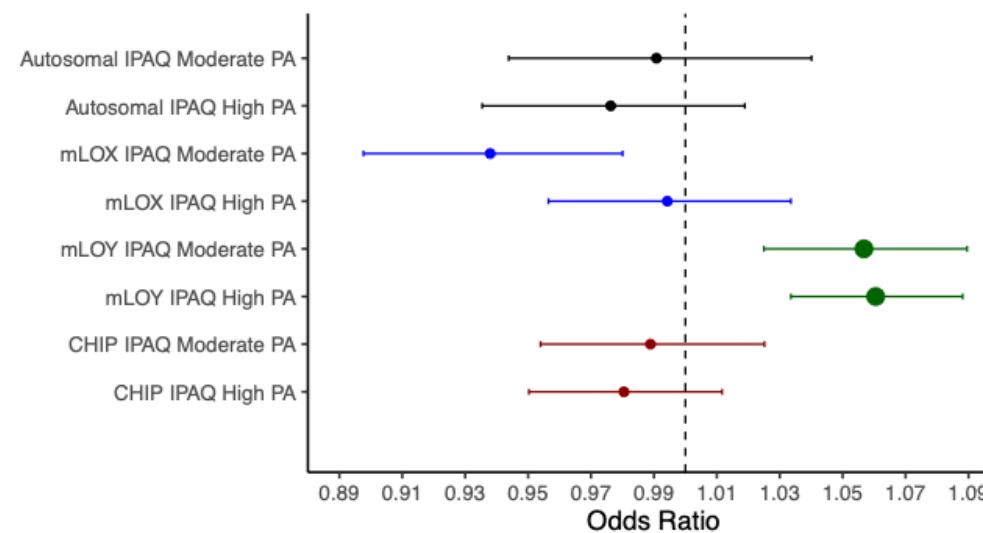


Do Other Modifiable Risk Factors Impact CH Risk?

Alcohol Consumption



Physical Activity



Mosaic Y Loss and Air Pollution (PM_{10})

Table 3

Cross-sectional associations between annual average outdoor residential air pollution concentrations and leukocyte mosaic loss of chromosome Y.

Air pollutant	Age-adjusted				Multivariable-adjusted			
	β	95% CI Lo	95% CI Up	p-Value	β^{Ψ}	95% CI Lo	95% CI Up	p-Value
Coarse particulate matter (PM_{10})								
Tertile 2 vs. 1	-0.0059	-0.0105	-0.0013	0.01*	-0.0044	-0.0094	0.0007	0.09
Tertile 3 vs. 1	-0.0070	-0.0116	-0.0023	3.50E-03*	-0.0054	-0.0104	-0.0003	0.04*
Ordinal p-trend				2.80E-03*				0.03*
Carbon monoxide (CO)								
Tertile 2 vs. 1	0.0001	-0.0047	0.0048	0.98	0.0003	-0.0050	0.0056	0.91
Tertile 3 vs. 1	-0.0007	-0.0055	0.0041	0.77	0.0023	-0.0029	0.0076	0.39
Ordinal p-trend				0.77				0.37
Nitrogen dioxide (NO_2)								
Tertile 2 vs. 1	-0.0011	-0.0058	0.0037	0.66	-0.0027	-0.0081	0.0027	0.32
Tertile 3 vs. 1	0.0014	-0.0033	0.0062	0.56	-0.0012	-0.0082	0.0059	0.75
Ordinal p-trend				0.56				0.72
Ozone (O_3)								
Tertile 2 vs. 1	-0.0036	-0.0083	0.0011	0.14	-0.0029	-0.0076	0.0019	0.23
Tertile 3 vs. 1	0.0004	-0.0044	0.0051	0.88	0.0064	0.0003	0.0126	0.04*
Ordinal p-trend				0.94				0.20
Sulfur dioxide (SO_2)								
Tertile 2 vs. 1	-0.0006	-0.0053	0.0040	0.79	0.0018	-0.0039	0.0075	0.54
Tertile 3 vs. 1	0.0020	-0.0027	0.0066	0.41	0.0010	-0.0126	0.0146	0.88
Ordinal p-trend				0.44				0.62

The outcome was the fluorescent signal intensity (median log R ratio (mLRR)) of a male-specific region of ChrY: 7502455–21979204 measured using Illumina 370CNV-Duo genotyping arrays. Decreased mLRR reflects greater mosaic loss of chromosome Y (mLOY) in a population of leukocytes. The effective sample size was 933 men for the age-adjusted analyses and 897 for the multivariable-adjusted analyses due to missing smoking pack-year data. Tertiles (PM_{10} : < 28.48, ≥ 28.48 to < 36.08 , $\geq 36.08 \mu\text{g}/\text{m}^3$; CO: < 0.98, ≥ 0.98 to < 1.28 , $\geq 1.28 \text{ ppb}$; NO_2 : < 17.25, ≥ 17.25 to < 25.95 , $\geq 25.95 \text{ ppb}$; O_3 : < 19.60, ≥ 19.60 to < 27.95 , $\geq 27.95 \text{ ppb}$; SO_2 : < 9.87, ≥ 9.87 to < 16.12 , $\geq 16.12 \text{ ppb}$).

* p-Values < 0.05 were considered statistically significant.

Ψ Single pollutant linear regression models were adjusted for clinic, race/cohort, age, BMI, smoking status and pack-years.



Arsenic and Glyphosate use associated with mLOY

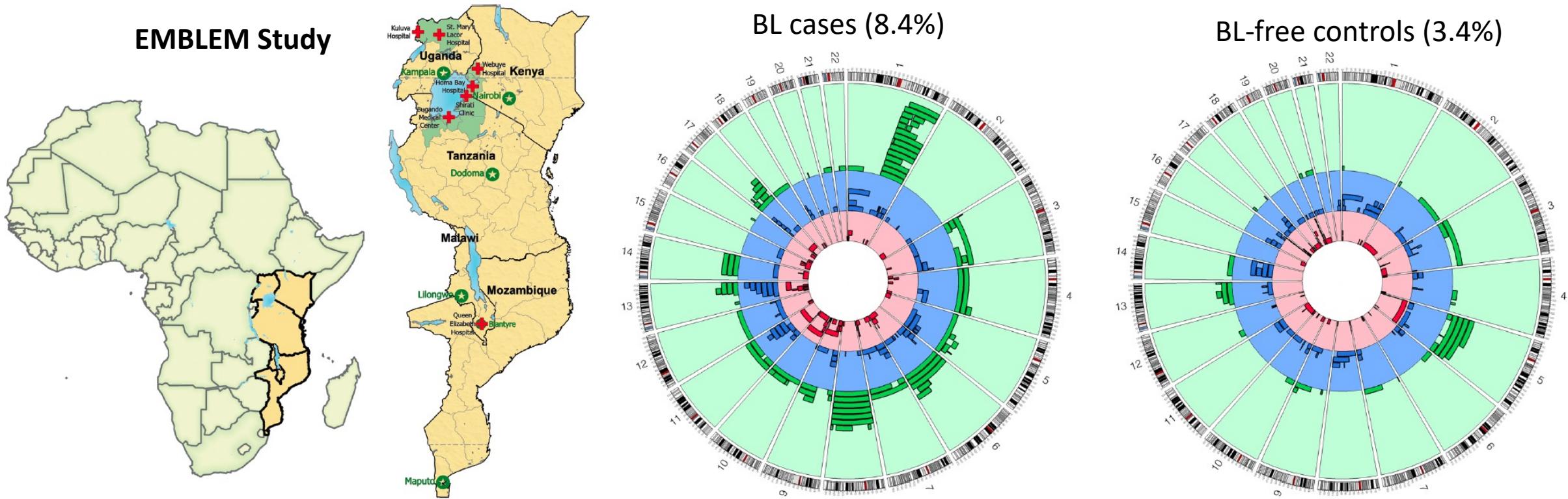
Arsenic in drinking water in Bangladeshi men

	n	Percentage of cells with LoY (%)		
		Mean (SD)	β (SE)	P
Water arsenic ($\mu\text{g/L}$)				
<17	195	0.39 (3.18)	Ref.	–
17–71	191	0.78 (3.25)	0.42 (0.27)	0.11
72–153	196	1.06 (3.35)	0.55 (0.27)	0.043
>153	196	1.35 (3.66)	0.75 (0.27)	0.0054
Ordinal ^b			0.24 (0.09)	0.0056

Occupational glyphosate use male farmers in the Agricultural Health Study

Glyphosate use	N _{total}	Expanded mLOY ^b		
		n _{no}	n _{yes}	OR (95% CI) ^c
Total lifetime days				
Quartile 1 (0–26)	402	351	51	1.00 (Ref)
Quartile 2 (>26–76)	402	363	39	1.01 (0.62, 1.65)
Quartile 3 (>76–161)	403	371	32	1.23 (0.72, 2.07)
Quartile 4 (>161–3,063)	399	363	36	1.75 (1.00, 3.07)
				<i>p</i> _{trend^d} = 0.03

Elevated Frequencies in Children Living in Sub-Saharan Africa



- mCA frequencies in SSA children are comparable to cancer-free adults of European ancestry in the US
- mCA frequencies in cancer-free men from SSA are at least 2-fold higher than in cancer-free European-American adults
- These findings point to a possible role of environmental factors (e.g., malaria or EBV) in elevating the frequency of mCAs in SSA individuals

Telomere Length and Autosomal Mosaicism

Table 4. Association between genetically-inferred telomere length and autosomal SCNAs by event type and copy number change

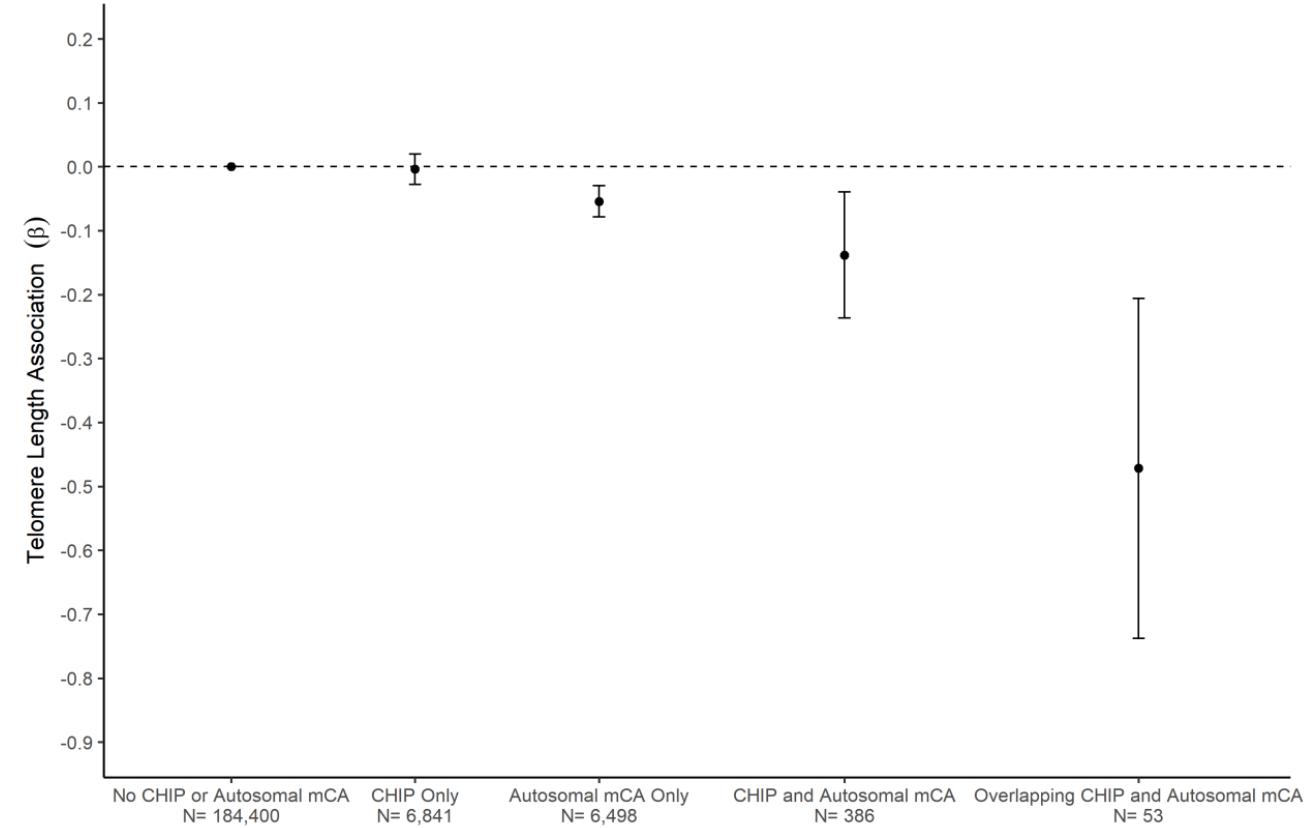
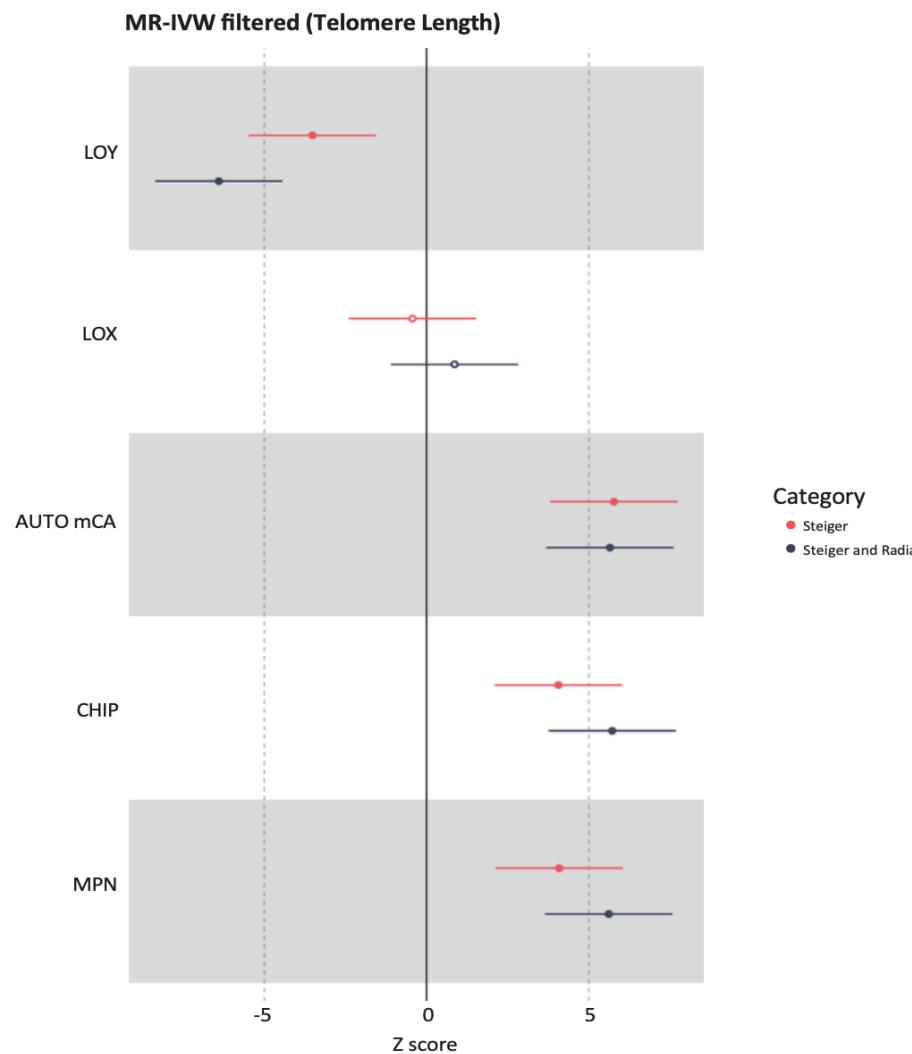
	Univariable Model		Multivariable Model		p-value*
	OR (95% CI)	p-value	OR (95% CI)	p-value	
Overall	1.05 (1.033-1.067)	3.61x10 ⁻⁹	1.07 (1.052-1.087)	1.61x10 ⁻¹⁵	
Event Type					0.3707
Telomeric	1.05 (1.027-1.067)	2.19x10 ⁻⁶	1.07 (1.045-1.086)	1.13x10 ⁻¹⁰	
Interstitial	1.08 (1.042-1.129)	6.47x10 ⁻⁵	1.10 (1.057-1.146)	3.51x10 ⁻⁶	
Whole	1.05 (1.005-1.088)	0.0262	1.07 (1.028-1.114)	0.0009	
Copy Number Change					0.0502
Gain	1.04 (0.995-1.089)	0.0788	1.06 (1.009-1.106)	0.0191	
Loss	1.10 (1.064-1.147)	1.99x10 ⁻⁷	1.13 (1.085-1.171)	8.46x10 ⁻¹⁰	
Neutral	1.05 (1.026-1.075)	3.92x10 ⁻⁵	1.07 (1.044-1.095)	3.31x10 ⁻⁸	
Undetermined	1.04 (1.006-1.067)	0.0185	1.06 (1.025-1.088)	3.51x10 ⁻⁴	

Multivariable models control for sex, age, age², genetic ancestry, and detailed smoking status

*Denotes test of heterogeneity

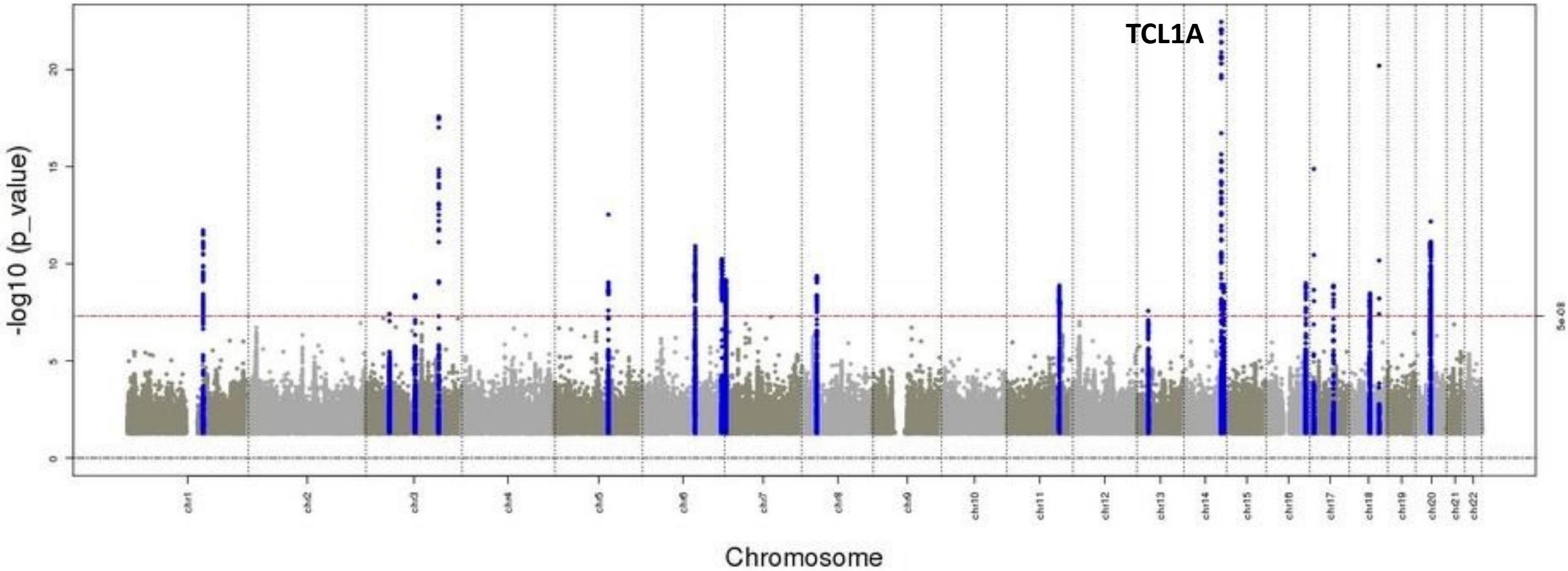
-- positive association between telomere length and mosaic fraction, suggesting impact on replicative potential

Dynamic Association of CH with Telomere Length

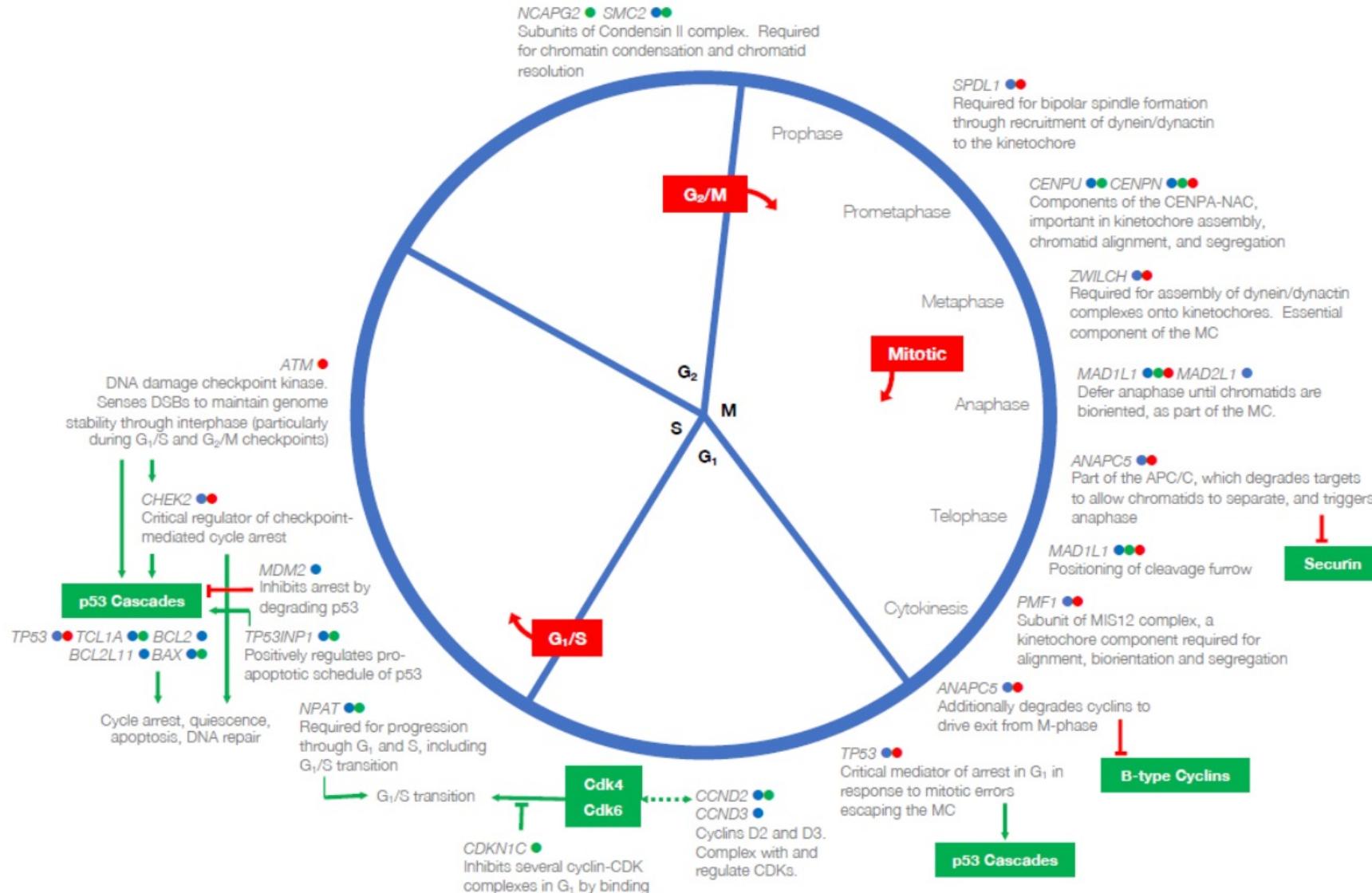


Mosaic Y Loss and Germline Susceptibility Variants

156 Independent Autosomal Genomic Regions

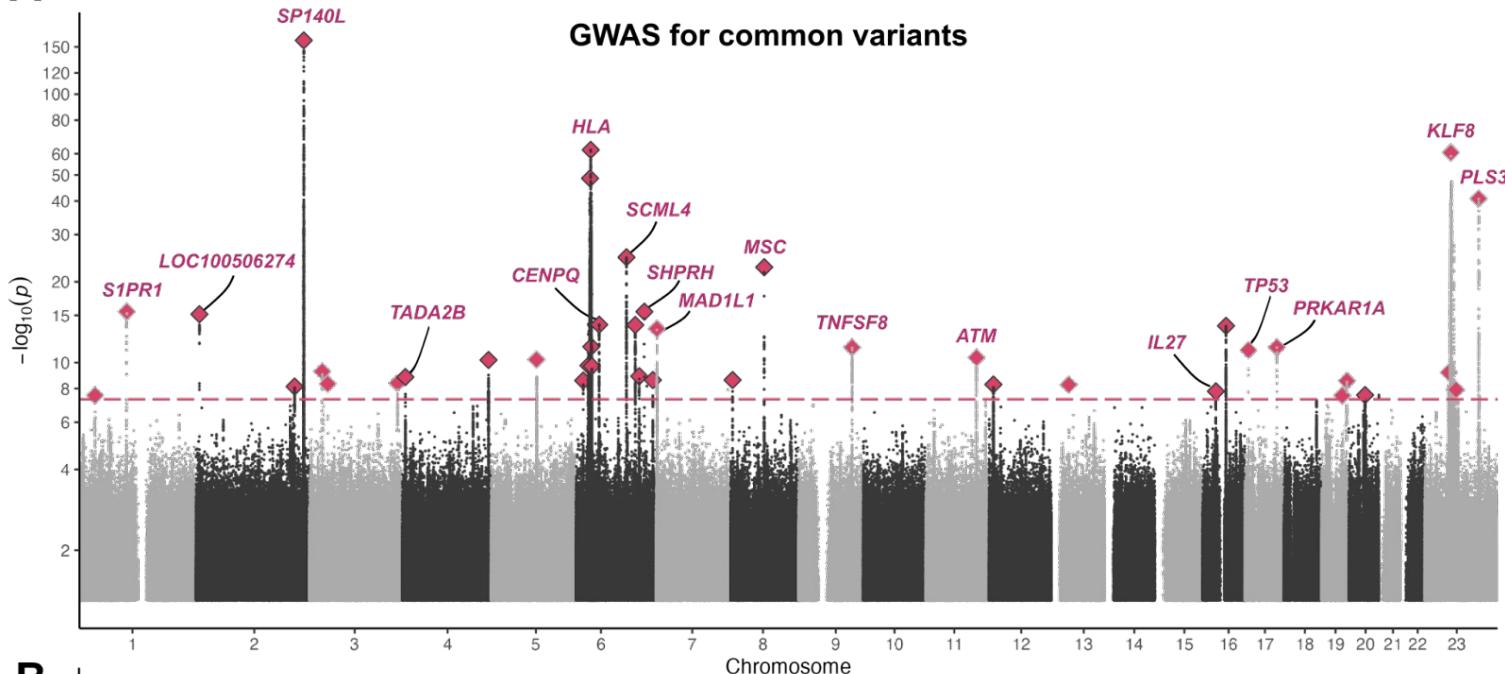


Cell Cycle Regulation and DNA Damage Response

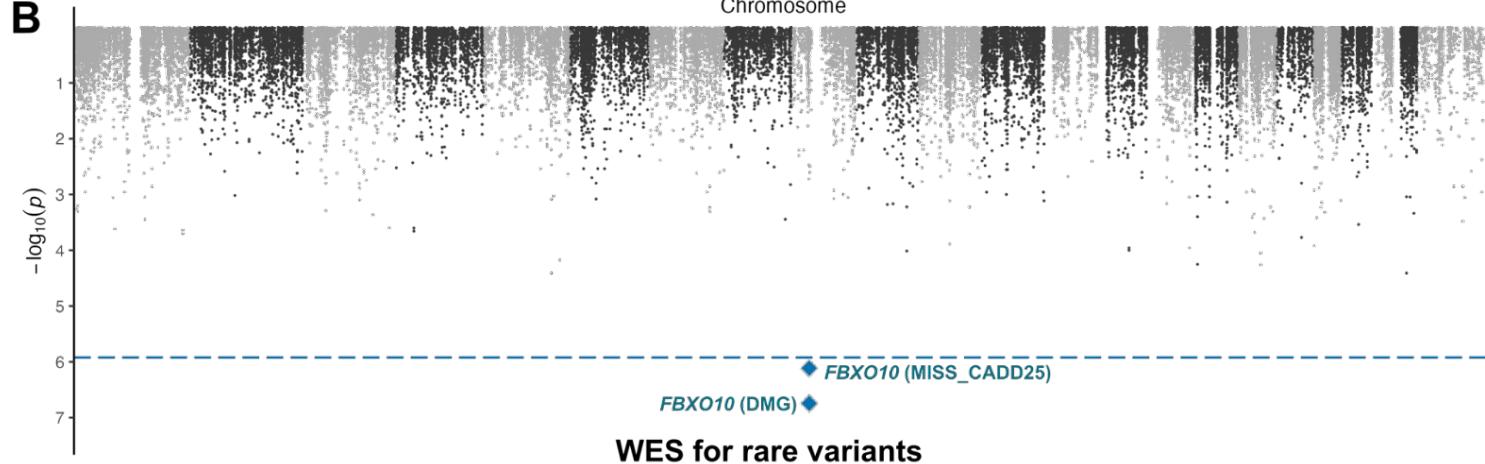


Mosaic X Loss and Germline Susceptibility

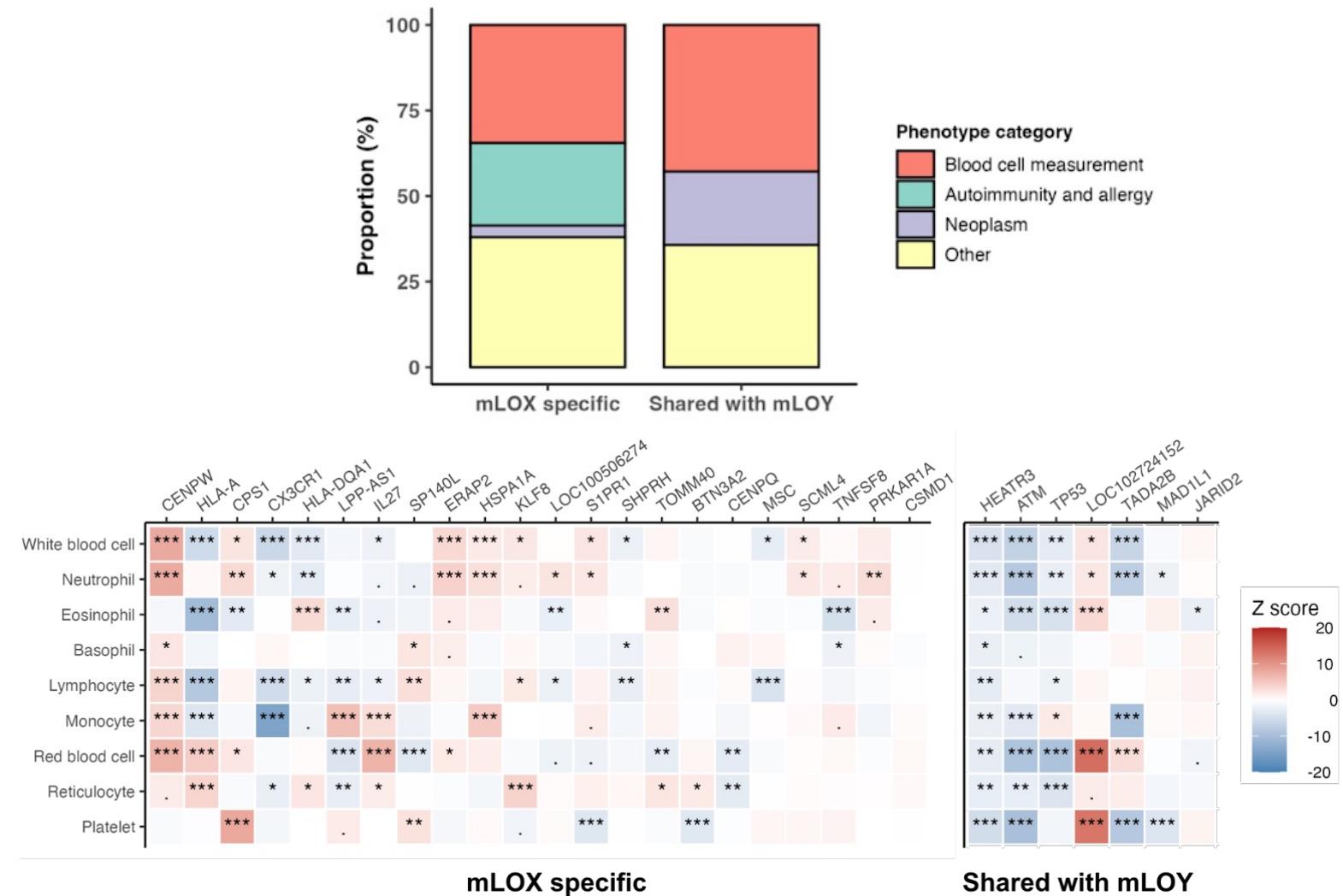
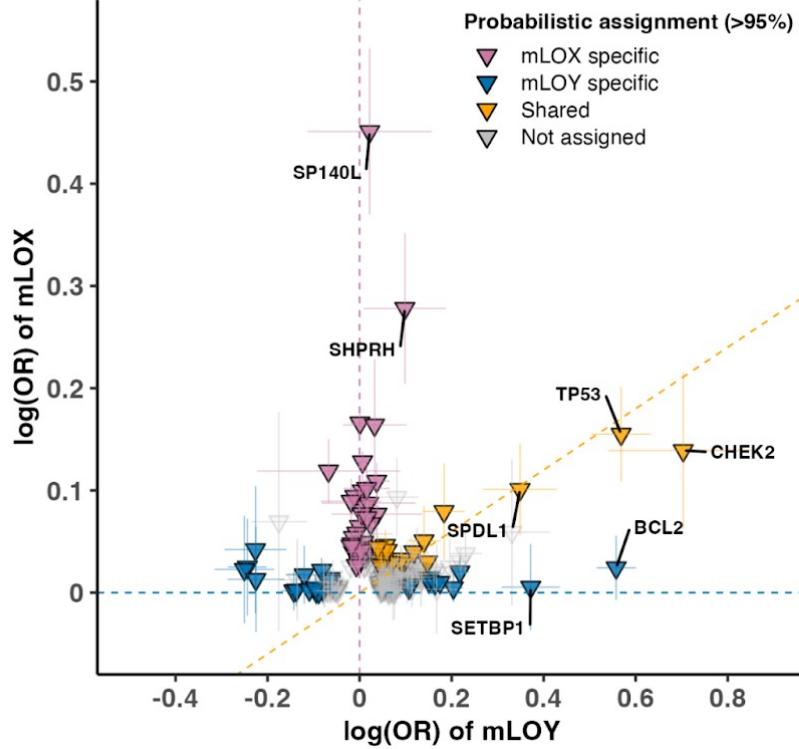
A



B

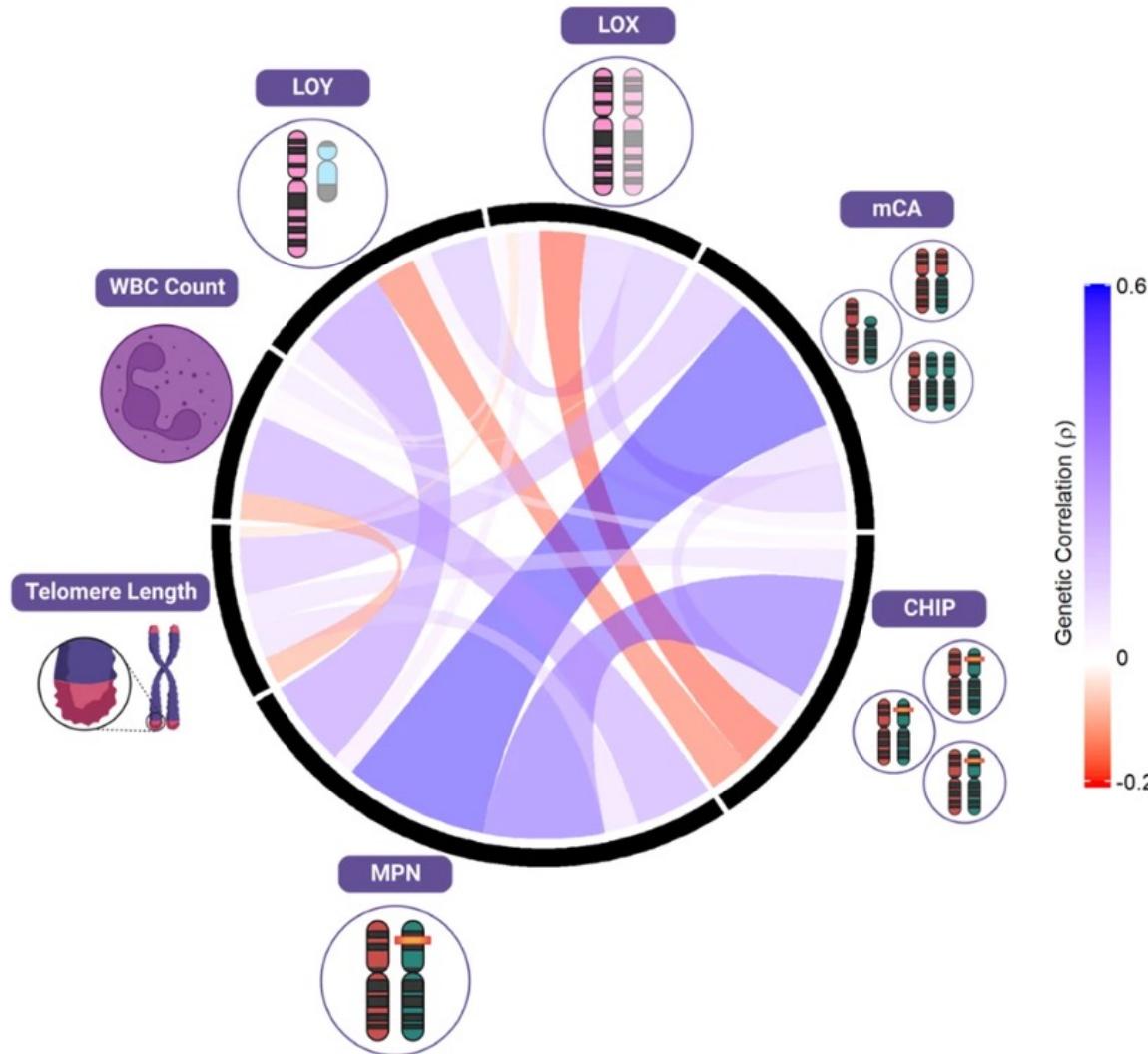


Shared mLOY and mLOX Germline Susceptibility

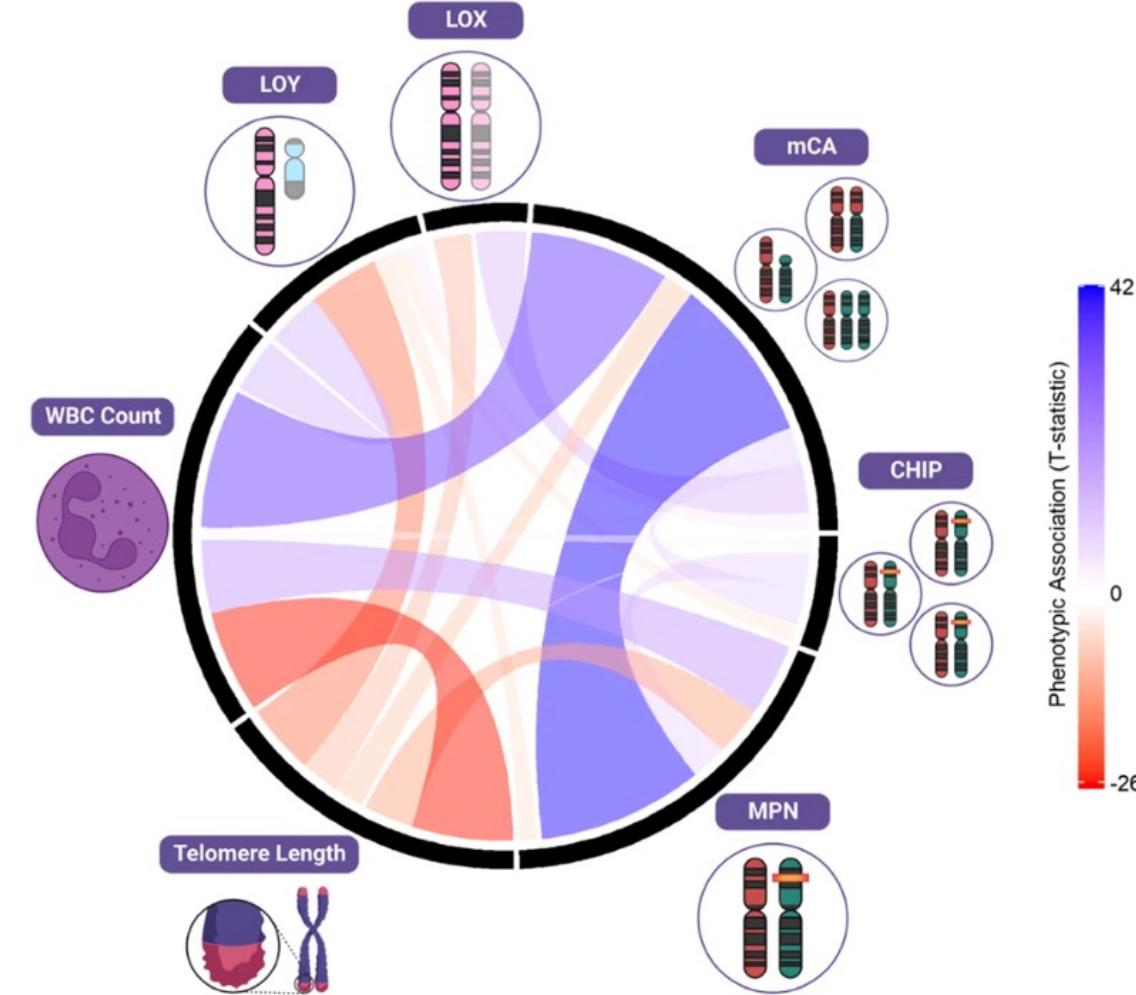


Shared Genetic Etiologies Across Types of CH

Genetic Correlation (HDL)

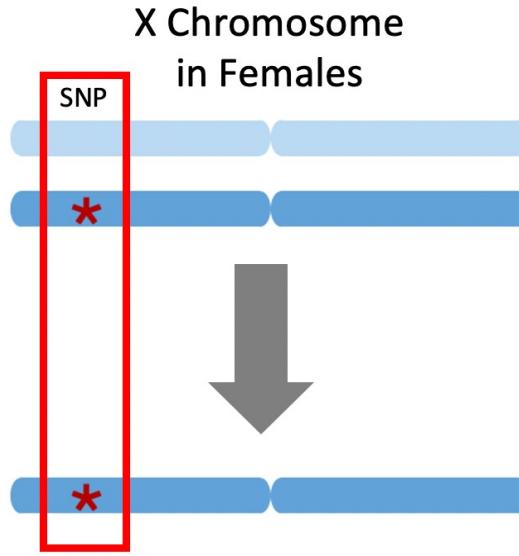


Phenotypic Correlation

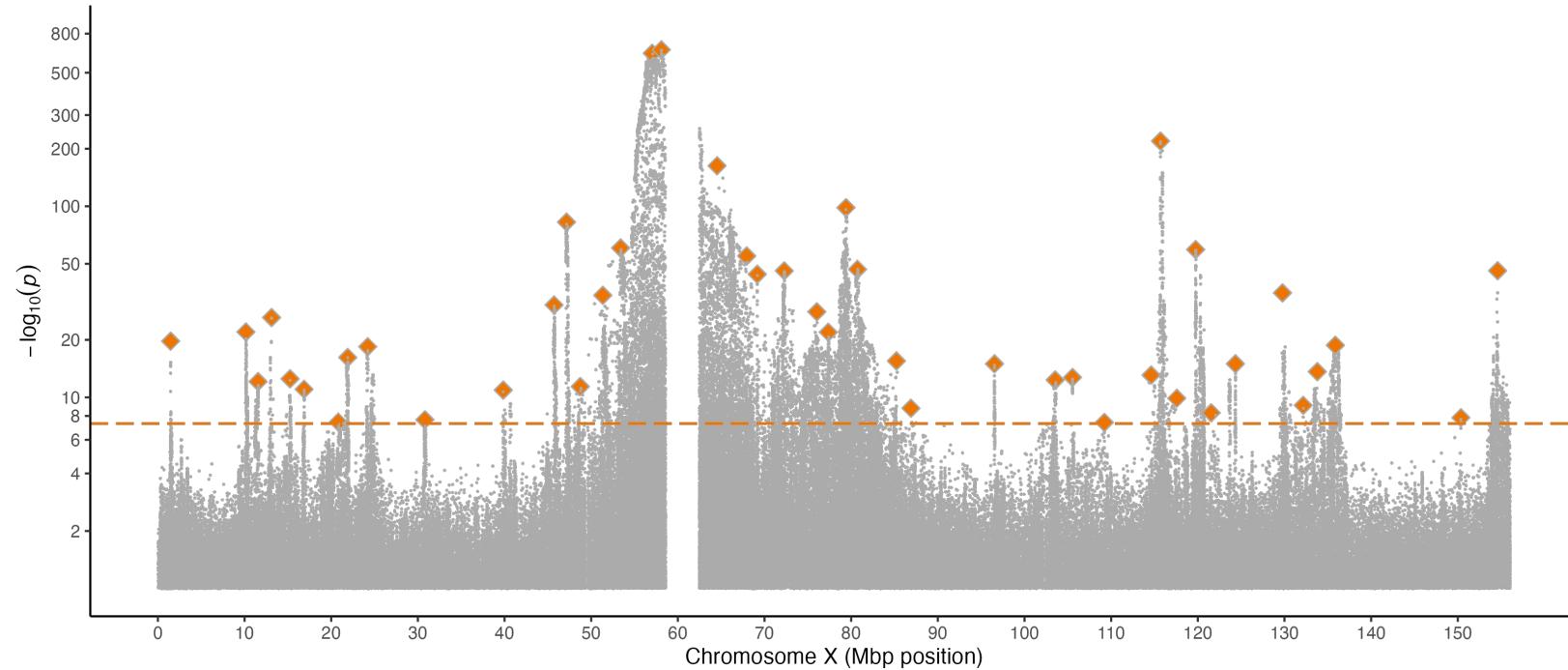


Chromosome X Alleles are Preferentially Retained

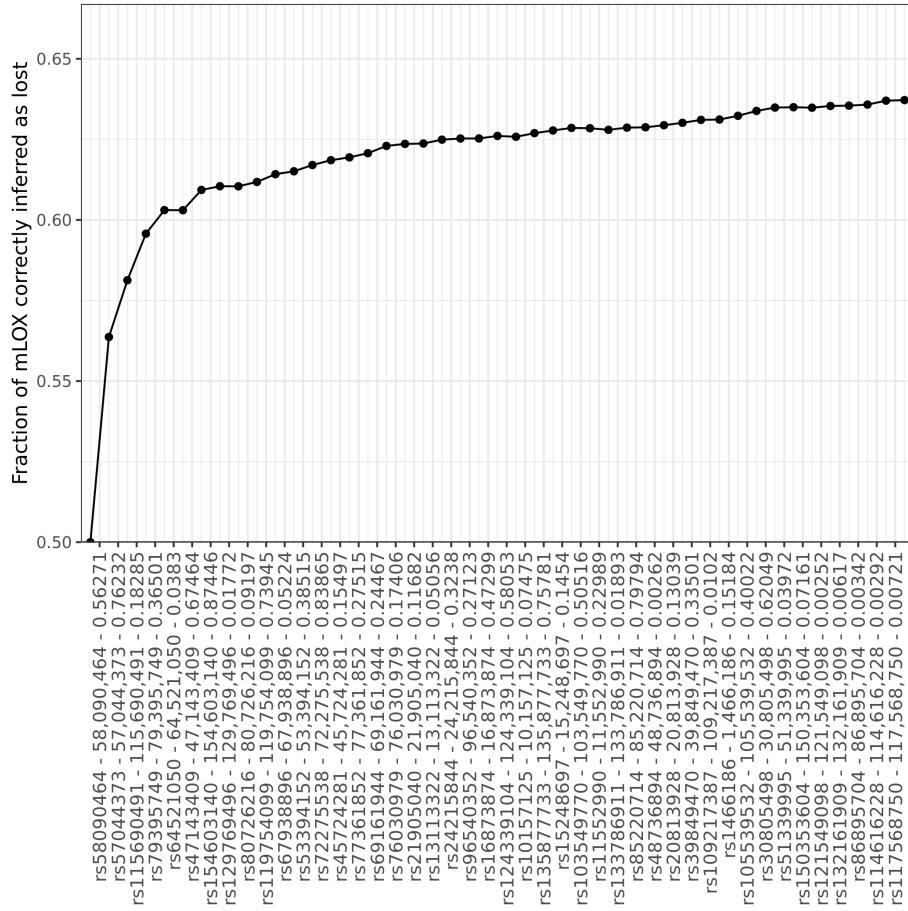
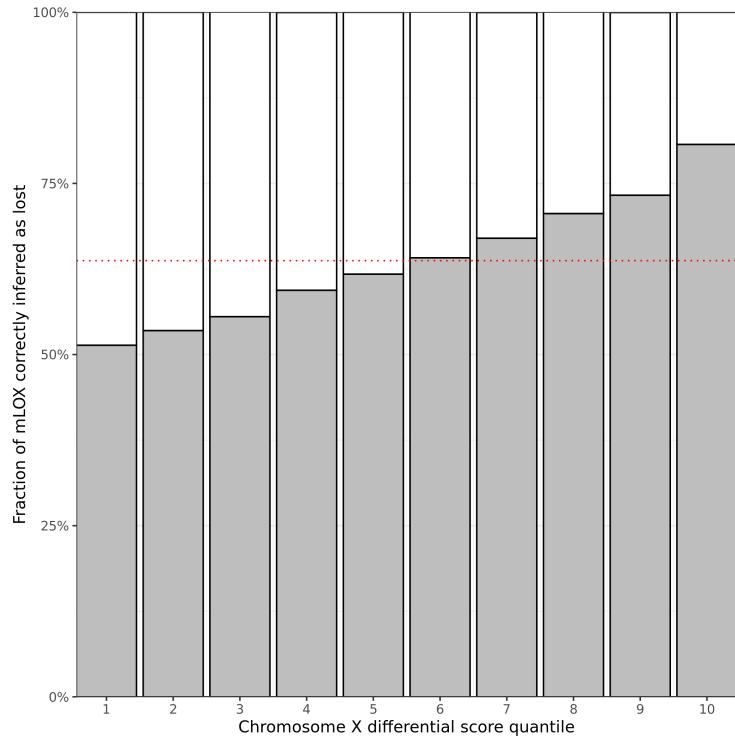
Allelic Shift Analysis



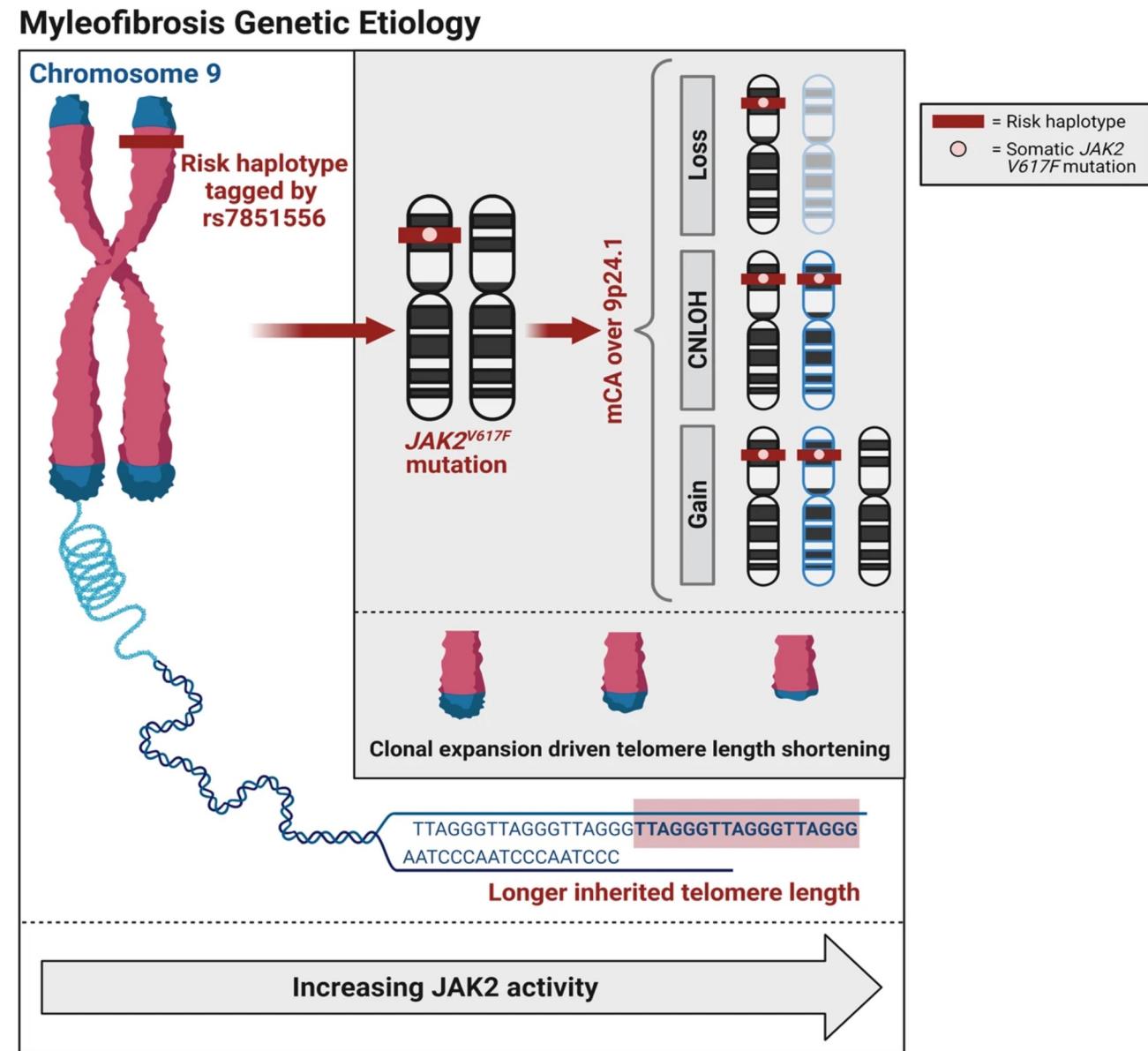
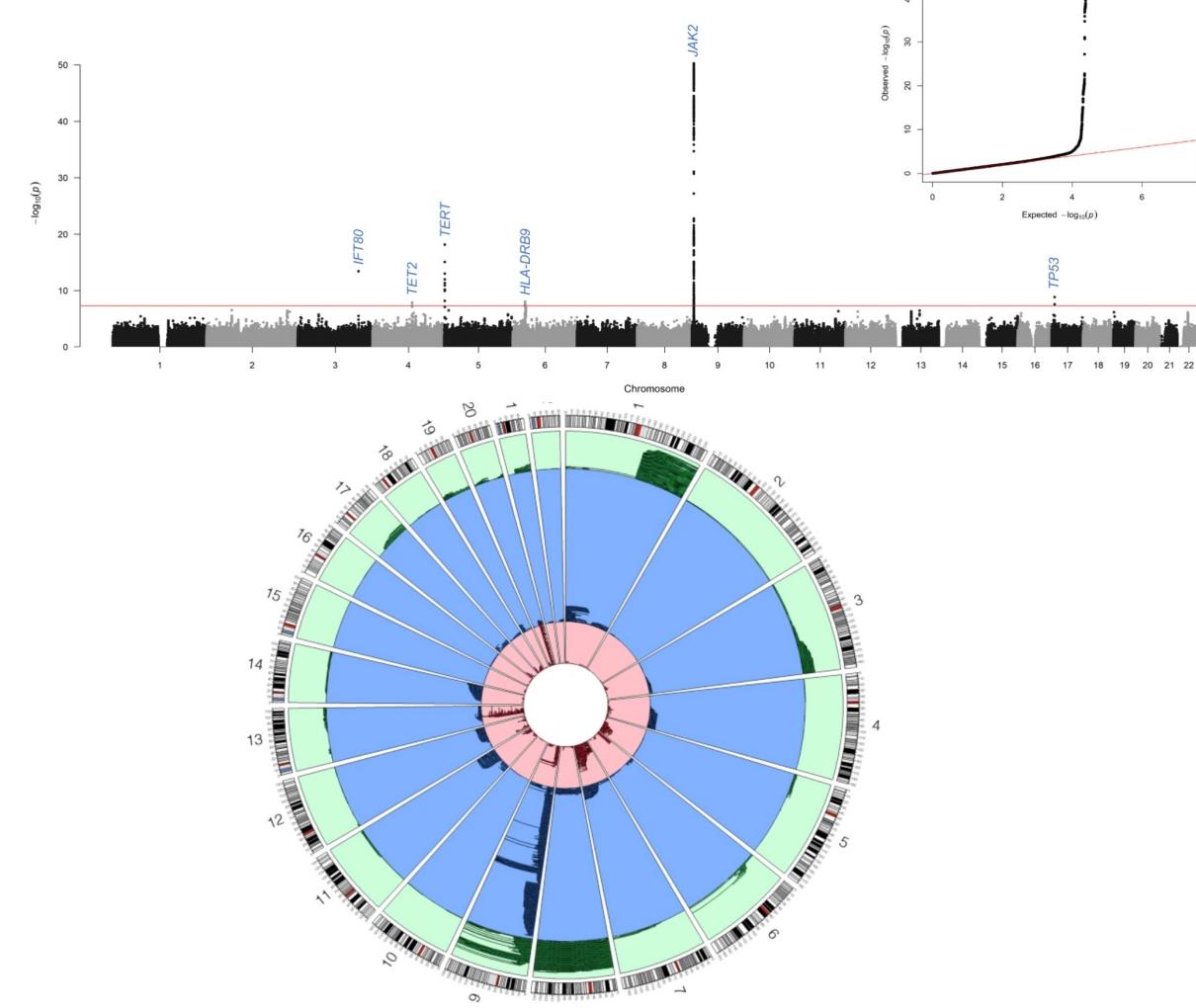
- similar to a transmission disequilibrium test (TDT) in family studies
- essentially a Fisher exact test



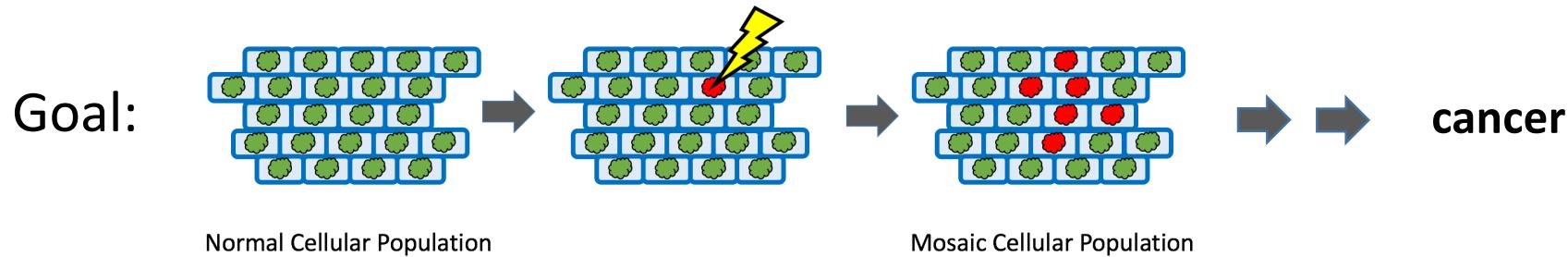
Chromosome X Differential Score Predicts Which X Chromosome is Retained when mLOX Occurs



cis Relationships in Myelofibrosis



Outline



1. How are mosaic chromosomal alterations detected?
2. What risk factors are associated with the development and clonal expansion of mosaic chromosomal alterations?
3. Do mosaic chromosomal alterations alter disease risk (e.g., cancer)?

Phenotypic Spectrum



Apparently Healthy



Eye pigmentation



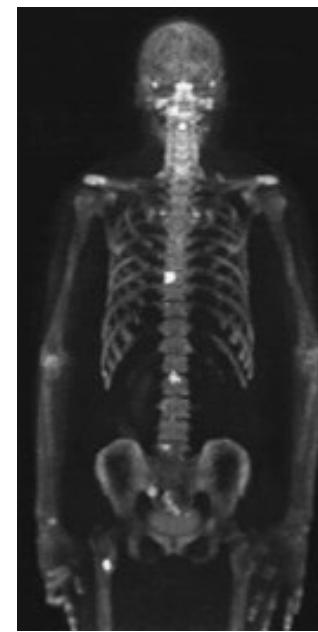
Blaschko's lines



Proteus Syndrome



CVD



Cancer

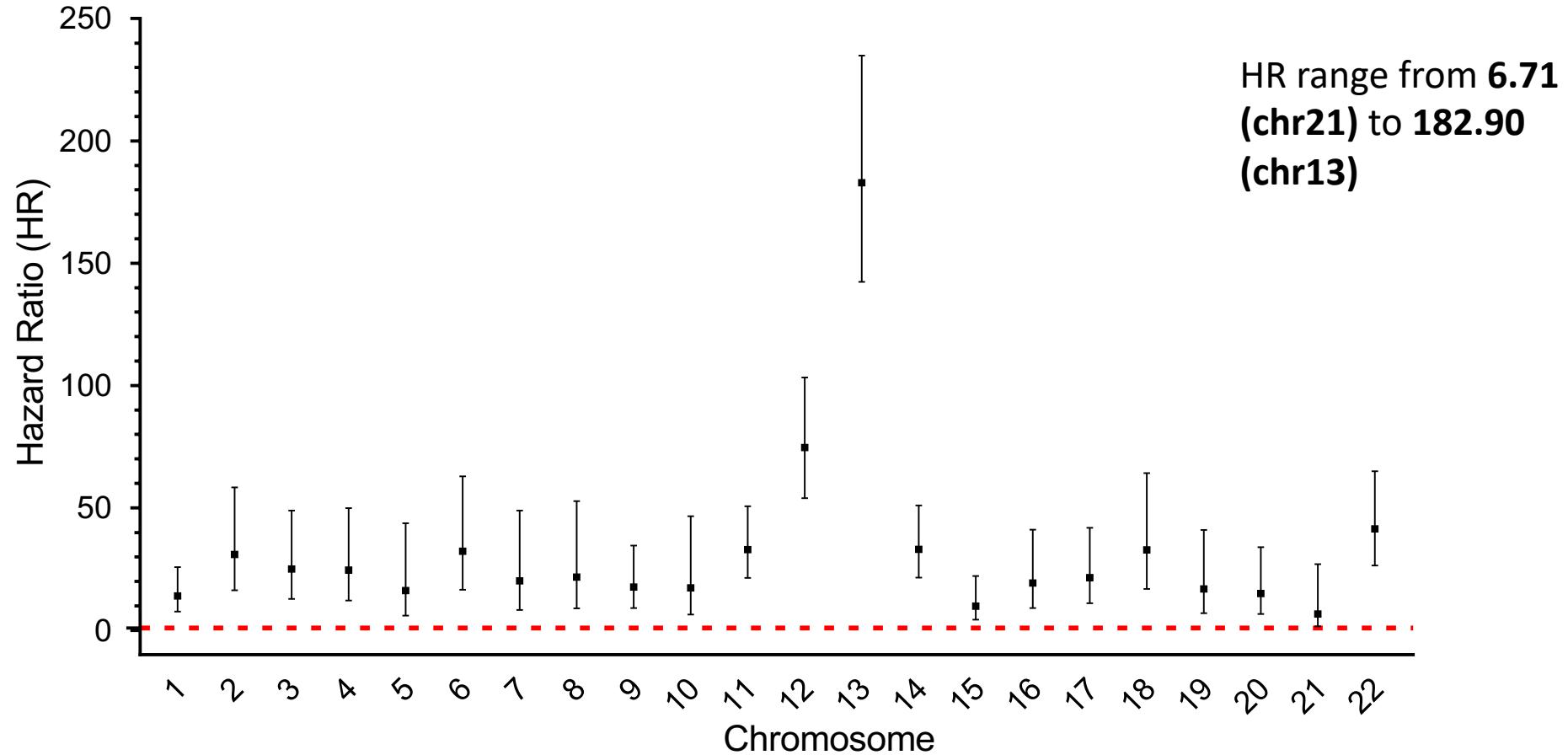
Dependent on:

- Developmental timing
- Tissue type affected
- Genomic location of alteration
- Percentage of cells affected

Autosomal Mosaicism and Incident Cancer Risk

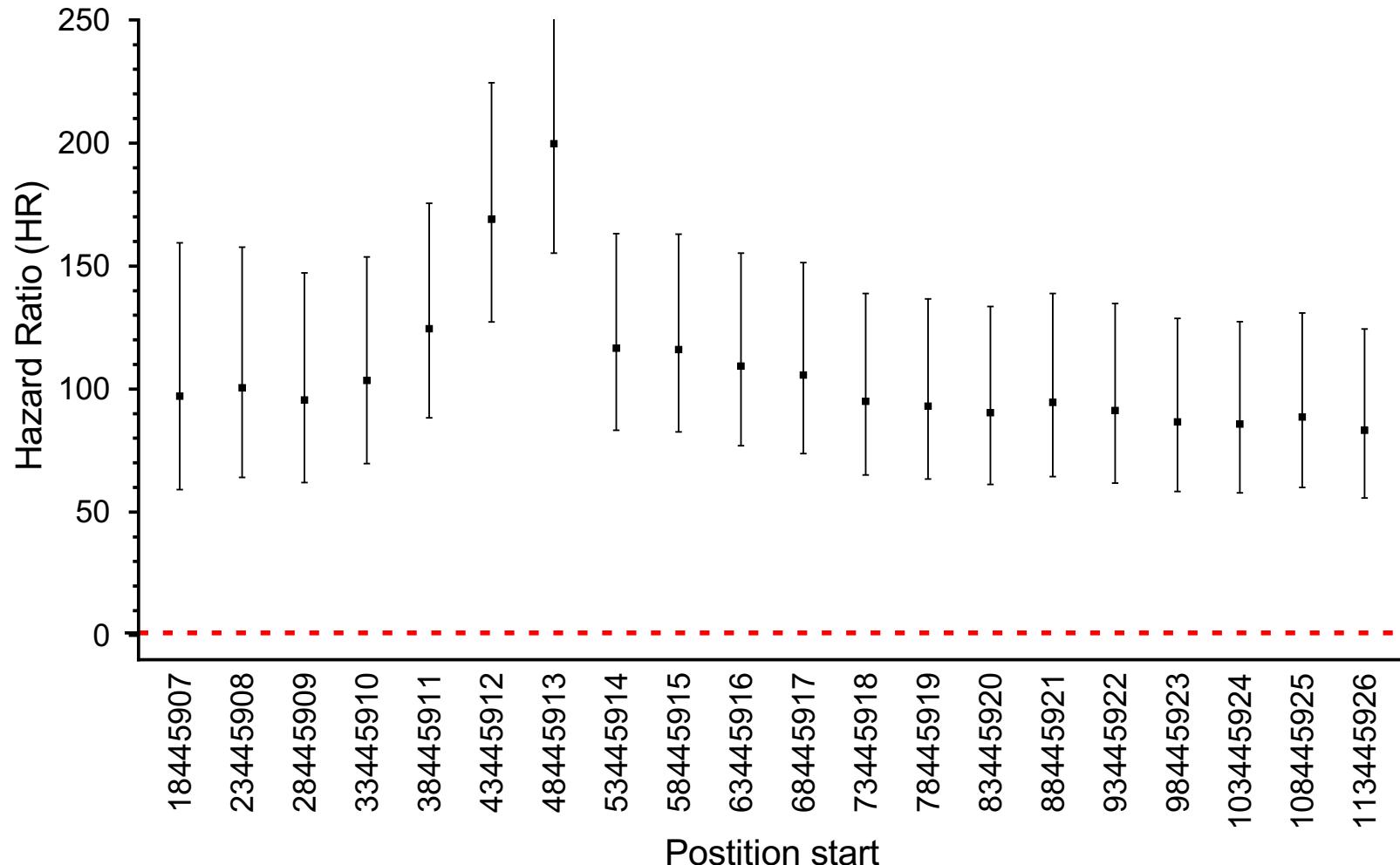
	Effect of expanded mCAs on Incident Cancer	HR	95% CI
Autosomal			
Hematologic Cancer	+	16.87	[15.03; 18.9]
Lymphoid Leukemia	→	99.35	[77.86; 126.8]
CLL	→→	121.92	[93.56; 158.9]
MPN	→	17.21	[13.14; 22.5]

Individual mCAs with CLL risk



Regional mCAs and CLL Risk

Chromosome 13



Building Integrative Models of CLL Risk

	Test Set AUC
Model 0: Age, age-squared, sex, genetic ancestry, smoking + blood cell traits	0.8372
Model 1: Model 0 + CLL PGS	0.8617
Model 2: Model 1 + any autosomal mCA	0.8824
Model 3: Model 1 + any CHIP mutation	0.8770
Model 4: Model 1 + any autosomal mCA + any CHIP mutation	0.8864

5 vs 10 Year CLL Prediction Models

	5-yr AUC	10-yr AUC
Model 0: Age, age-squared, sex, genetic ancestry, smoking + blood cell traits	0.8372	0.8091
Model 1: Model 0 + CLL PGS	0.8617	0.8304
Model 2: Model 1 + any autosomal mCA	0.8824	0.8586
Model 3: Model 1 + any CHIP mutation	0.8770	0.8446
Model 4: Model 1 + any autosomal mCA + any CHIP mutation	0.8864	0.8648

Mosaic Y Loss and Incident Solid Tumor Risk

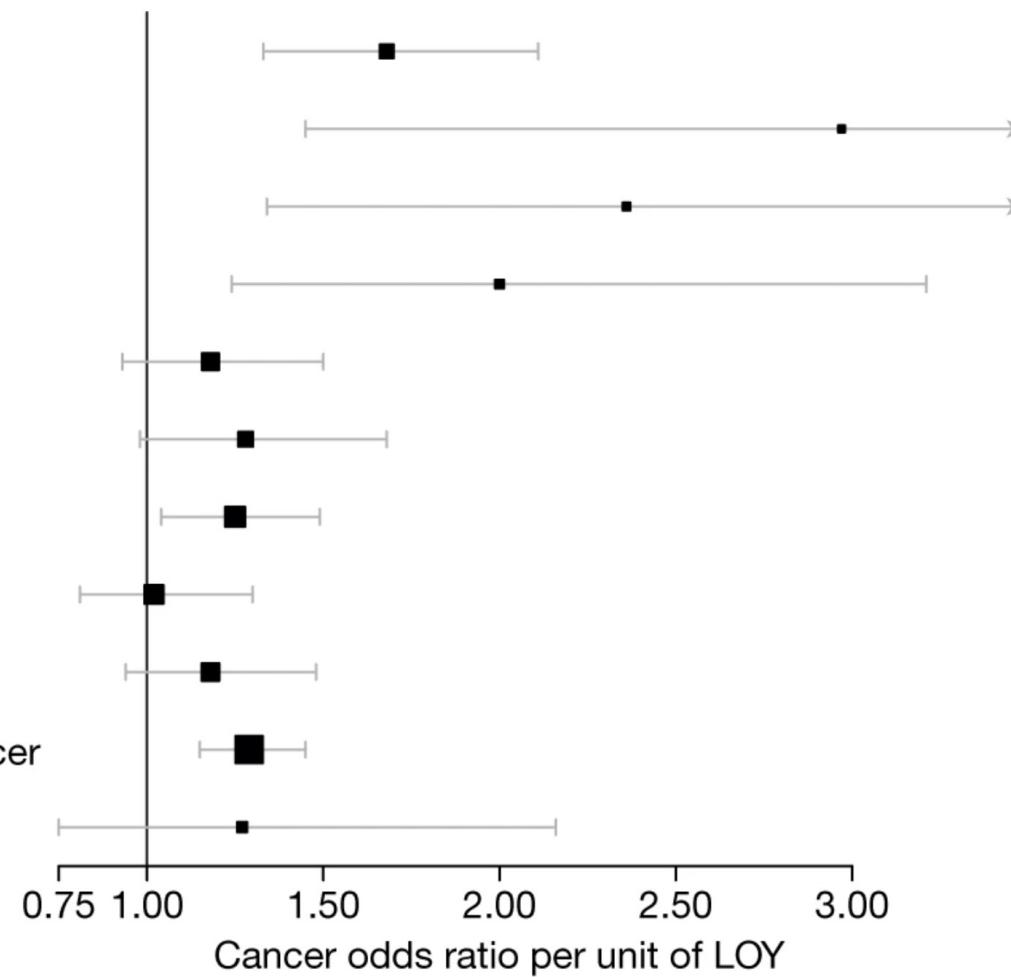
Group site	Incident cases N	Percent Y loss		mLOY (LRR \leq 0.15)		mLOY (LRR \leq 0.40)		SD of continuous LRR ^b	
		LRR \leq 0.15	LRR \leq 0.40	HR (95% CI)	P	HR (95% CI)	P	HR (95% CI)	P
All solid tumors cancer	13,895	3.0%	0.6%	1.18 (1.07-1.30)	0.001	1.36 (1.09-1.71)	0.007	1.04 (1.02-1.05)	7.5×10^{-7}
Oral cavity and pharynx	281	4.3%	1.1%	1.77 (0.98-3.20)	0.061	2.32 (0.73-7.36)	0.153	1.07 (0.97-1.17)	0.164
Esophagus	298	3.7%	0.3%	1.26 (0.68-2.31)	0.465	0.74 (0.10-5.29)	0.764	1.03 (0.94-1.13)	0.562
Stomach	190	3.7%	1.6%	1.30 (0.61-2.80)	0.500	3.58 (1.13-11.34)	0.030	1.11 (1.01-1.22)	0.026
Colon excluding rectum	848	2.5%	0.4%	1.00 (0.65-1.55)	0.992	0.92 (0.30-2.87)	0.886	1.03 (0.97-1.09)	0.335
Rectum & rectosigmoid junction	550	2.4%	0.9%	0.95 (0.54-1.65)	0.846	2.37 (0.98-5.75)	0.057	1.02 (0.95-1.10)	0.598
Liver & intrahepatic bile duct	112	1.8%	0.9%	0.63 (0.15-2.57)	0.518	1.97 (0.27-14.31)	0.503	0.98 (0.83-1.16)	0.788
Pancreas	229	4.8%	0.4%	1.59 (0.86-2.95)	0.140	0.87 (0.12-6.21)	0.866	1.08 (0.99-1.18)	0.089
Lung	783	5.6%	2.0%	1.11 (0.82-1.52)	0.498	2.25 (1.36-3.71)	0.002	1.06 (1.02-1.11)	0.003
Skin excluding basal/squamous	4,519	2.6%	0.4%	1.16 (0.97-1.40)	0.109	1.29 (0.83-2.01)	0.257	1.04 (1.01-1.06)	0.007
Prostate	4,345	2.8%	0.4%	1.13 (0.94-1.36)	0.195	0.98 (0.60-1.60)	0.928	1.03 (1.00-1.05)	0.050
Bladder	369	6.0%	0.8%	1.86 (1.20-2.90)	0.006	1.50 (0.48-4.70)	0.485	1.06 (0.99-1.14)	0.087
Brain & other nervous system	207	1.4%	0.0%	0.70 (0.22-2.21)	0.543	—	—	1.04 (0.92-1.17)	0.587

^aAdjusted for age using age as the underlying time metric in the Cox proportional hazards regression model as well as the following covariates: detailed smoking history [25-level variable incorporating current smoking status, smoking intensity (current and former smokers); time since quitting (former smokers), and cigar and pipe use (current and former smokers), race/ethnicity (white, black, Asian, mixed, or other race)]; alcohol drinking (never drinker, former drinker, infrequent drinker (<1 drink/week), occasional drinker (>1 drink/week but <1 drink/day), moderate daily drinker (1-3 drinks/day), or heavy daily drinker (>3 drinks/day); education level (college or university degree, A levels/AS levels or equivalent, O levels/GCSEs or equivalent, CSEs or equivalent, NVQ or HND or HNC equivalent, or other professional qualifications); and BMI (<18.5, 18.5-<25, 25-<30, 30-<35, or \geq 35 kg/m²).

^bScaled by the – (SD) of mLRR, such that the HR corresponds to a one SD decrease in mLRR.

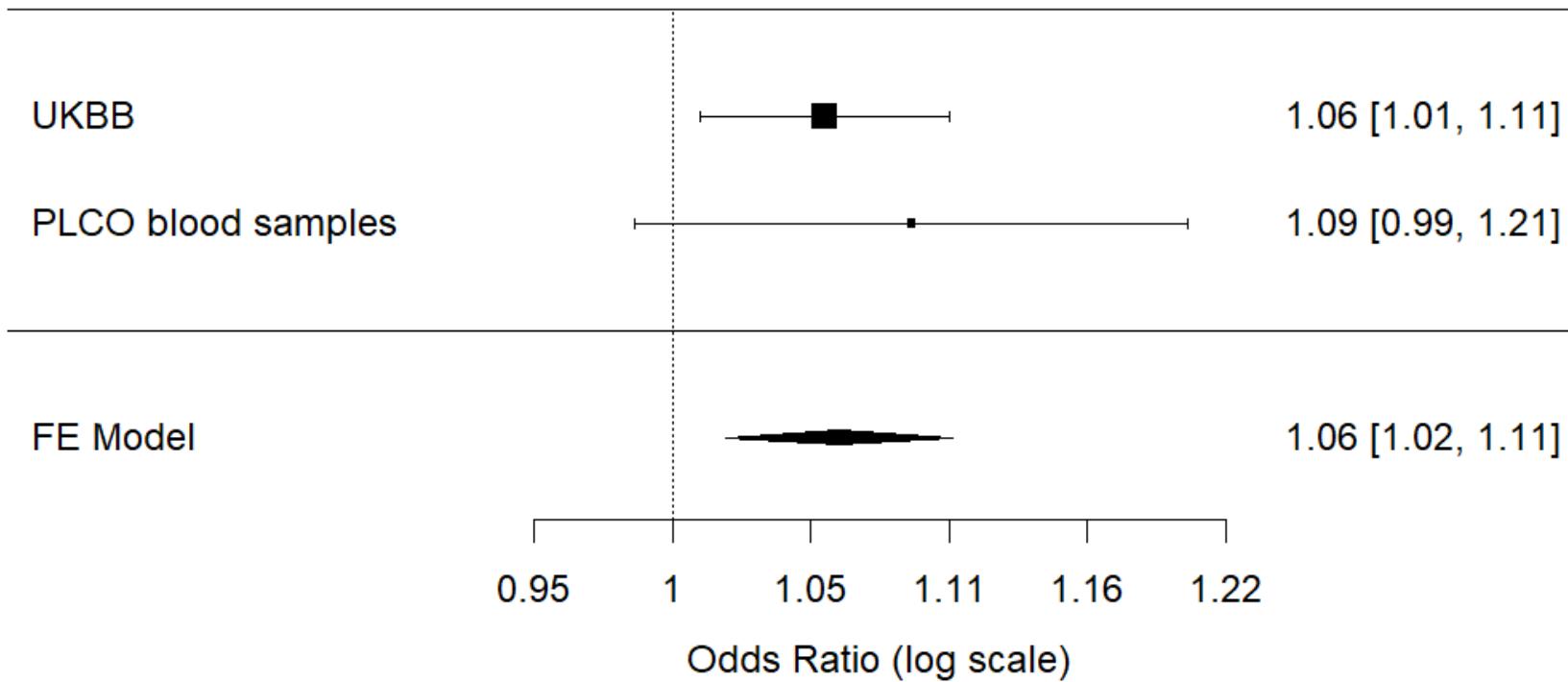
Y Loss Variants Associated with Increased Cancer Risk

Prostate cancer
Testicular germ cell tumour
Glioma
Renal cell carcinoma
Colorectal cancer
Lung cancer
Breast cancer
Ovarian cancer
Endometrial cancer
Breast, prostate, ovarian and endometrial cancer
Chronic lymphocytic leukaemia



The genetic risk score comprises the 156 LOY-associated loci identified in the UK Biobank discovery analysis ($n = 205,011$). Squares are proportional to the sample size. Error bars denote 95% confidence intervals around the point estimate OR effect, with the value 1 denoting no effect.

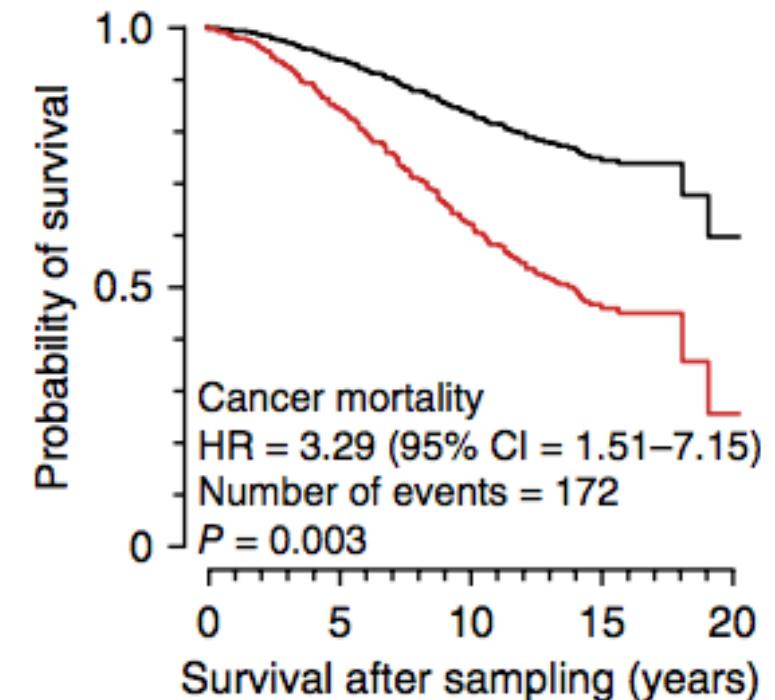
mLOY and Prostate Cancer



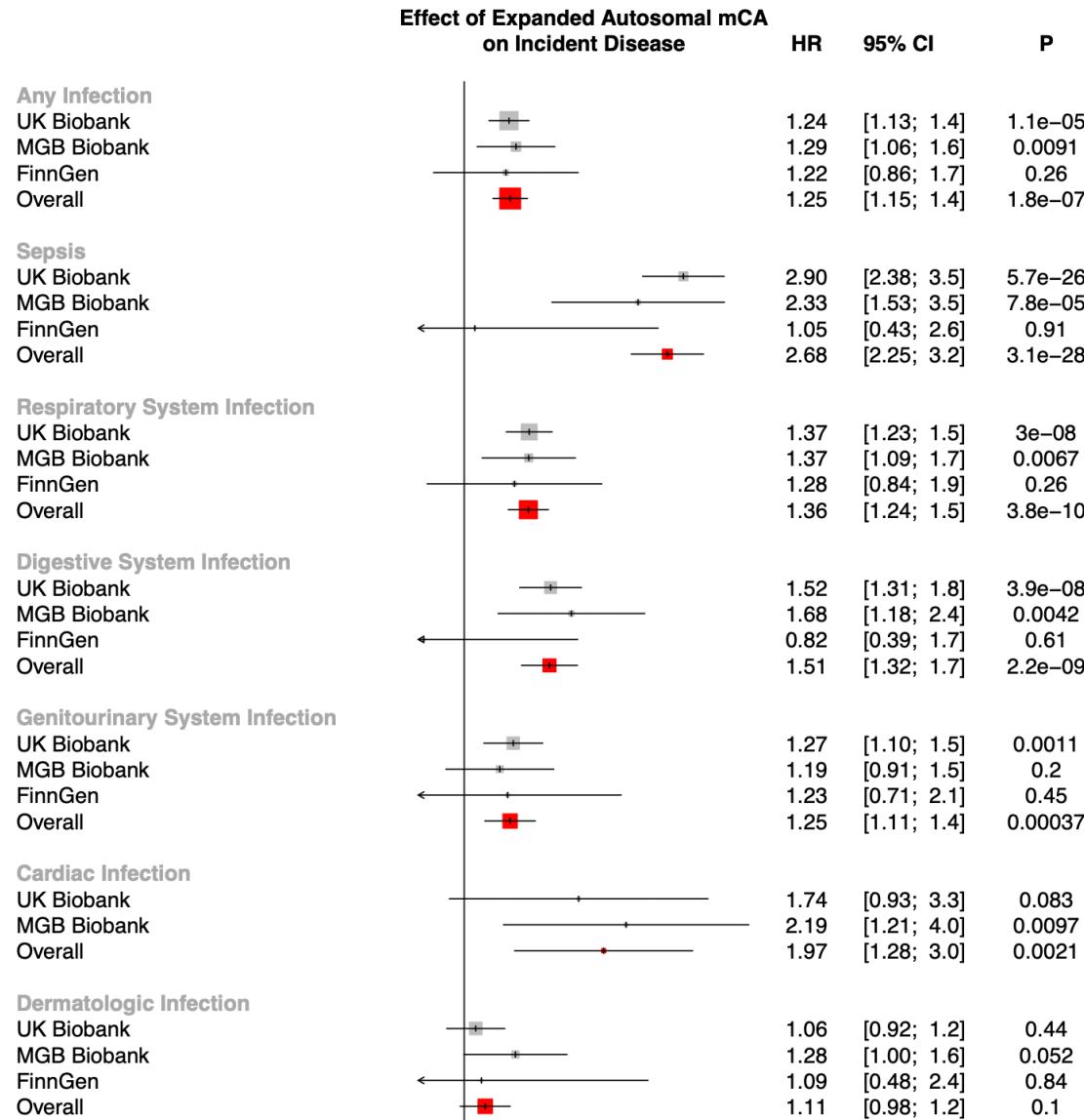
FE meta-analysis – incident PCa	total n	odds ratio	95% CI	p-value
mLOY	236,898	1.062	1.020,1.107	0.00398

Mosaic Y Loss and Mortality

	All-cause mortality			Cancer mortality		
	(Number of events = 637)			(Number of events = 172)		
	HR	95% CI	P value	HR	95% CI	P value
Genotyping age	1.09	1.06–1.13	<0.0001*	1.06	1.00–1.12	0.039*
Hypertension	1.68	1.35–2.08	<0.0001*	1.30	0.88–1.93	0.191
Exercise habits	0.49	0.32–0.74	<0.0001*	1.12	0.35–3.53	0.852
Smoking	1.38	1.14–1.68	<0.0001*	1.37	0.94–2.00	0.105
Diabetes	1.43	1.09–1.86	0.010*	0.98	0.55–1.76	0.949
BMI	0.99	0.96–1.02	0.425	1.01	0.96–1.07	0.674
LDL cholesterol	0.97	0.88–1.07	0.501	0.95	0.79–1.16	0.625
HDL cholesterol	0.89	0.69–1.15	0.359	0.64	0.38–1.09	0.099
Education level	0.95	0.88–1.02	0.151	0.98	0.85–1.13	0.784
Autosomal gain (>2 Mb)	0.94	0.42–2.12	0.880	1.41	0.34–5.82	0.635
Autosomal LOH (>2 Mb)	1.41	0.86–2.32	0.172	1.19	0.43–3.26	0.742
LOY	2.13	1.08–4.19	0.029*	3.76	1.21–11.67	0.022*

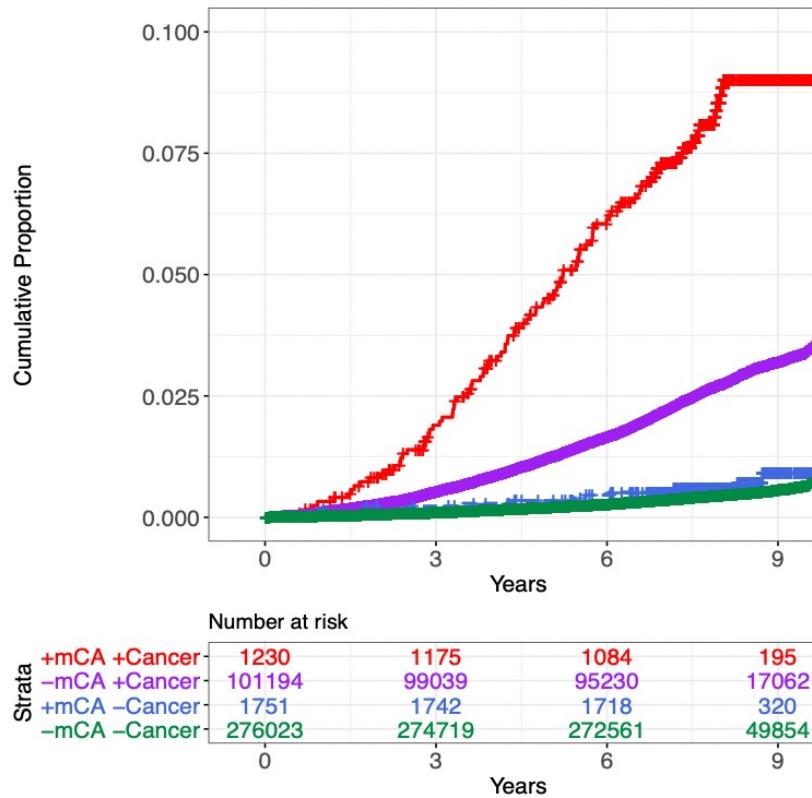


Autosomal Mosaicism and Risk of Infections

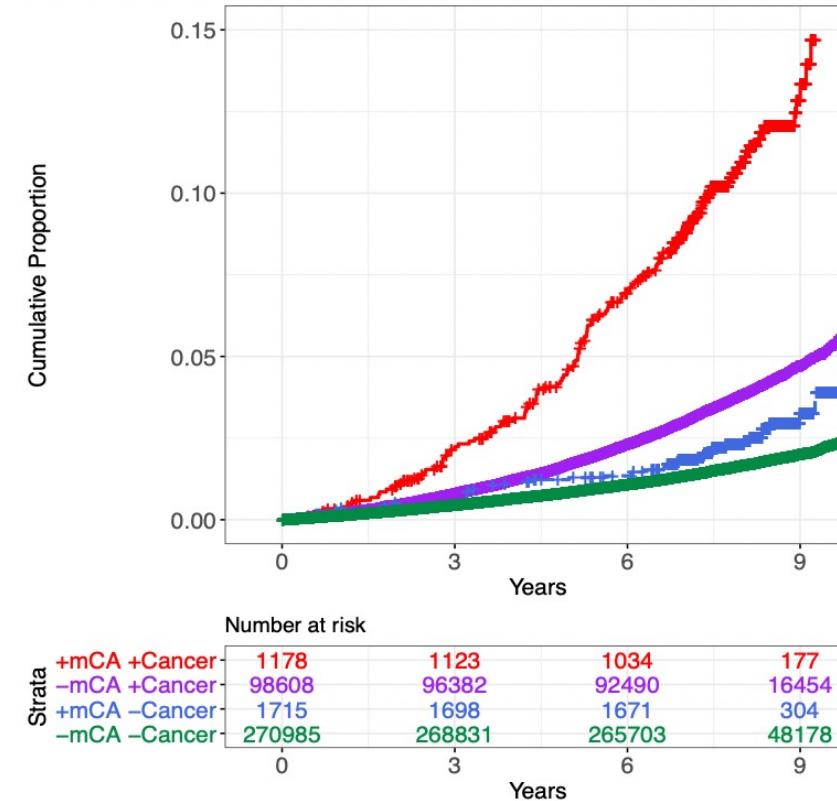


Infection Risk Stronger for Cancer Patients

Sepsis



Pneumonia



Thinking Causally

Clonal Mosaicism and Cancer (or other outcome)

(1) Risk Factor

Mosaicism → Cancer

(2) Common Cause

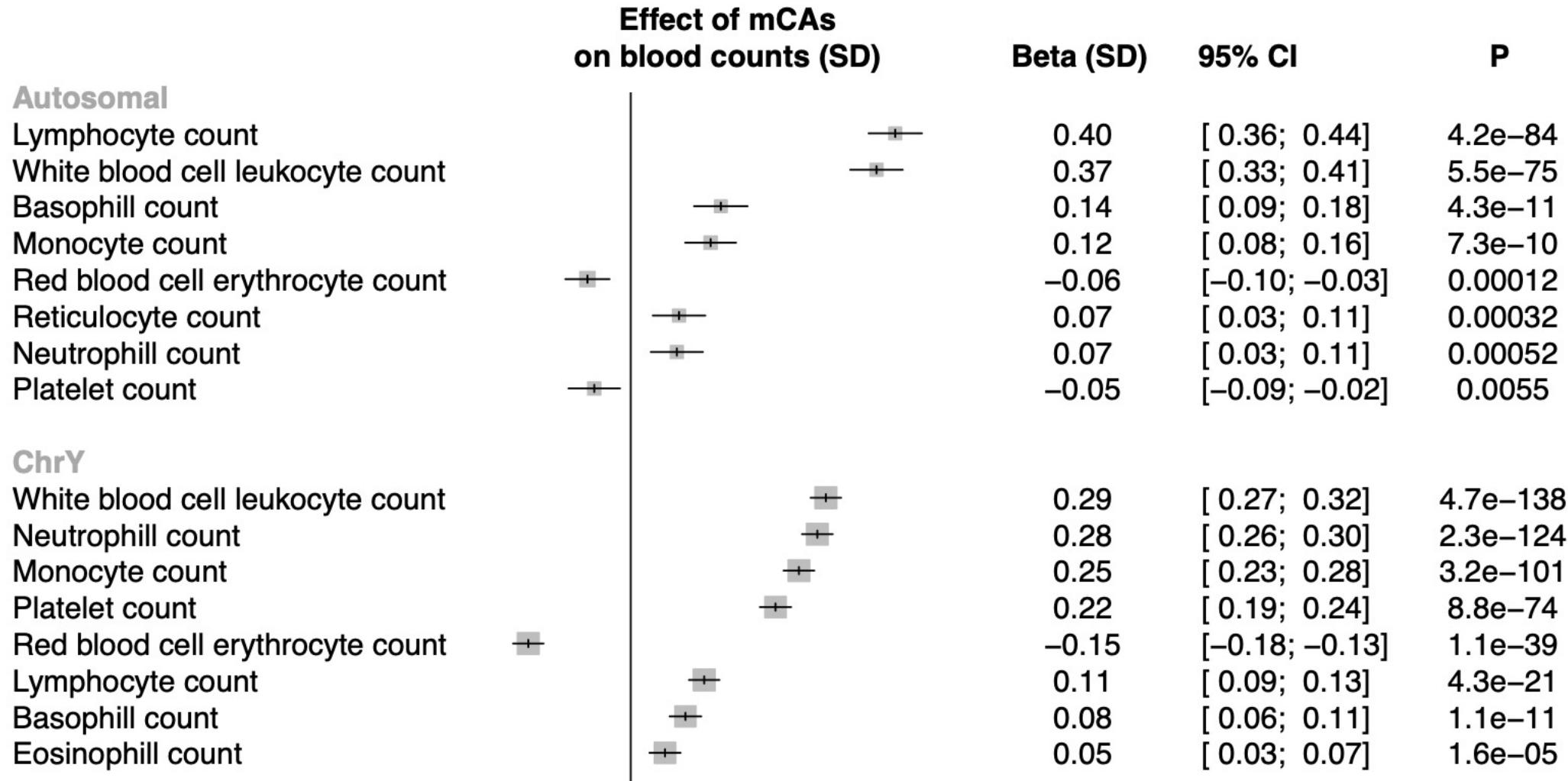


e.g., X = residual effect of age or smoking

(3) Reverse Causation

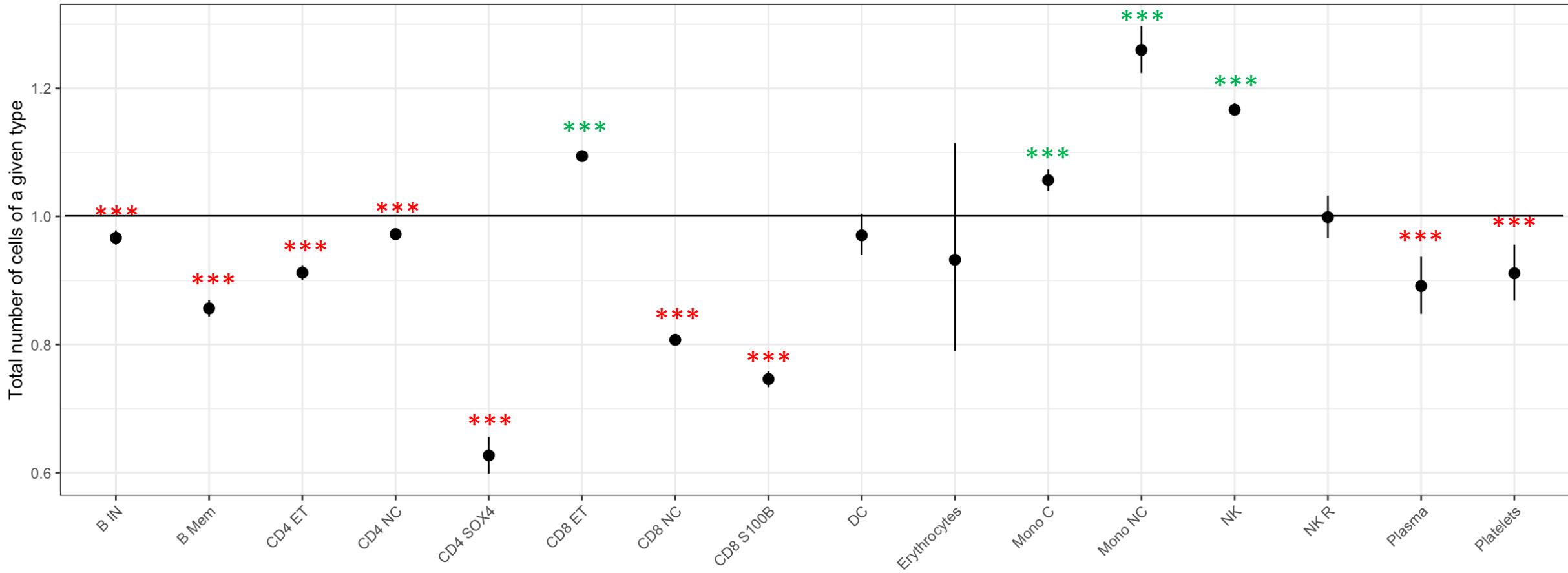
Cancer → Mosaicism

Clonal Mosaicism and Blood Cell Counts

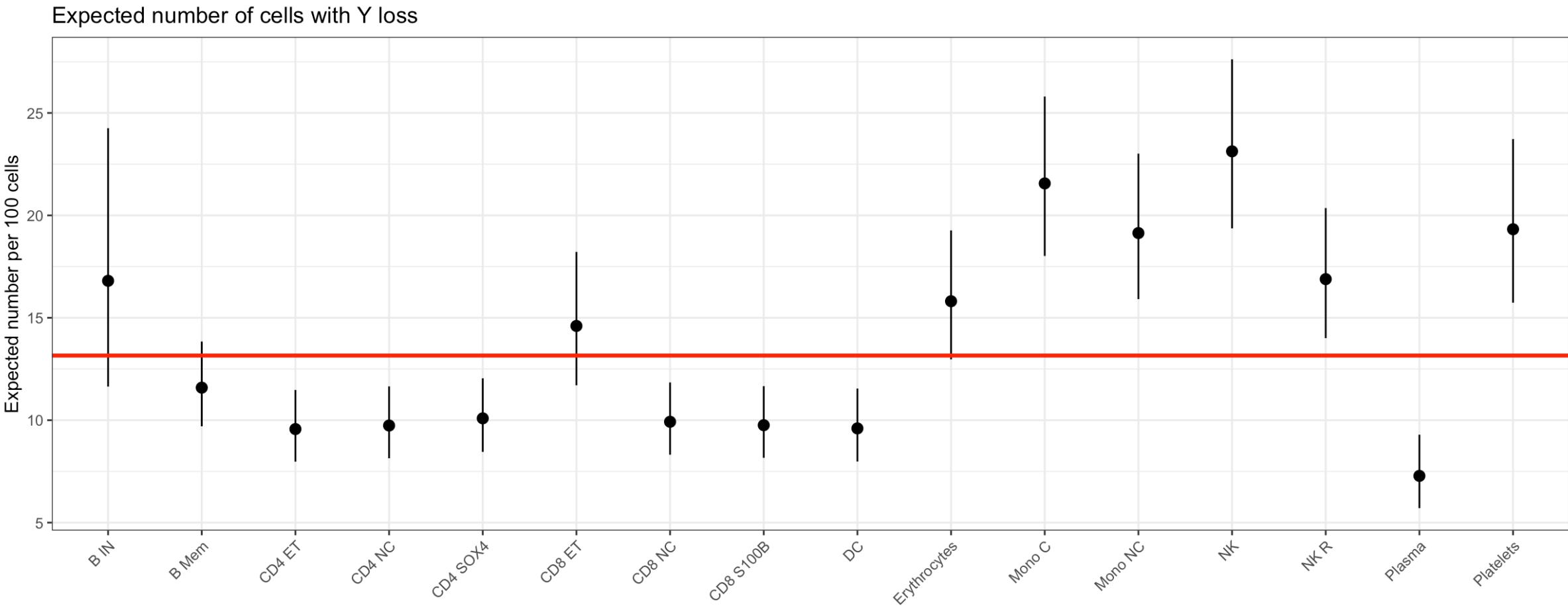


Single Cell RNA Sequencing Data from the OneK1K

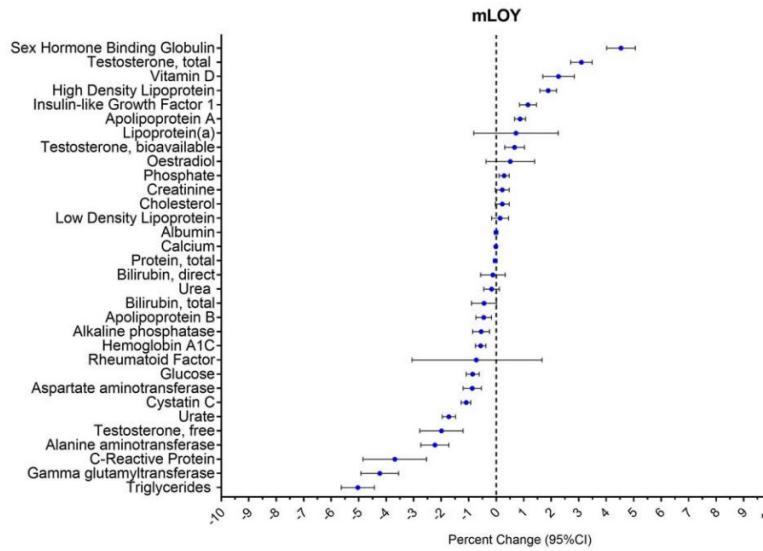
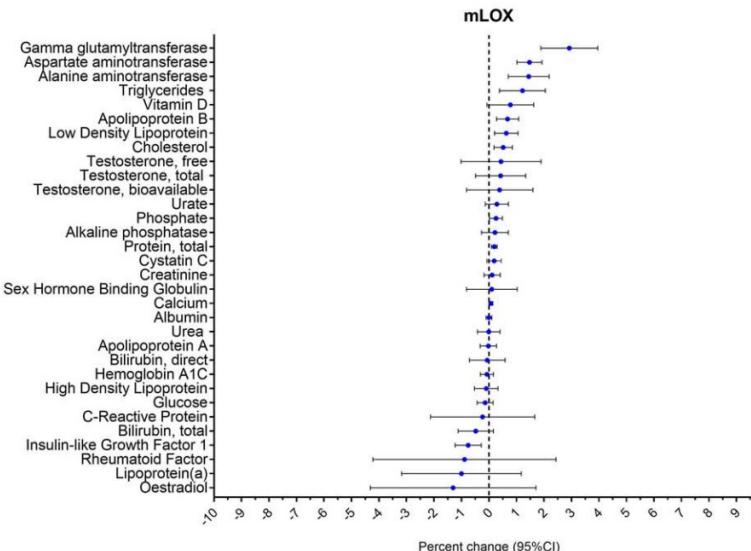
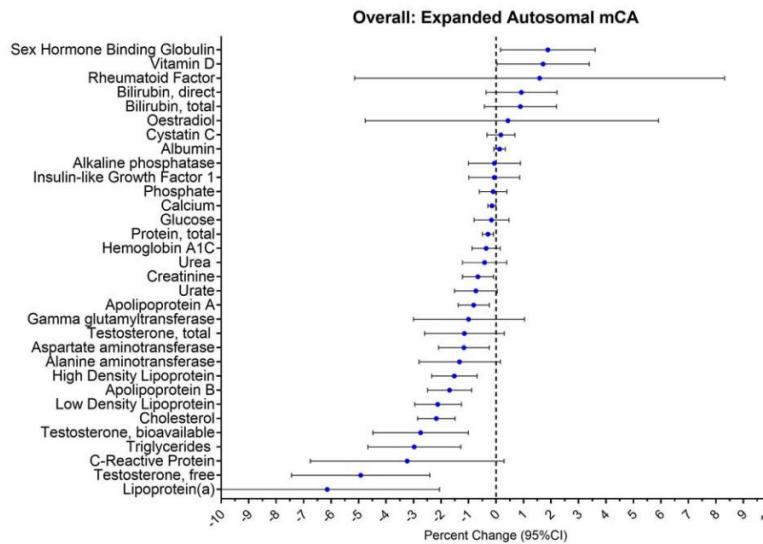
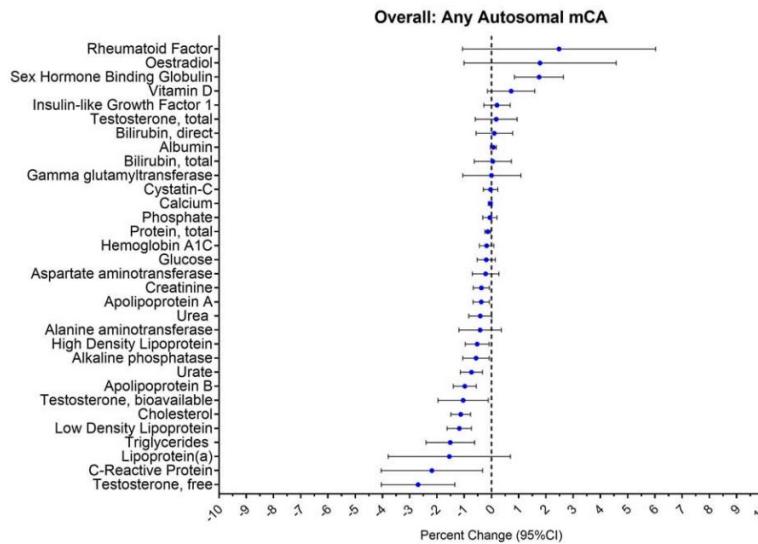
Total number of cells of a given type: `glm(n ~ age family=poisson,data = input_data)`



mLOY Frequency Differs by Cell Type



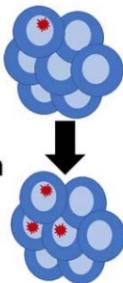
Clonal Mosaicism and Serum Biomarkers



Serum
Biomarkers



Clonal
Expansion



Discussion

- Higher frequency of CH in healthy populations
 - male X << autosomes < female Xi << Y loss
- The frequency of individuals with CH increases with age (regardless of event type)
- Dynamic process – the proportion of cells affected changes over time but generally increases with increasing age
- Smoking is associated with increased risk of CH, evidence for other exposures as well
- Germline susceptibility loci have been identified associated with different types of CH
 - Shared etiology as well as key differences
- CH is associated with hematologic malignancies – both lymphoid and myeloid
 - Not all individuals with CH develop a hematologic malignancy
- CH is associated with select solid tumors as well as infection risk
 - Ongoing studies to disentangle whether associations are causal
- Initial evidence for biologic impact of CH (e.g., changes in blood cell counts), but mechanisms for effects outside of hematologic compartment are unclear

Collaborators



FinnGen

Andrea Ganna, Aoxing Liu, Timo Petteri Sipilä, Awaisa Ghazal, Mervi Aavikko, Mark Daly, Aarno Palotie



UK Biobank (UKBB)

John R.B. Perry, Yajie Zhao, Giulio Genovese, Maryam Zekavat



Estonian Biobank (EBB)

Reedik Mägi, Georgi Hudjashov, Vlad Tuzov, Fanny-Dhelia Pajuste



Biobank Japan (BBJ)

Chikashi Terao, Azusa Kouno



Breast Cancer Association Consortium (BCAC)

Joe Dennis



Million Veteran Program (MVP)

Bryan Gorman, Maryam Zekavat



Mass General Brigham Biobank (MGB)

Giulio Genovese, Maryam Zekavat



Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial (PLCO)

Weiyin Zhou, Derek Brown, Wen-Yi Huang, Neal Freedman

Core Analysis

Aoxing Liu (Meta-analysis & FinnGen specific analyses)
Giulio Genovese (MoChA pipeline)
Yajie Zhao (WES & GWAS in UKB)
Katherine Kentistou (GWAS to genes)
Matti Pirinen (Bayesian line model to cluster mLOX and mLOY loci)
Derek Brown (heat-map)
Kai Yu (pathway analysis)
Zhiyu Yang (genetic correlation)
Po-Ru Loh (chromosome X analysis)

Leadership

Aoxing Liu, Giulio Genovese, Yajie Zhao, Chikashi Terao, Po-Ru Loh, Andrea Ganna, John R.B. Perry



Interested in Genetic Mosaicism?



Join the DCEG Genetic Mosaicism
Working Group
for information email
mitchell.machiela@nih.gov



mitchell.machiela@nih.gov



Interested in genetic epidemiology of cancer? Research and training opportunities available at DCEG.

<https://dceg.cancer.gov/fellowship-training>