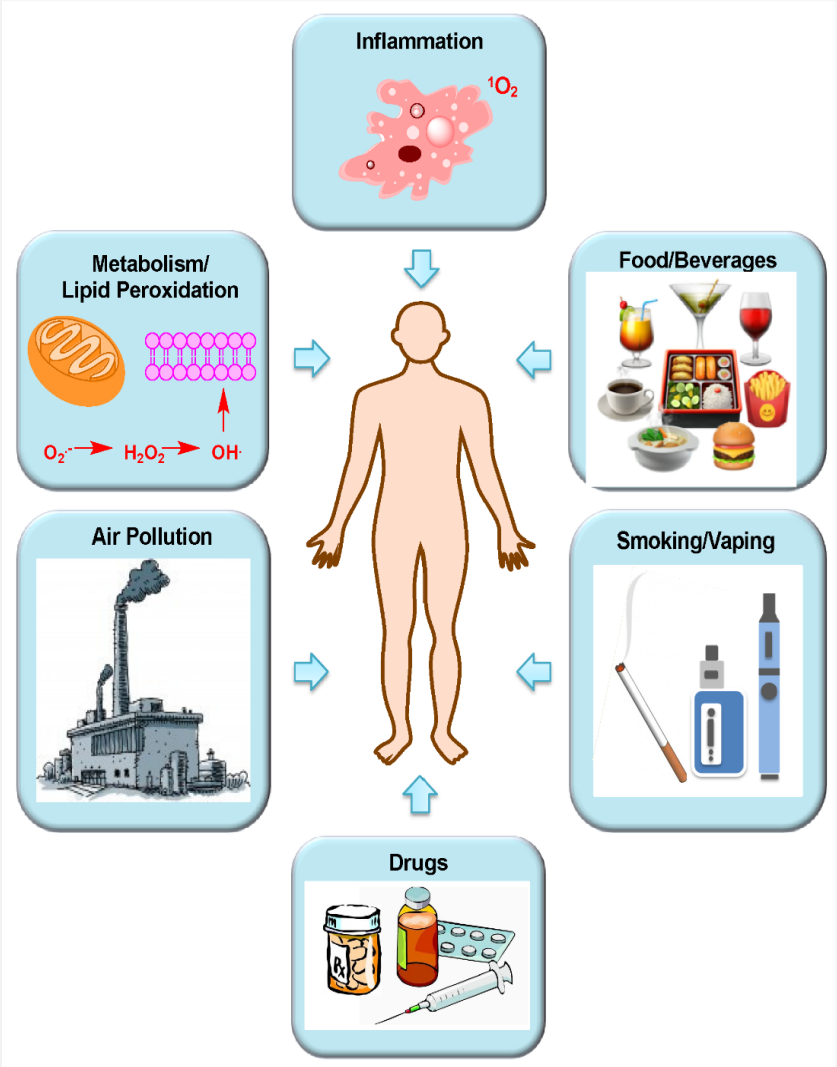
The characterization of mutational signatures of Acetaldehyde and Formaldehyde to investigate their role in carcinogenesis

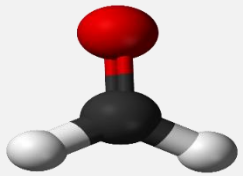
Mahanish Jung Thapa
PhD candidate
Supervisor
Dr. Kin Chan

Ottawa Conference
April 12, 2023

Aldehydes and relevance

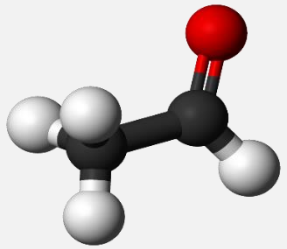


Sources of acetaldehyde and formaldehyde
Dator et al., 2019



Formaldehyde (FA)

HCHO



Acetaldehyde (AA)

CH_3CHO

International Agency for Research on Cancer

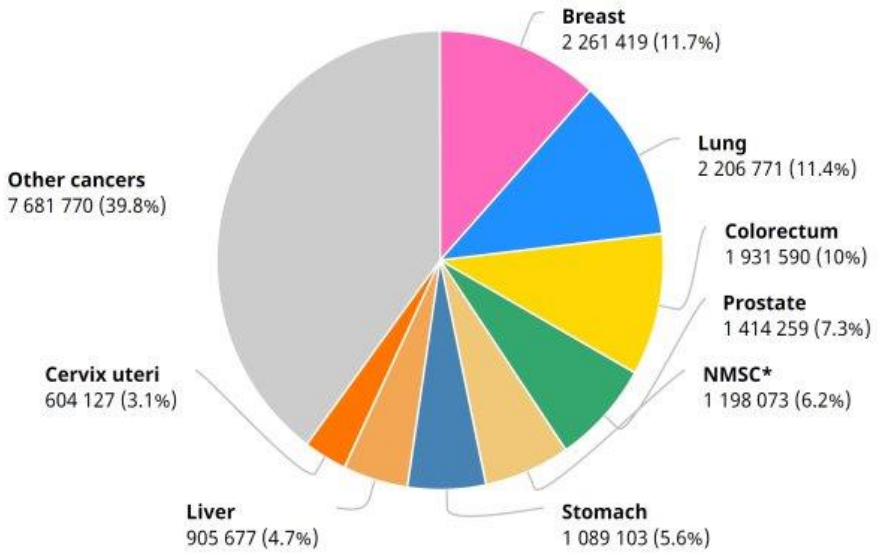
World Health Organization

Chemical	Group	Remarks
Formaldehyde	Group I	Carcinogenic
Acetaldehyde	Group 2B	Possible carcinogenic to humans
Acetaldehyde from ethanol consumption	Group I	Carcinogenic

IARC Monographs 71, 1999

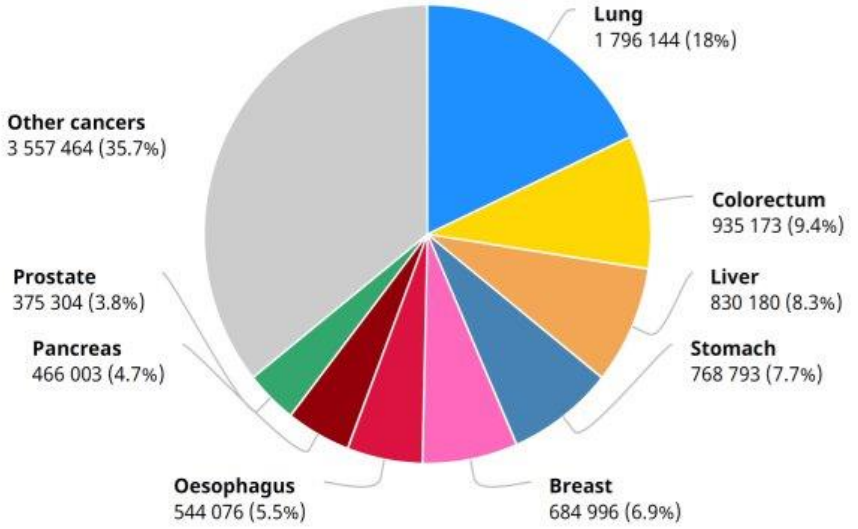
All cancers

Number of new cases in 2020, both sexes, all ages



Total: 19 292 789 cases

Number of deaths in 2020, both sexes, all ages



Total: 9 958 133 deaths

International Agency for Research on Cancer
World Health Organization

GLOBAL CANCER OBSERVATORY

#GCO365

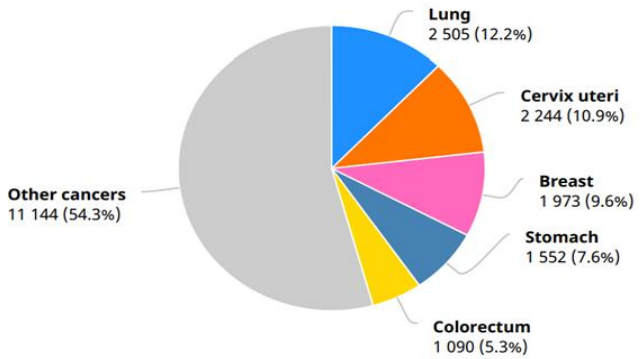
IARC

International Agency for Research on Cancer
World Health Organization

Nepal

Source: Globocan 2020

Number of new cases in 2020, both sexes, all ages



Total: 20 508

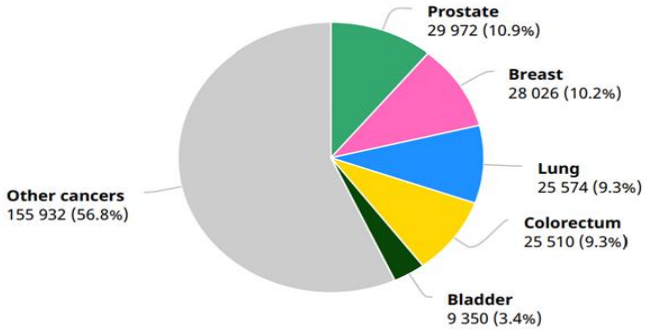
Geography

International Agency for Research on Cancer
World Health Organization

Canada

Source: Globocan 2020

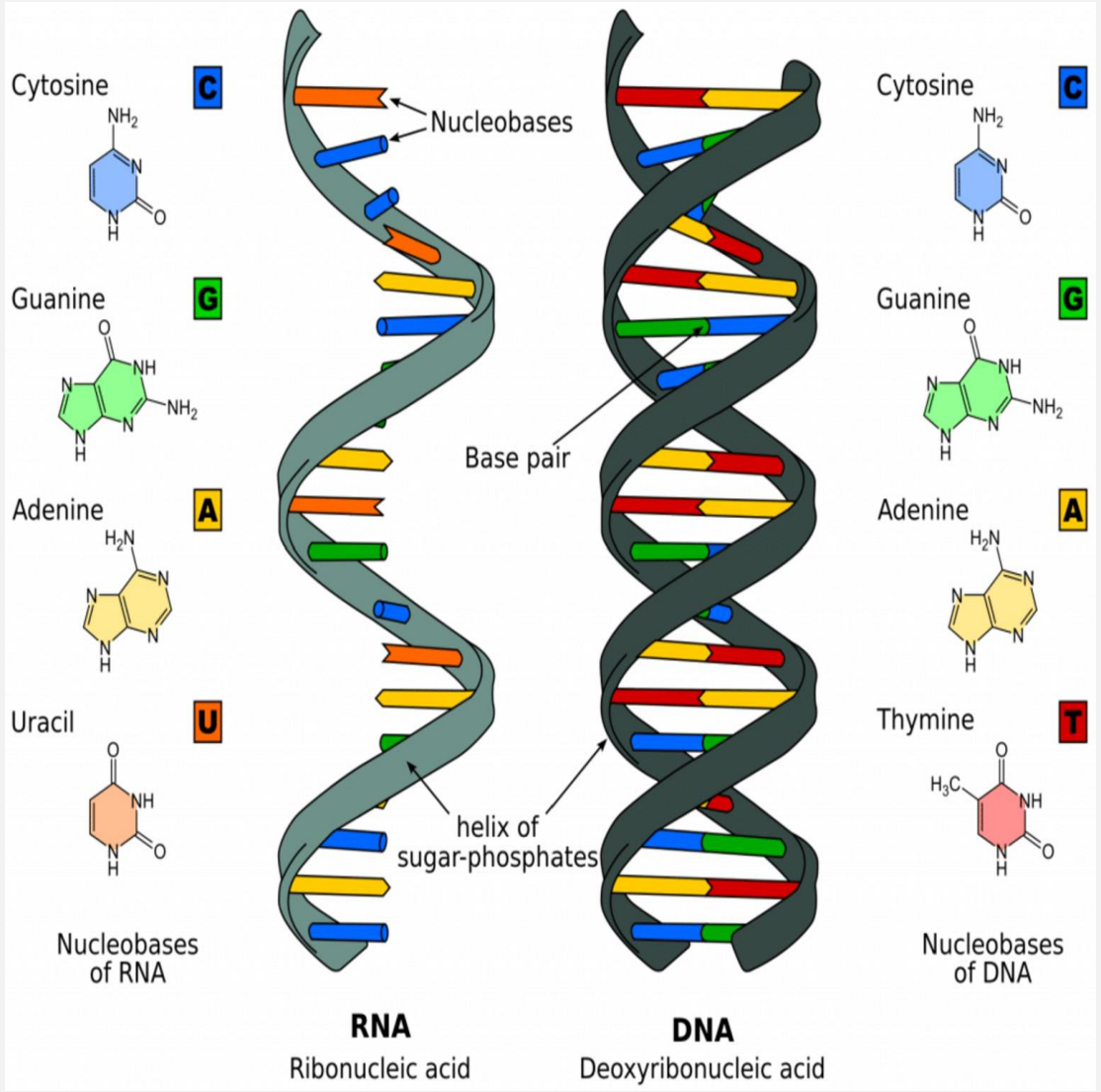
Number of new cases in 2020, both sexes, all ages



Total: 274 364

Geography

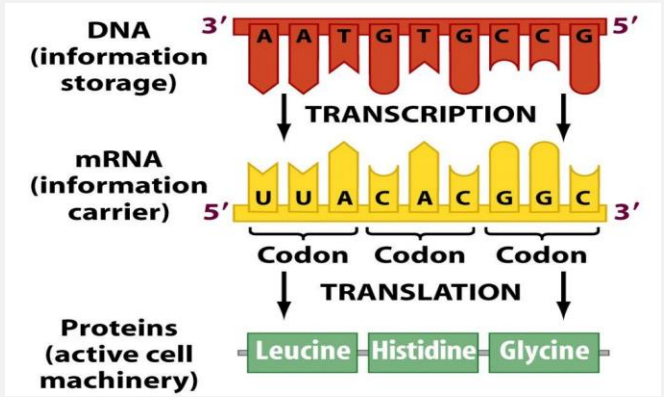
DNA, RNA, amino acid, and proteins



https://teachmephysiology.com/biochemistry/cell-growth-death/dna-replication/

Triplet codon table

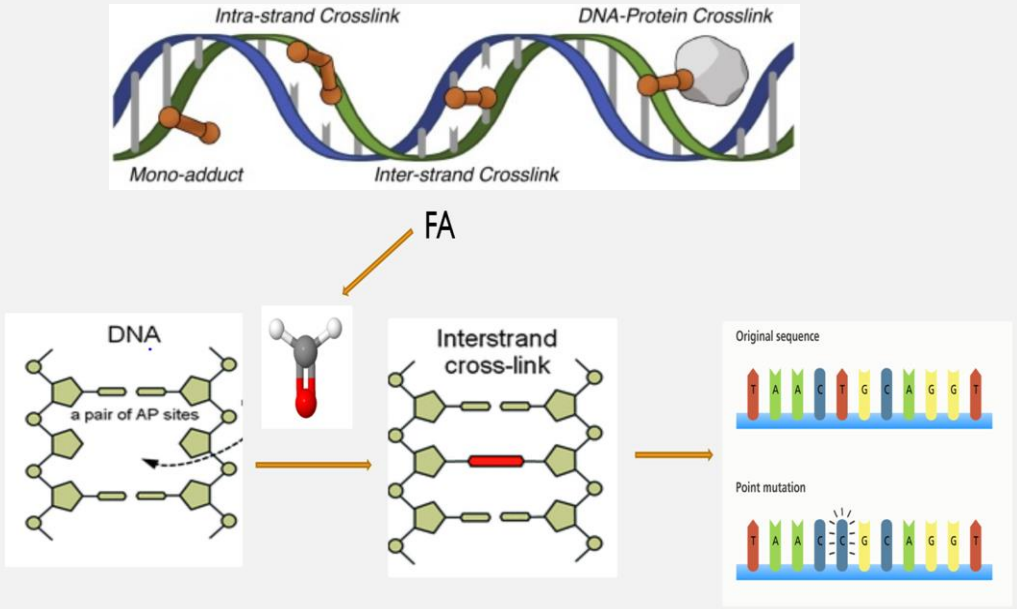
1st base	2nd base T		2nd base C		2nd base A		2nd base G		3rd base
T	TTT	Phe (F)	TCT	Ser (S)	TAT	Tyr (Y)	TGT	Cys (C)	T
	TTC		TCC		TAC		TGC		C
	TTA		TCA		TAA	STOP	TGA	STOP	A
	TTG		TCG		TAG		TGG	Trp (W)	G
C	CTT	Leu (L)	CCT	Pro (P)	CAT	His (H)	CGT	Arg (R)	T
	CTC		CCC		CAC		CGC		C
	CTA		CCA		CAA	Gln (Q)	CGA		A
	CTG		CCG		CAG		CGG		G
A	ATT	Ile (I)	ACT	Thr (T)	AAT	Asn (N)	AGT	Ser (S)	T
	ATC		ACC		AAC		AGC		C
	ATA		ACA		AAA	Lys (K)	AGA	Arg (R)	A
	ATG	Met (M)	ACG		AAG		AGG		G
G	GTT	Val (V)	GCT	Ala (A)	GAT	Asp (D)	GGT	Gly (G)	T
	GTC		GCC		GAC		GGC		C
	GTA		GCA		GAA	Glu (E)	GGA		A
	GTG		GCG		GAG		GGG		G

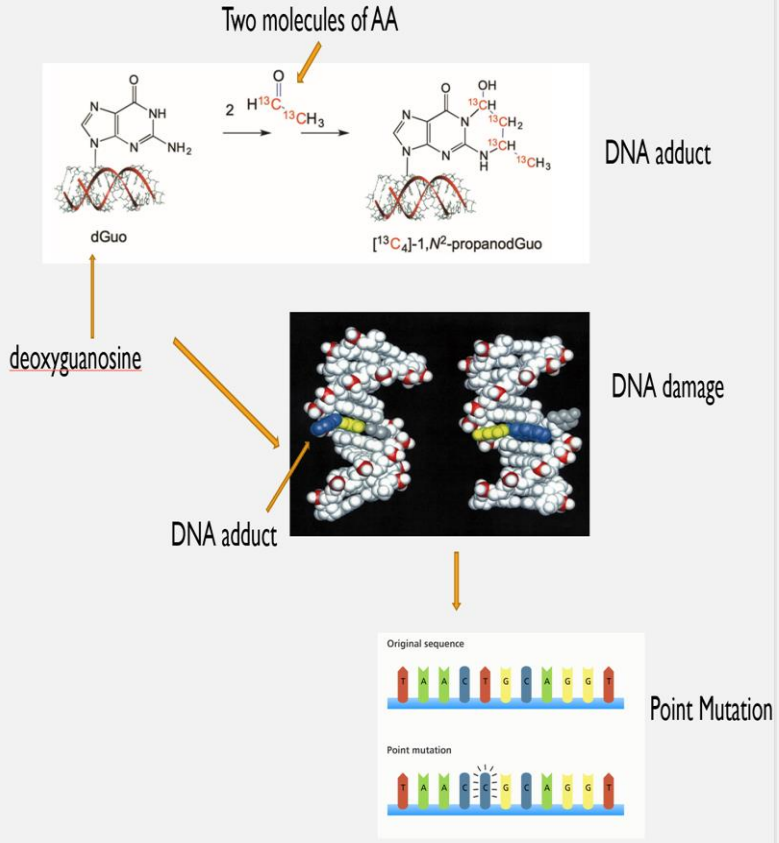


Amino acids to proteins

FA & AA DNA damage and mutation

Both formaldehyde and acetaldehyde causes different types of DNA damage- point mutations, DNA adducts, DNA single-strand breaks, DNA double-strand breaks, and DNA-DNA inter-strand crosslinks, DNA-DNA intra-strand crosslinks, and DNA-protein crosslinks.





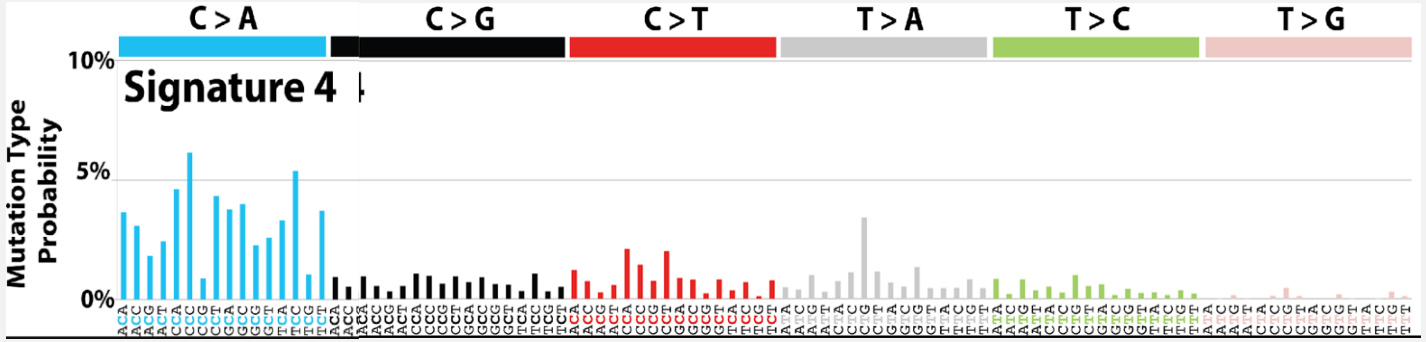
Kawanishi et al., 2014; Rycenga and Long , 2018

Ichikawa et al., 2012; Genome Research Limited
https://www.molinstincts.com/structure/formaldehyde-cstr-CT1001024308.html

Medeiros et al., 2019; Hingerty 2008; Genome Research Limited

Mutational Signatures

- Mutational Signatures are unique combinations of different somatic mutation types by different mutation processes.
- COSMIC (Catalogue of Somatic Mutations in Cancer)



It is associated with tobacco smoking. Its profile is similar to the mutational spectrum observed in experimental systems exposed to tobacco carcinogens such as benzo[a]pyrene.

COSMIC, Alexandrov et al., 2016

Hypothesis

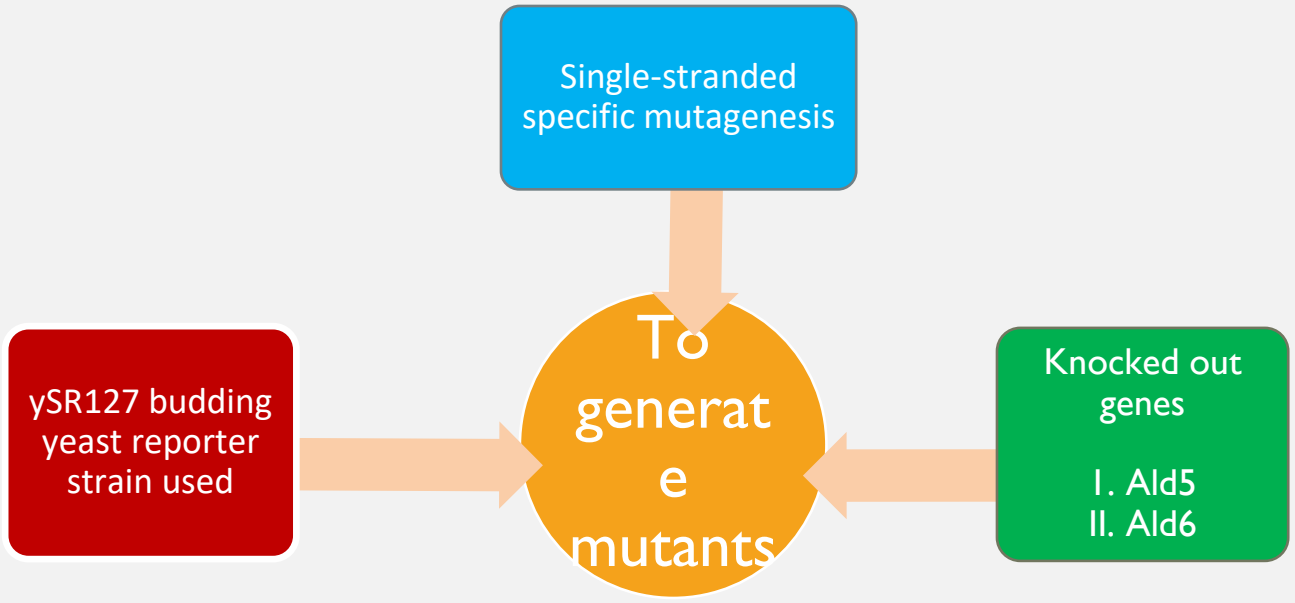
My hypothesis is that the Acetaldehyde (AA) creates many mutations in the DNA of our yeast model system, as previously shown by my lab mate Reena Fabros for Formaldehyde (FA).

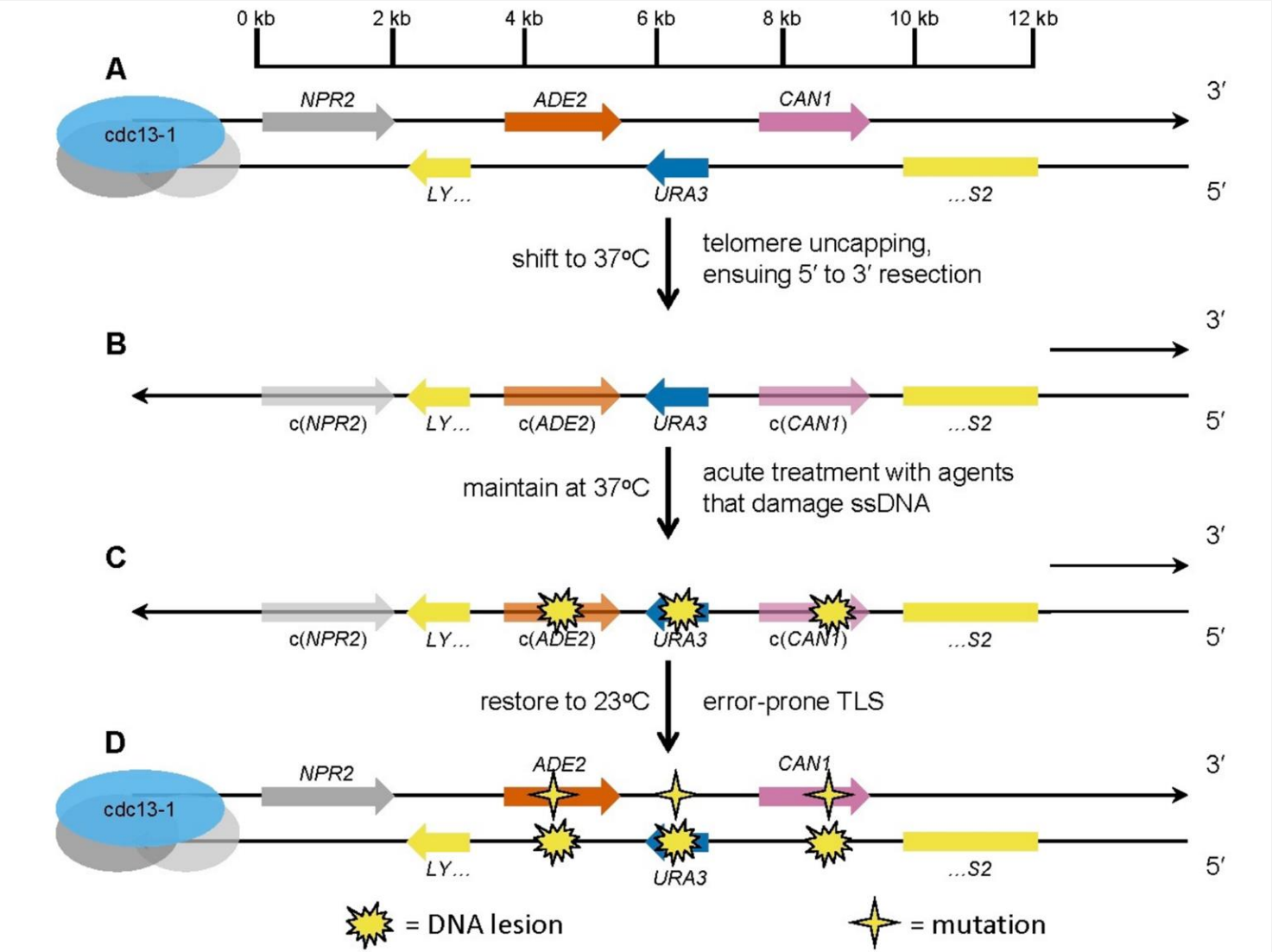
The mutational signatures produced by both AA and FA induced mutations can be characterized which may show similarities to the cancer related mutational signatures.

Research Objective I

- To generate yeast mutants through the exogenous exposure of AA and investigate the mutagenic effect of AA.

- Knocked out genes
  I. Ald5 : Mitochondrial aldehyde dehydrogenase
  II. Ald6 : Cytosolic aldehyde dehydrogenase





Schematic Diagram of our Yeast Model System

Chan et al., PLoS Genetics 2012

Experimental results summary

CELL VIABILITY(%)

Conc / Strain	0 mM	25 mM	50 mM	75 mM	100 mM
Wild type	96	92	85	42	20
ald5Δ	94	77	53	9	
ald6Δ	92	79	47	10	
ald5Δ/ ald6Δ	80	40	9	2	

X-axis=Conc=Concentration
Y-axis =Yeast strain
The cell viability values are approximated

Lowest viability even at 0 mM

MUTATIONAL FREQUENCY

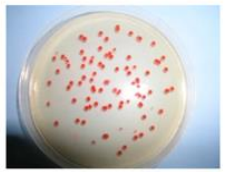
Conc / Strain	0 mM	25 mM	50 mM	75 mM	100 mM
Wild type	$0.2 * 10^{-4}$	$1 * 10^{-4}$	$2 * 10^{-4}$	$5.1 * 10^{-4}$	$0.23 * 10^{-4}$
ald5Δ	$0.25 * 10^{-4}$	$1.4 * 10^{-4}$	$5.2 * 10^{-4}$		
ald6Δ	$0.30 * 10^{-4}$	$2 * 10^{-4}$	$5.5 * 10^{-4}$		

X-axis=Conc=Concentration
Y-axis =Yeast strain
The mutational frequency values are approximated

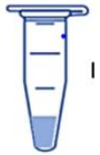
Highest mutational frequency for each strain

Lowest mutational frequency

Research Objective 2



Successful candidates are replicated and tested in different media



DNA extraction

Whole genome sequencing to generate paired end reads

Mapping of sequenced reads are done to reference genome ySR127 in local and global alignment using Bowtie2 and SAM tools and visualized through Integrative genomics viewer

To computationally characterize the mutational signature created by AA using yeast genomics data.

command-not-found.com

bcftools

genomic variant calling and manipulation of VCF/BCF files

VCF

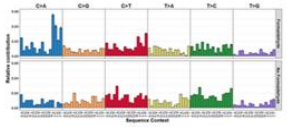
Generates final mutation file- Variant Call Format file (VCF)

Variant call format file

Phred Quality score

Redundancy

Used R package called MutationalPatterns



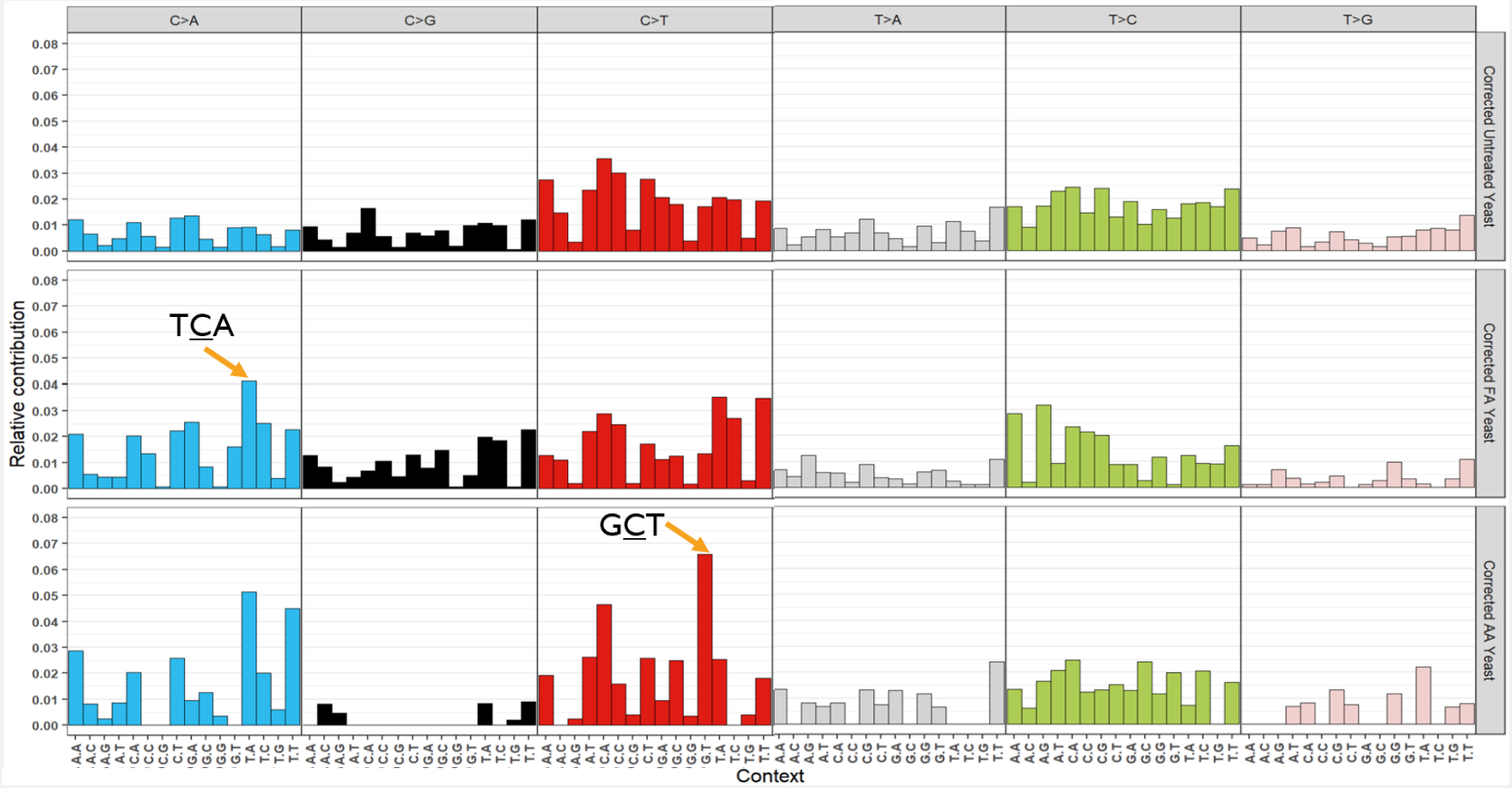
Mutational signature

CORRECTING TRINUCLEOTIDE FREQUENCY

- The trinucleotide frequency of reference ySR127 yeast genome is corrected to match with trinucleotide frequency of human genome.

- To match FA and AA signatures with COSMIC signatures and these COSMIC signatures are found in different human cancer types.

Single base substitutions

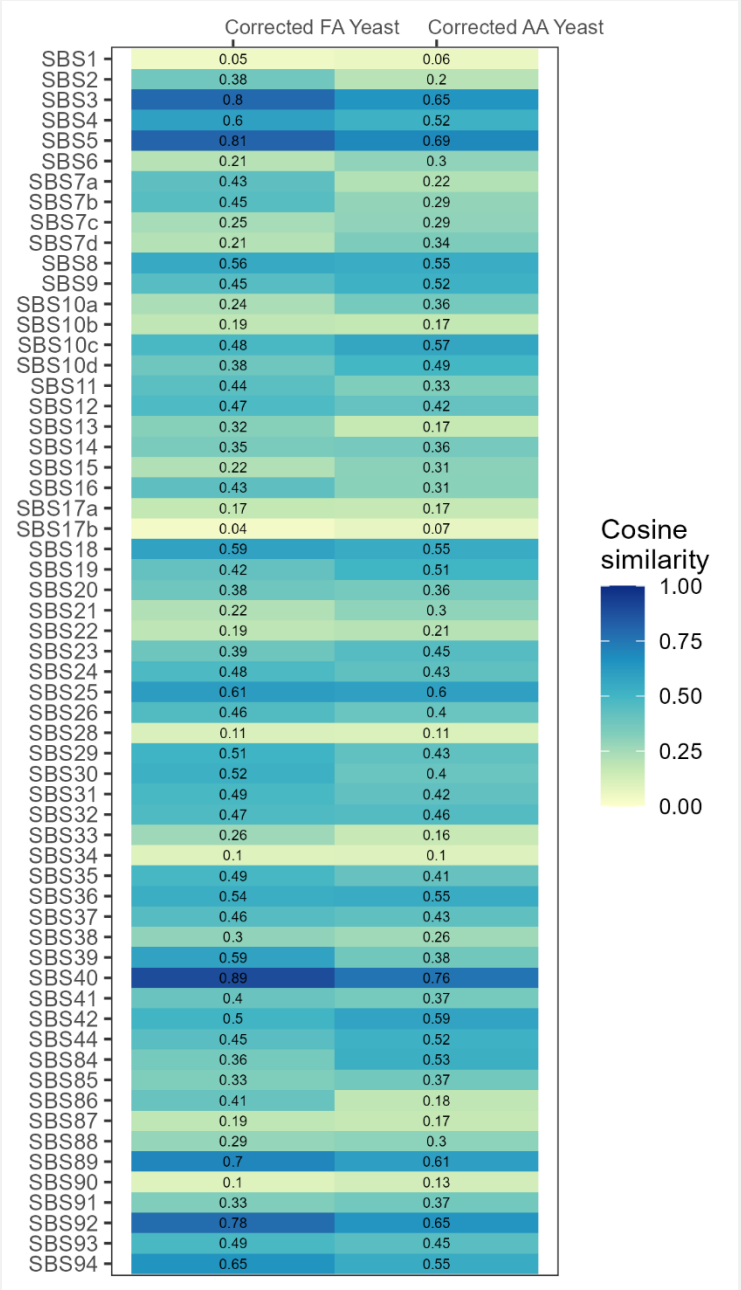


Thapa et al., 2022

Cosine similarity

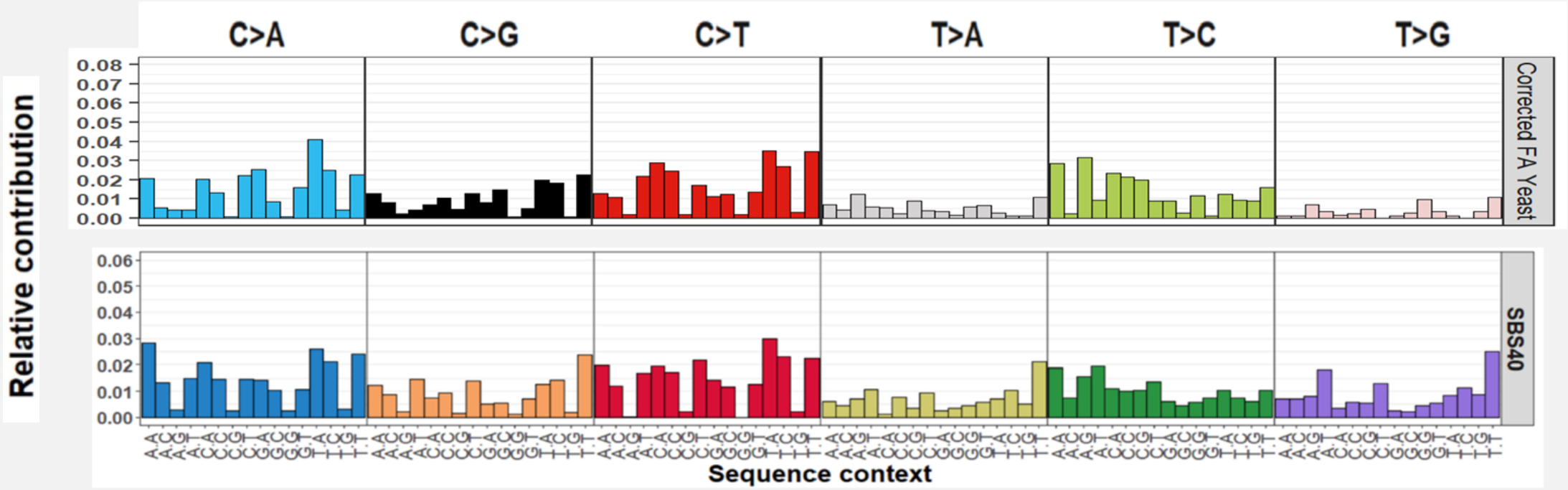
- Corrected FA signature shows more similarity towards SBS5 and SBS40 with the cosine similarity values of 0.81 and 0.89 respectively.
- Corrected AA signature shows similarity towards SBS5 and SBS40 with the cosine similarity values of 0.69 and 0.76.

Thapa et al., 2022



FA SIGNATURE AND SBS40

- SBS40 is the third most common mutational signatures found in atleast 28 cancer types.
- The etiology of the SBS40 is not confirmed yet.



Thapa et al., 2022
Alexandrov et al., 2020
COSMIC, 2021

Future Directions

Research Objective 3

- To investigate the prevalence of AA- and FA-induced mutational signatures in cancer data

- In collaboration with Dr. Alexandre Blais, I am analyzing gene expression data in cancer datasets to identify the upregulated and downregulated biological pathways correlated with cancer mutational signatures.

- I am working on DNA mutation genomics data and RNAseq (transcriptomics data) from 108 and 470 lung squamous cell carcinoma cancer patients data publicly available in cBioPortal cancer genomics domain.

CONCLUSION

- Mutational signatures can be used as biomarkers for early detection of cancer types and differentiate between different cancer types as prognostic indicators, predictors of therapeutic sensitivity, and as target of disease control.
- Formaldehyde signature resembles SBS40 based on their higher cosine similarity value (Thapa, Fabros, Alasmar, & Chan, 2022).
- Acetaldehyde signature shows less similarity to SBS5 and SBS40 compared to the formaldehyde signature. It may happen that the acetaldehyde signature may be also the composite signature of two or more signatures.

ACKNOWLDEGEMENTS

1. Principal Investigator: Dr. Kin Chan

2. TAC members: Dr. Adam Rudner, Dr. Mathieu Lavallée-Adam and Dr. Barbara Vanderhyden.

3. Lab members: Dr. Suzana Gelova, Reena Fabros, Bledar Xhialli, Ghadir Makki, Andrew Clement, Emma McGurk

THANK YOU !!

https://newlinetheatre.blogspot.com/2018/04/stasis-is-membrane-that-keeps-yeasts.html