

# Decoupling mutational processes in human germline

Vladimir Seplyarskiy

Sunyaev lab



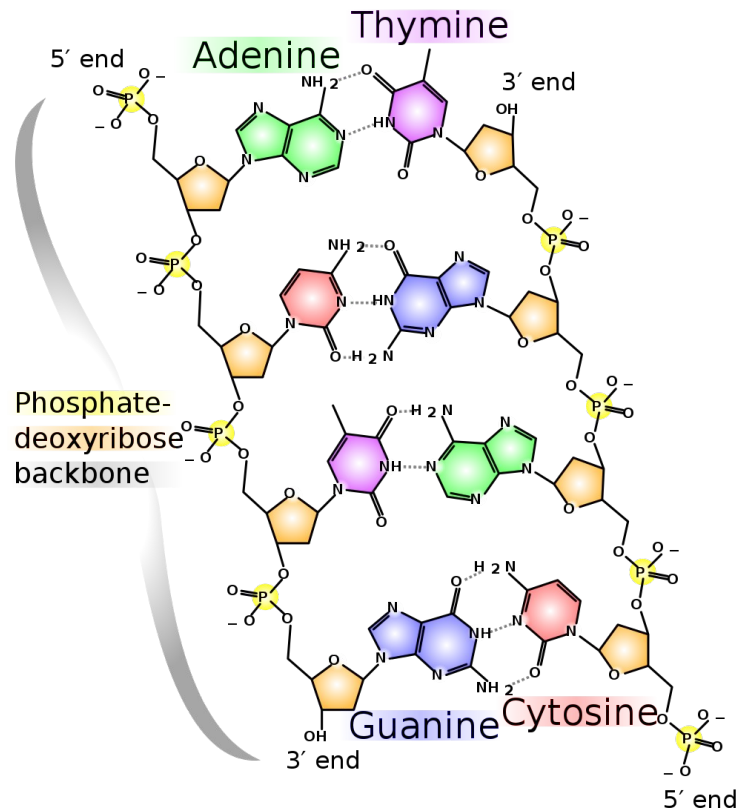
**Department of Biomedical Informatics**  
Harvard Medical School



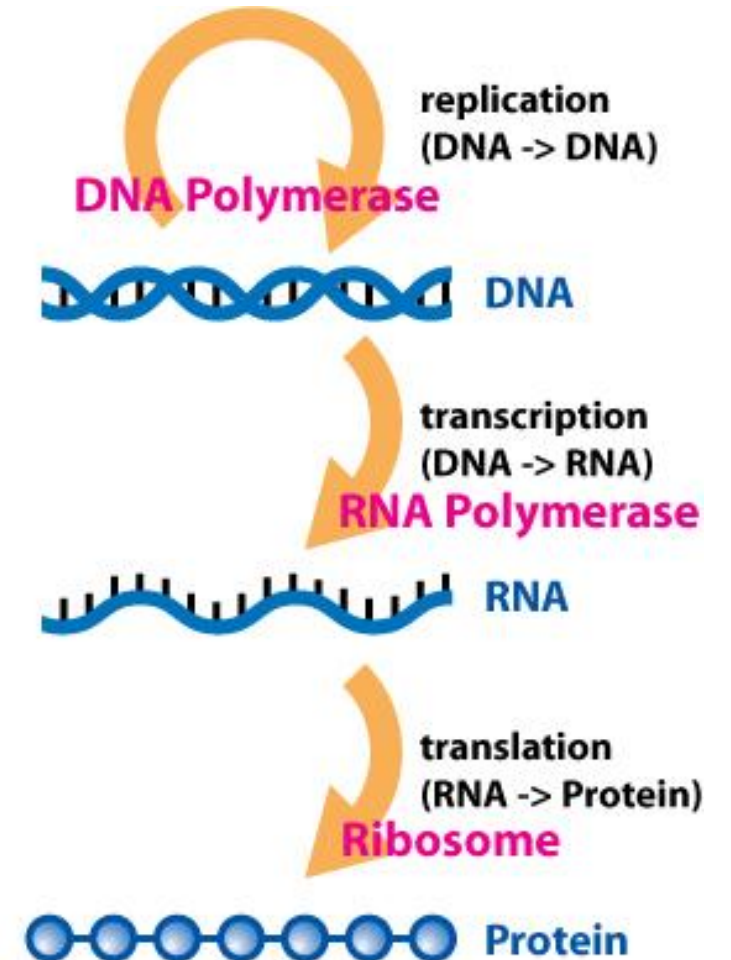
**Division of Genetics**  
Department of Medicine  
Brigham and Women's Hospital / Harvard Medical School

# DNA and it's function

DNA double stranded molecule  
Each strand have direction

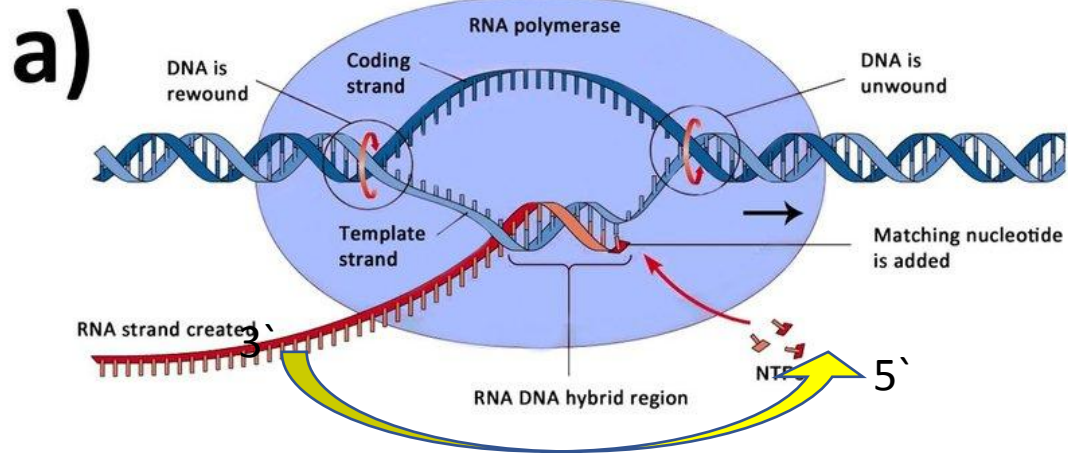


DNA->RNA->Protein

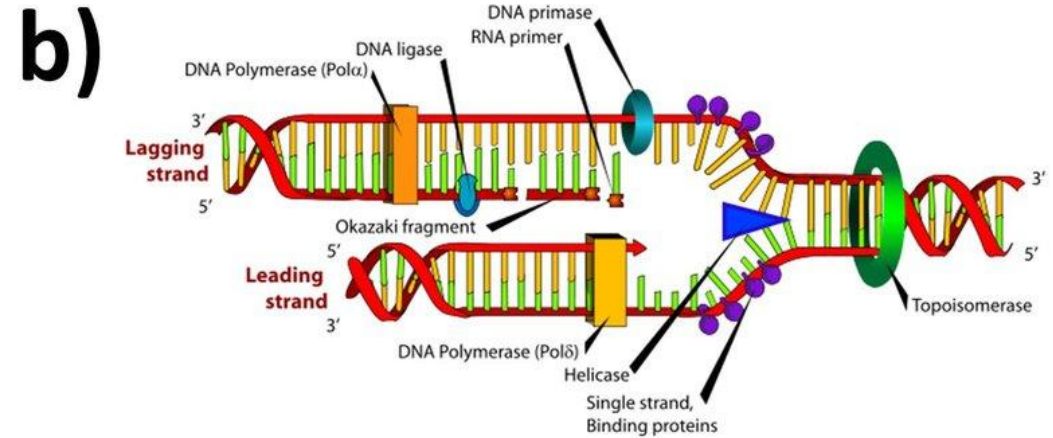


# Transcription and replication

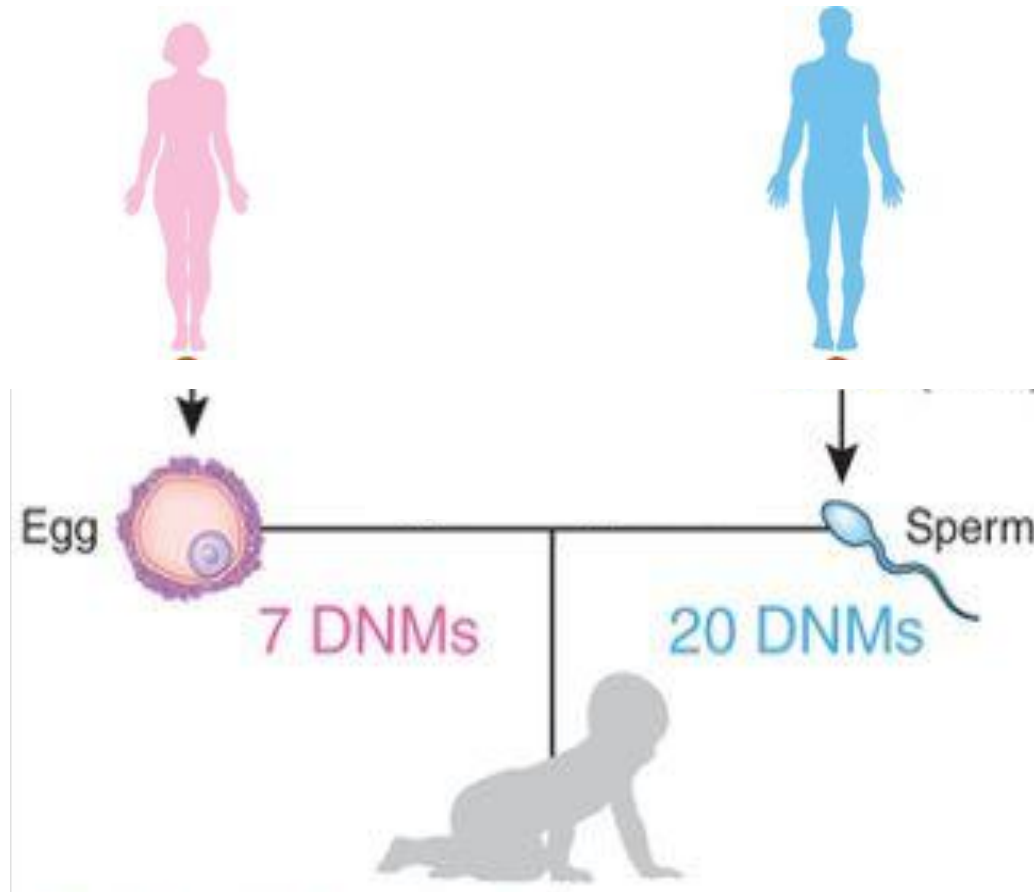
DNA→RNA



DNA→DNA

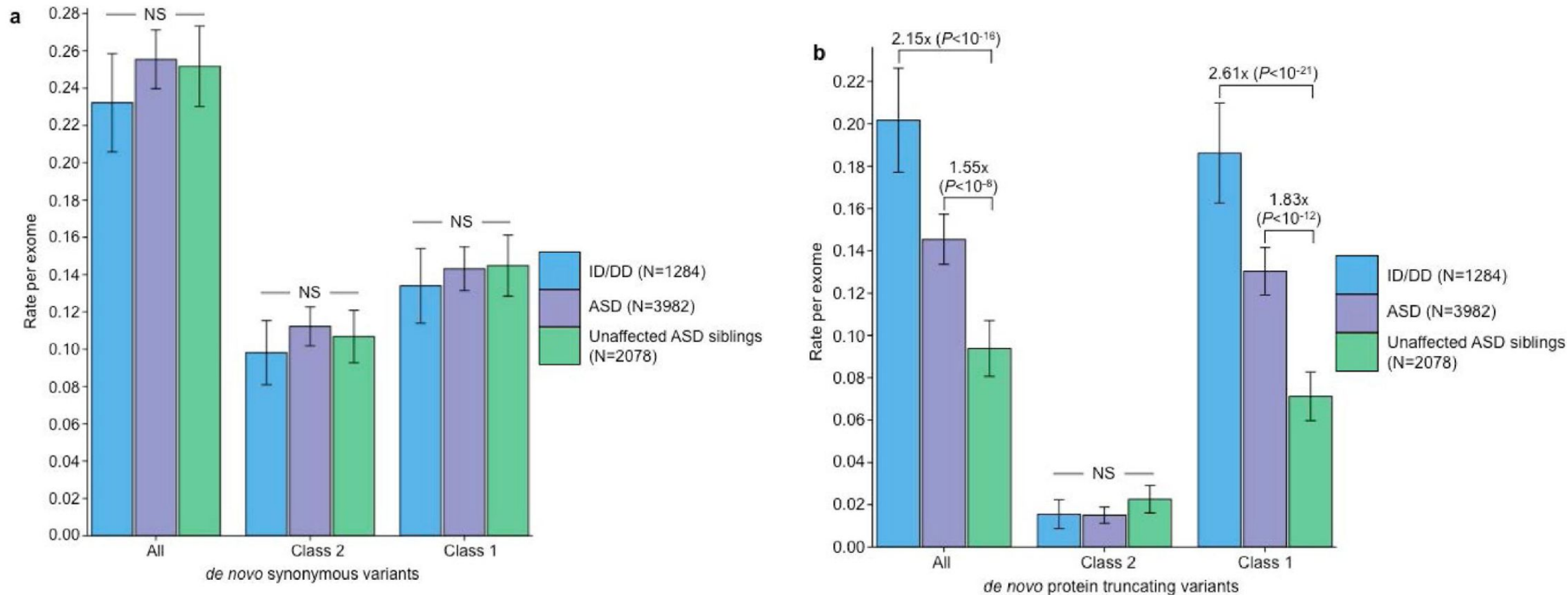


# Germline mutations



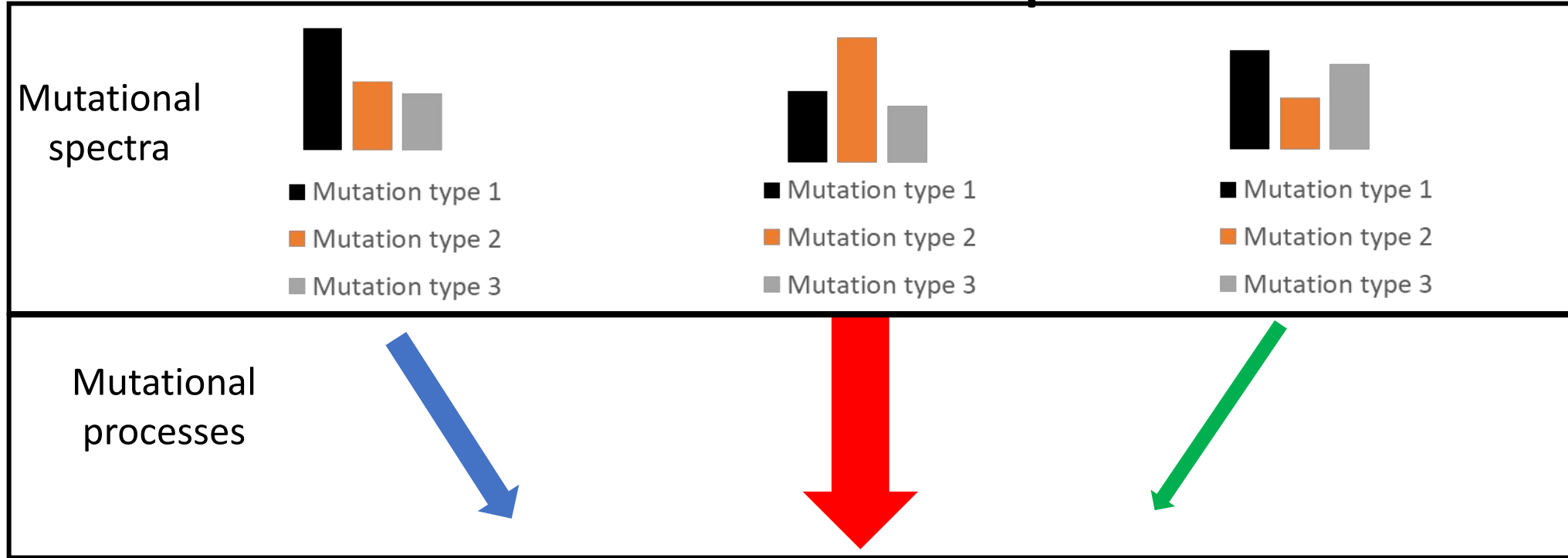
Goriely, Nat. Gen. 2016

# Role of mutations in genetic disease



Kosmicki et al, Nature Genetics 2017

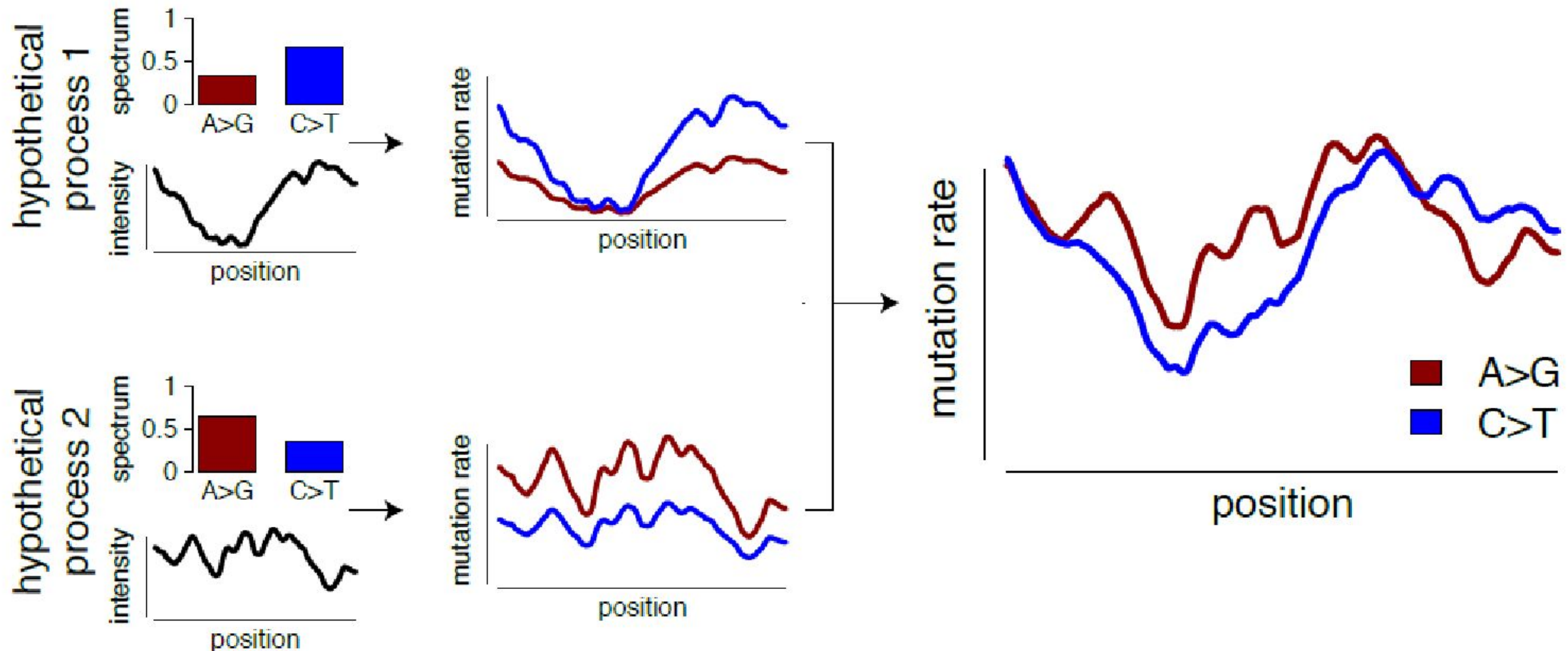
# Germline mutations are induced by a mixture of mutational processes



Observed mutations

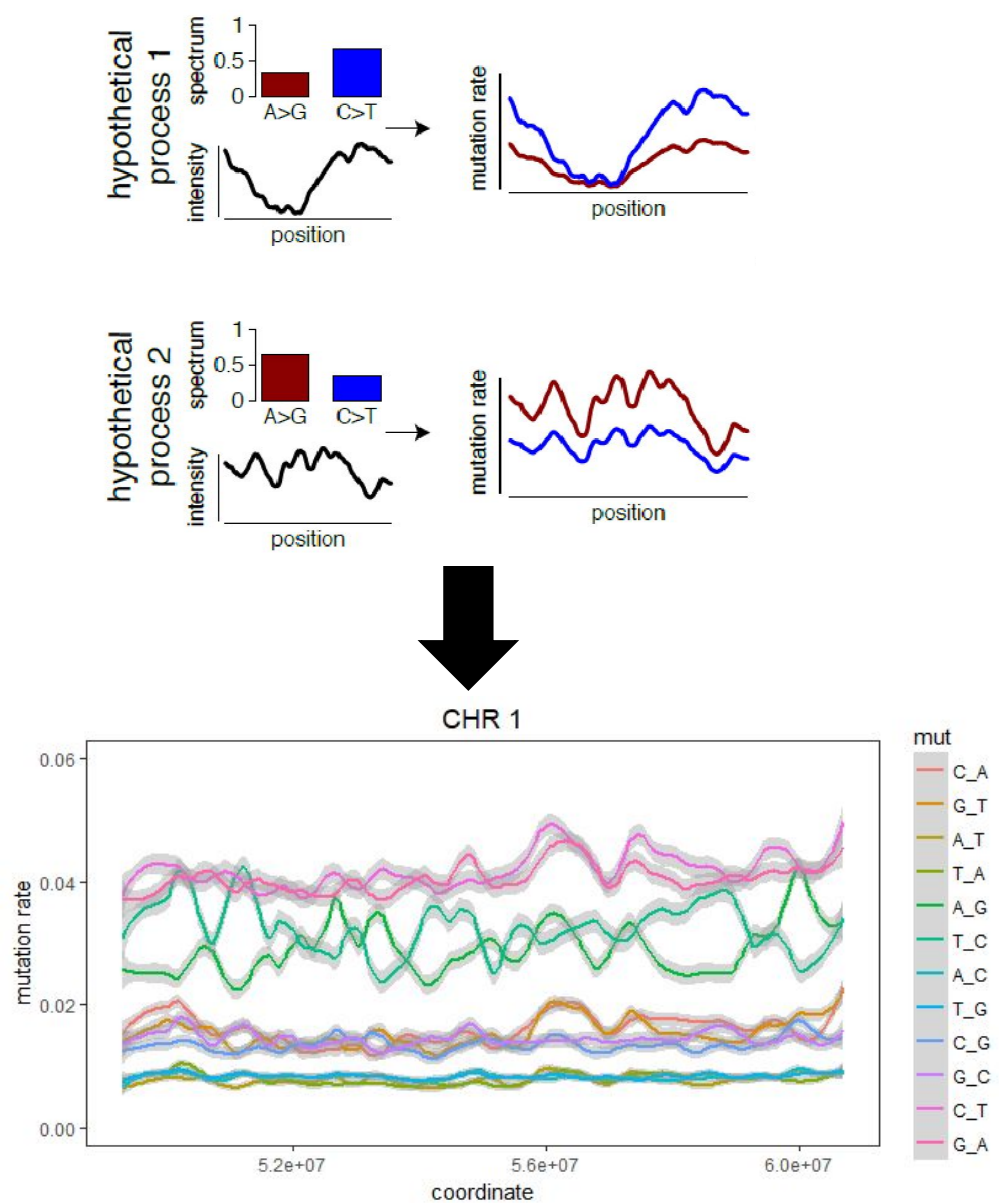


# Spatial variation in intensity of the mutational processes generates diverse mutational patterns





# Extracting mutational processes from the spatial variation



$$\begin{matrix} \text{\#Mutation types} \\ S \end{matrix} \times \begin{matrix} I \\ \text{\#Regions} \end{matrix} = \begin{matrix} \text{\#Signatures} \end{matrix}$$

$$\begin{matrix} \text{\#Signatures} \\ \text{(mutational processes)} \end{matrix}$$

$$\begin{matrix} \text{\#Mutation types} \\ M \end{matrix} = \begin{matrix} \text{\#Regions} \end{matrix}$$

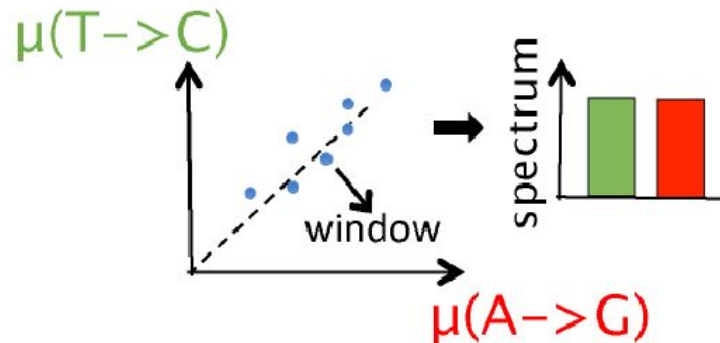
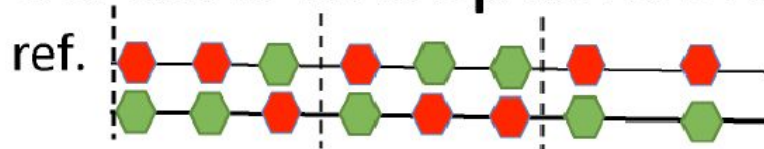


# Manifestation of strand dependent and independent mutational processes

3

mutational process

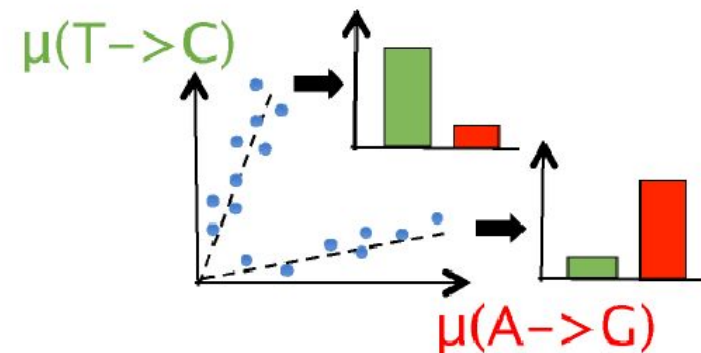
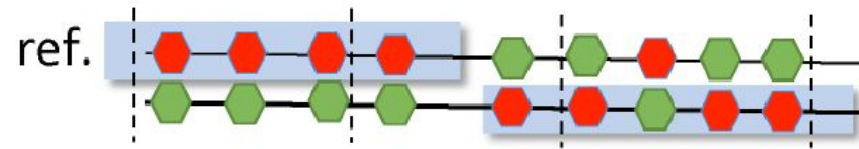
strand independent



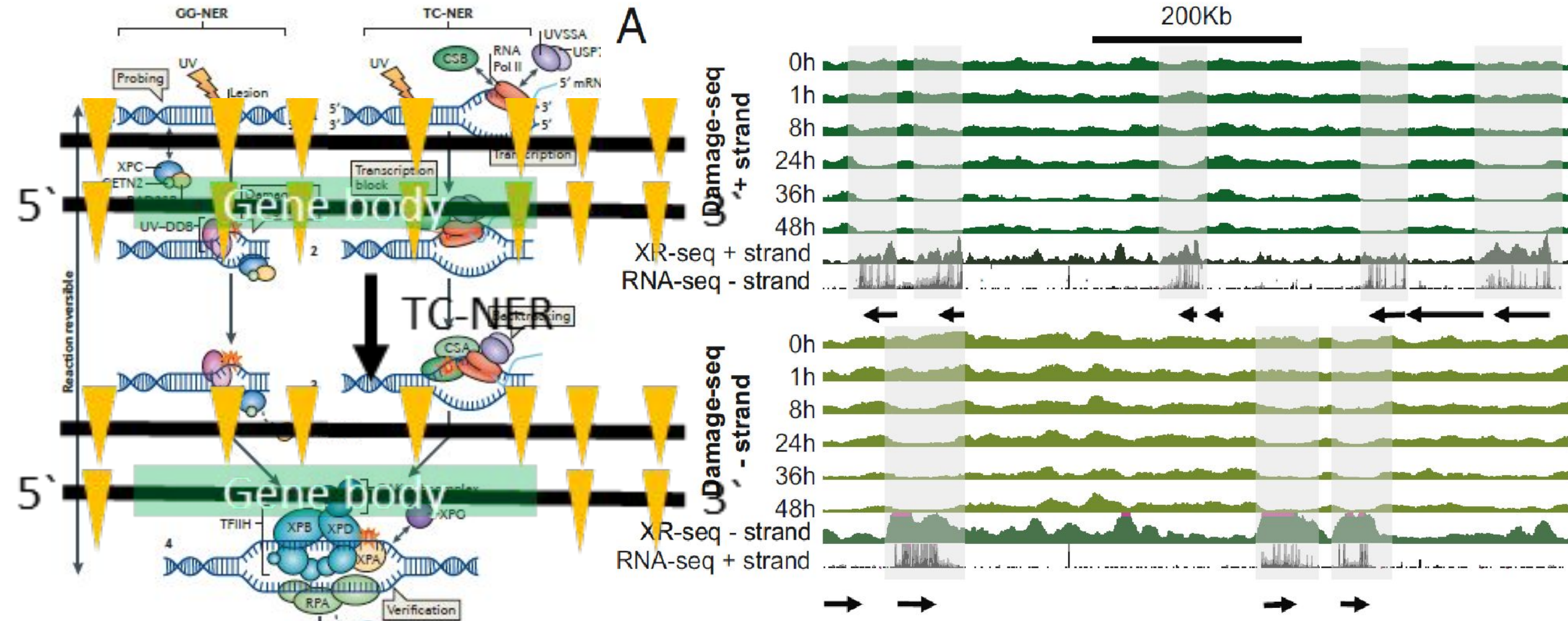
complementary mutations

red hexagon A -> G  
green hexagon T -> C

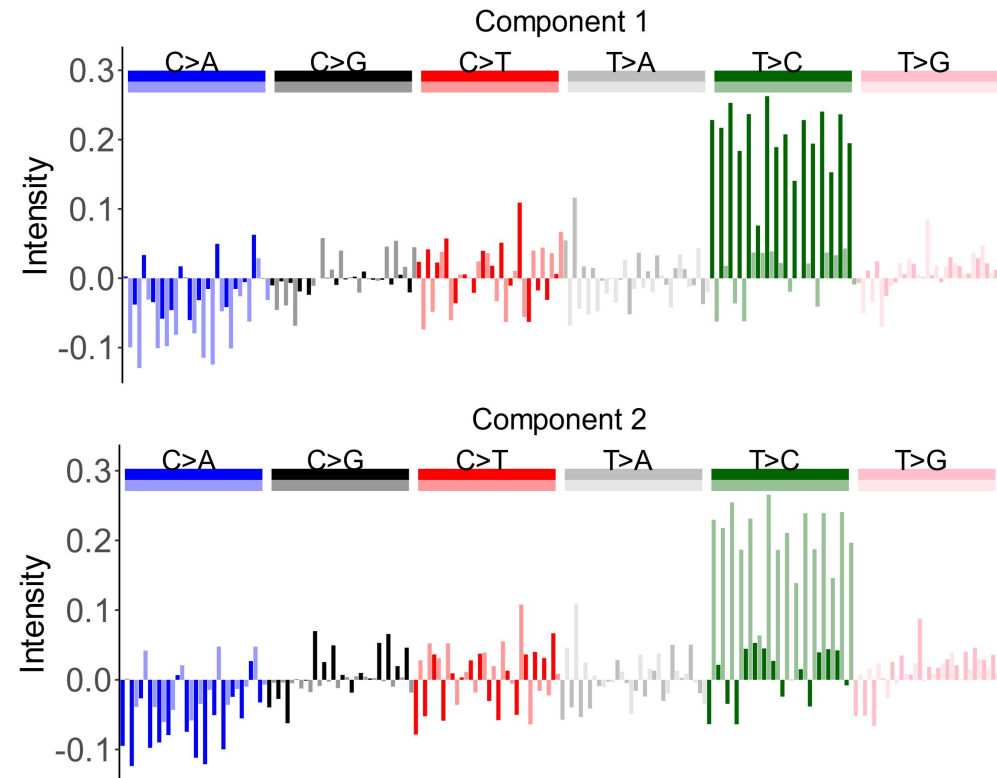
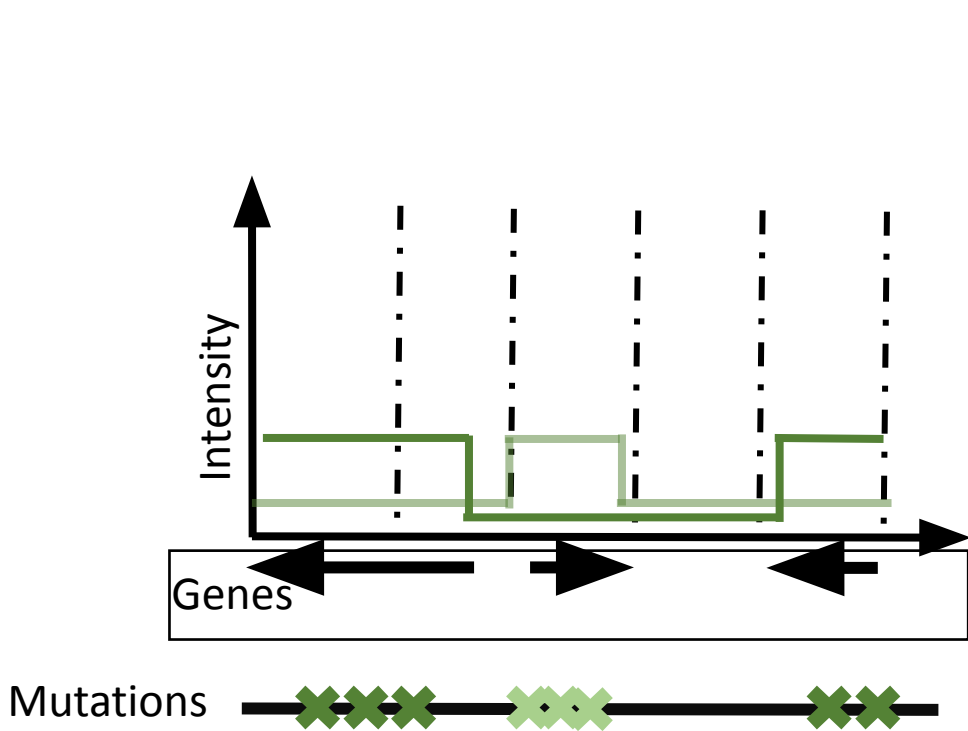
strand dependent



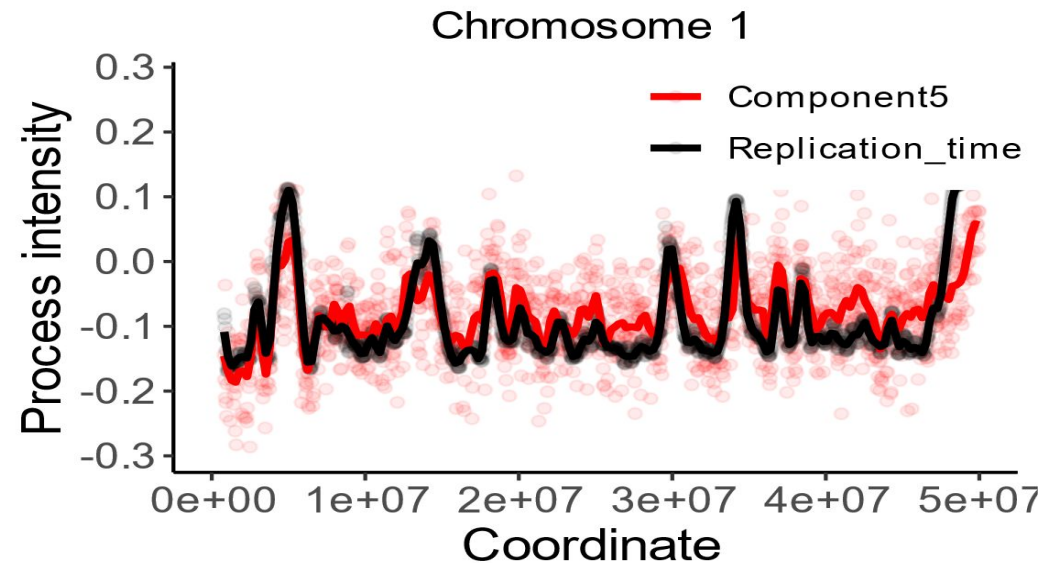
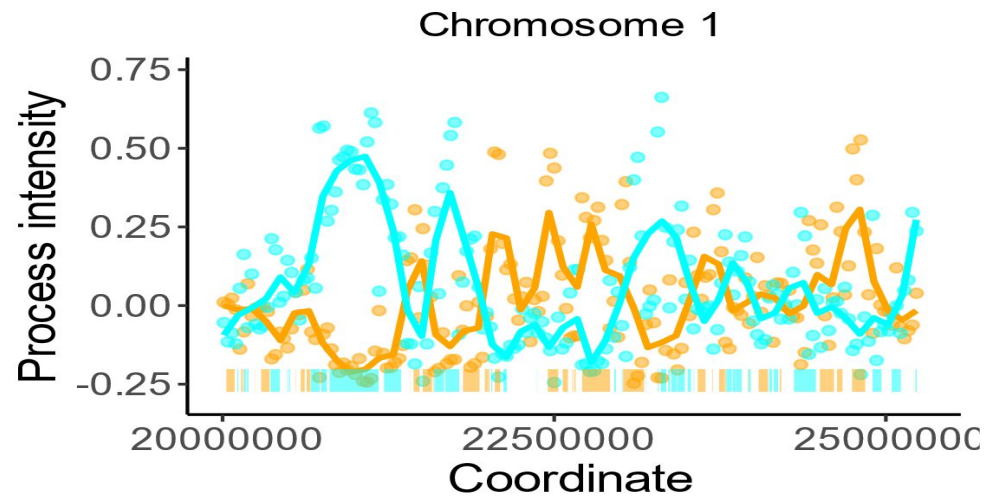
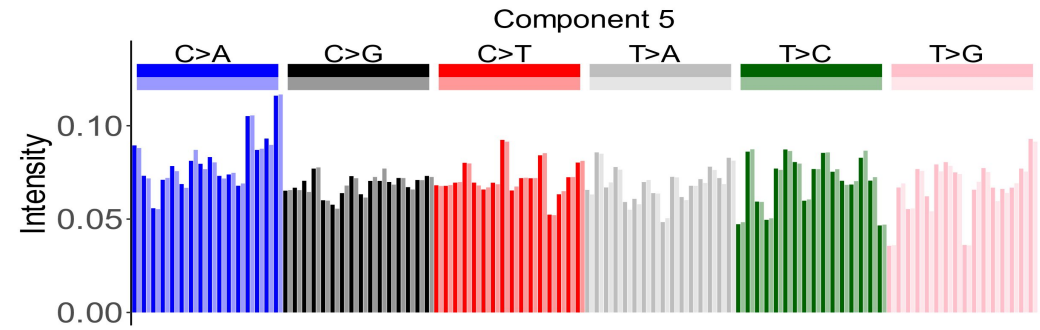
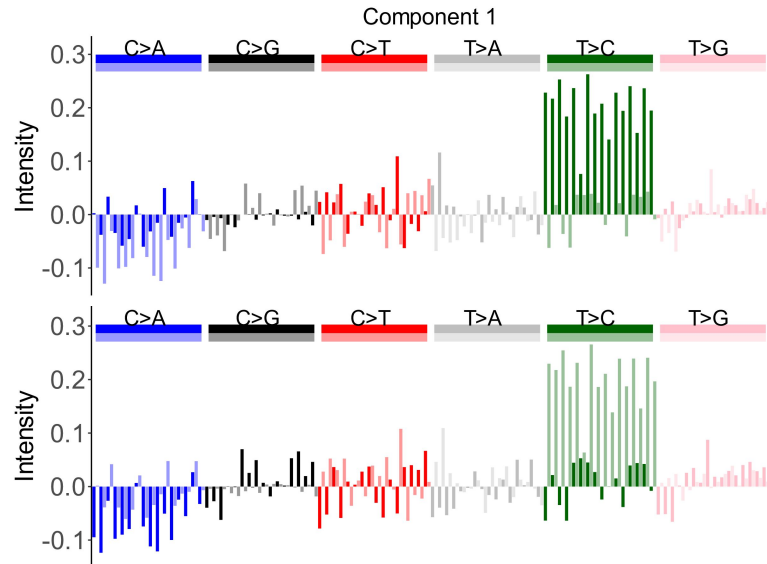
# Mutational asymmetry in genes



# Manifestation of strand dependent and independent mutational processes

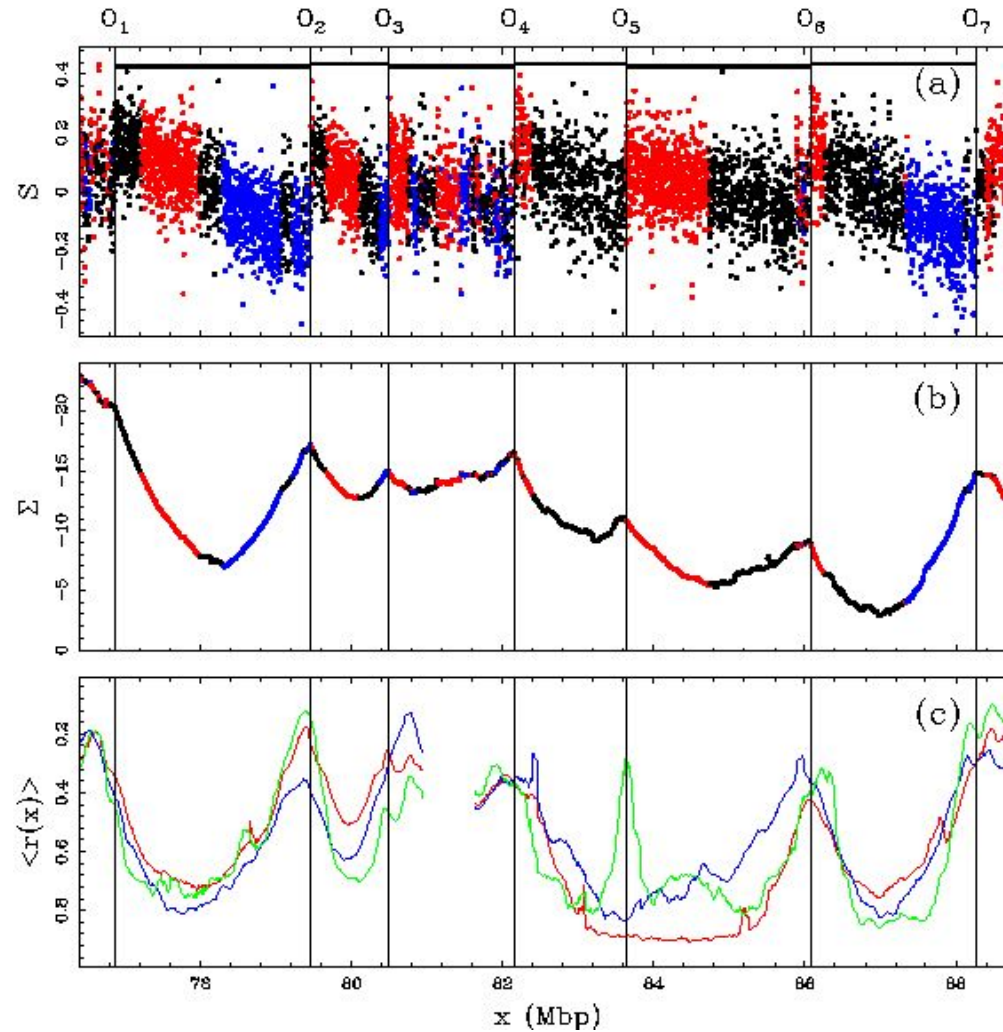


# What do we find?



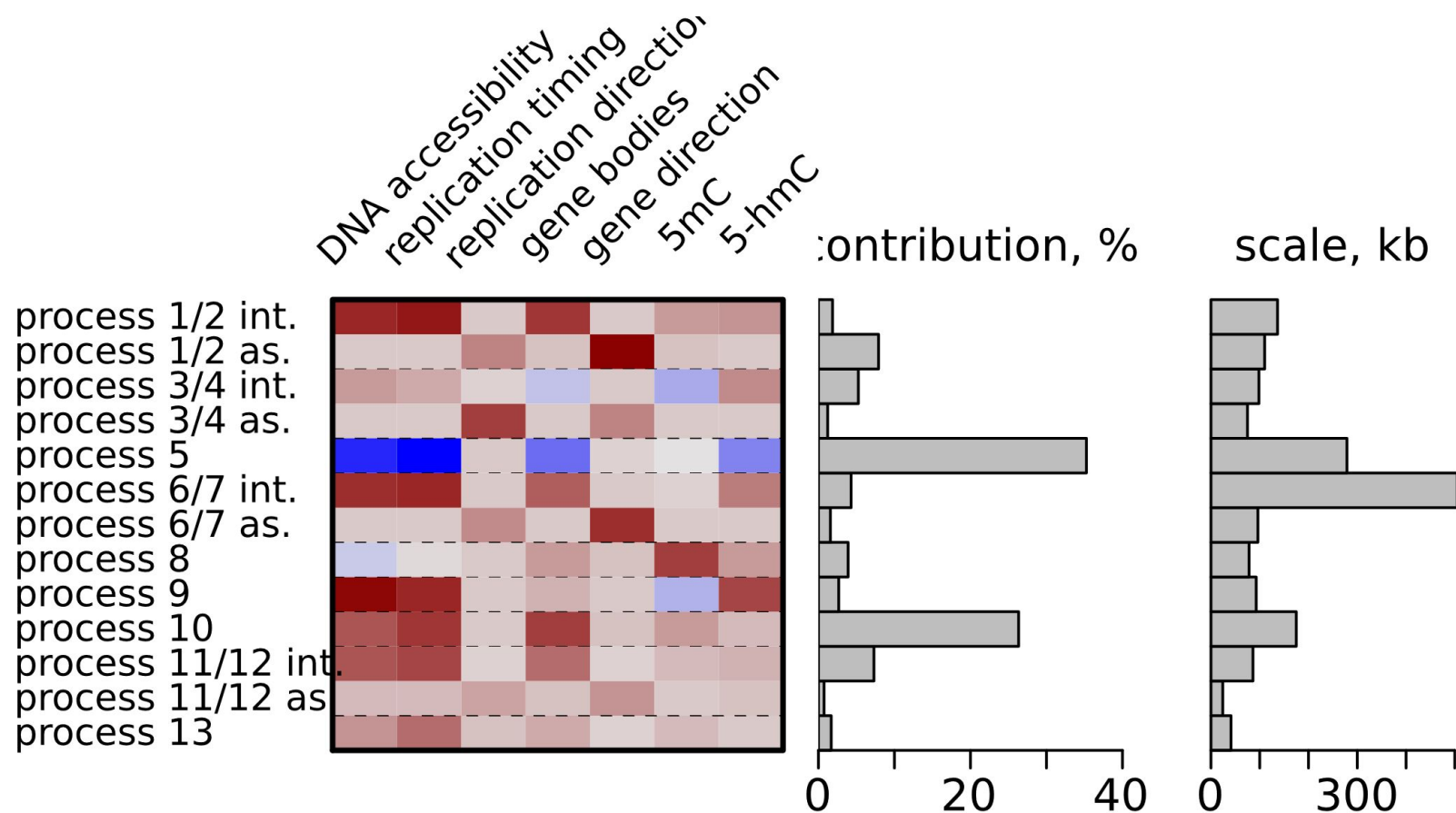


# Replication Program

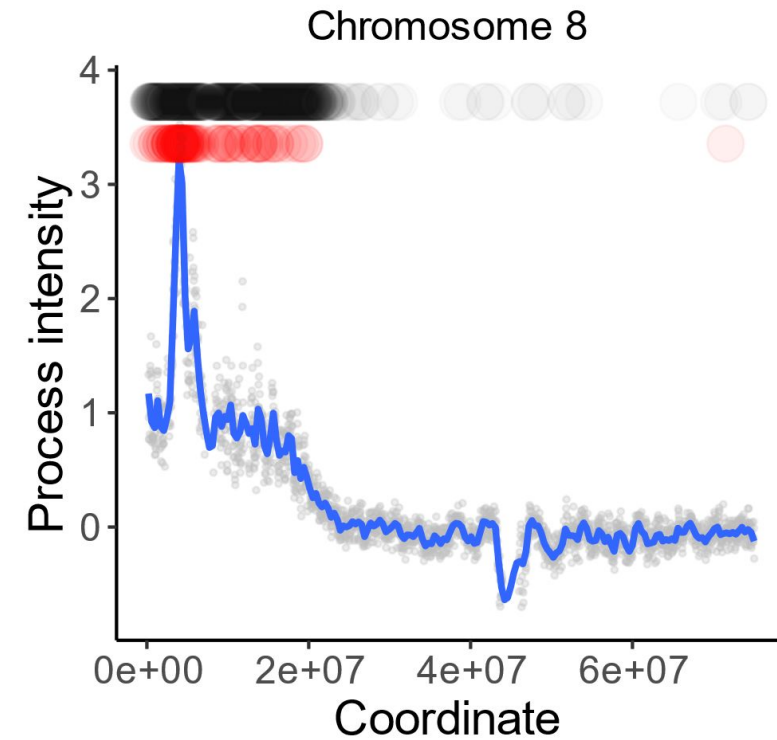
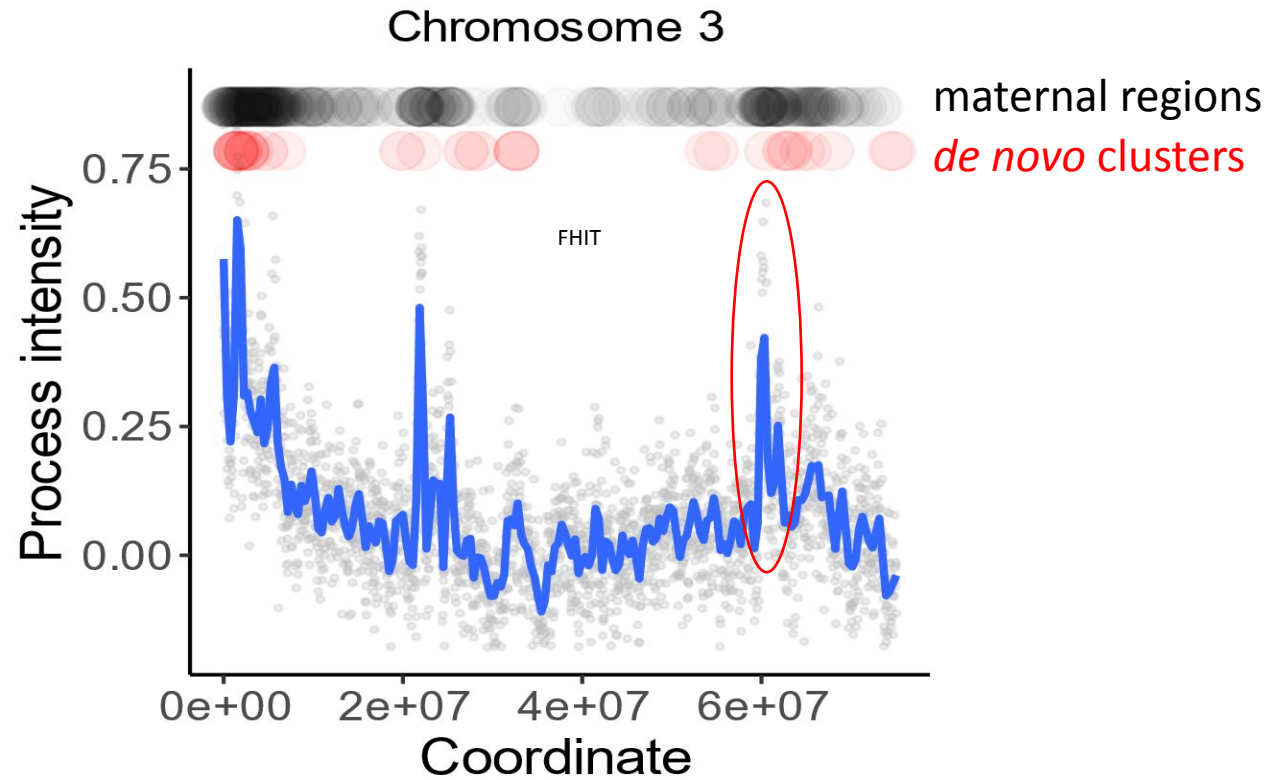


Arneodo lab, Plos Comp 2012

# What do we find

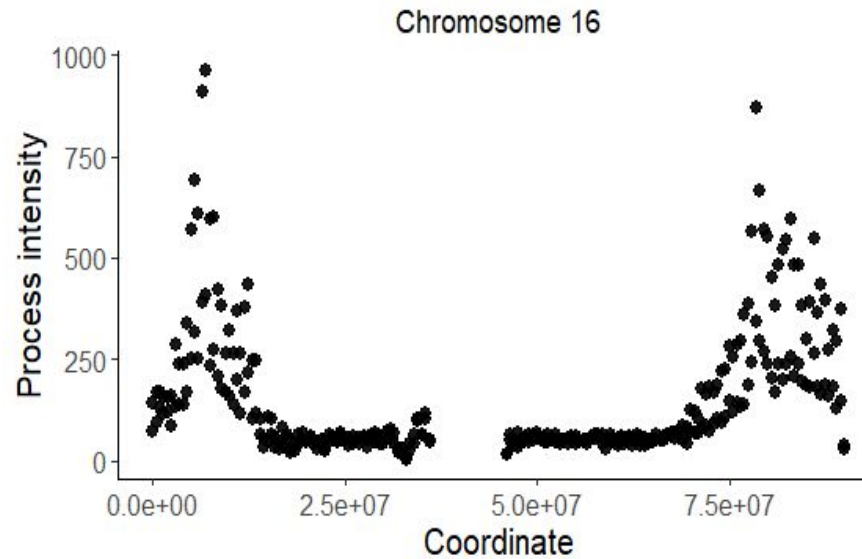
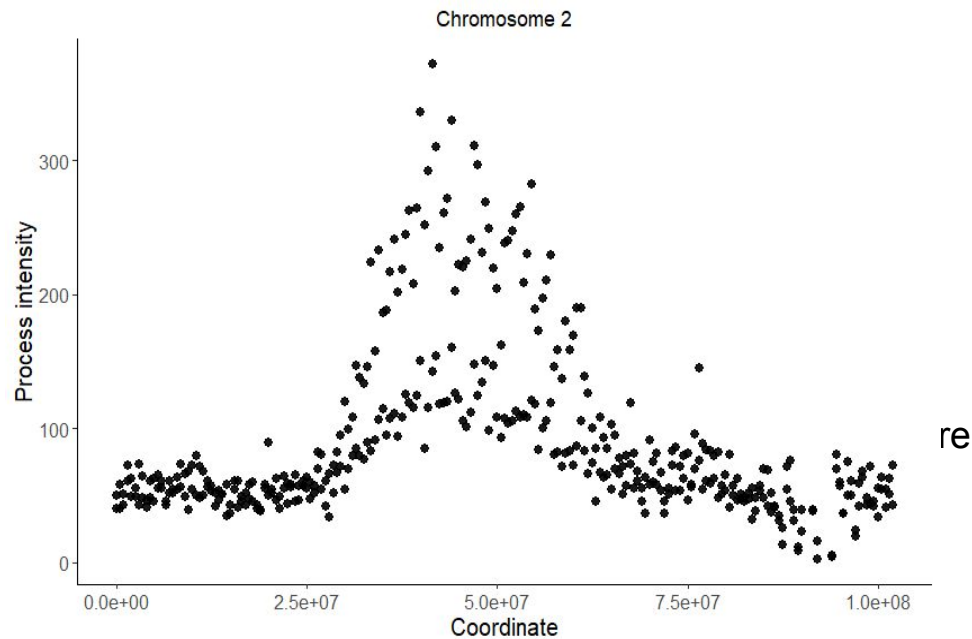
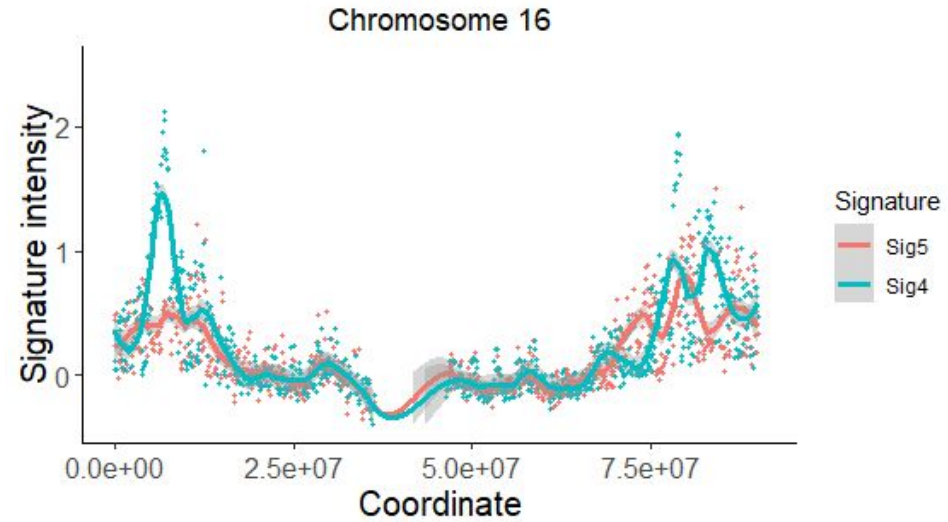
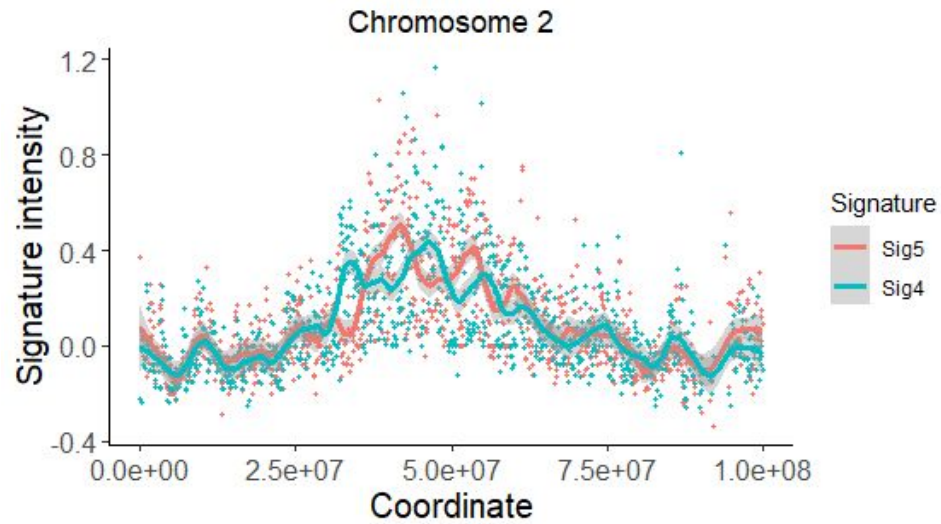


# Process 7/8 have local bursts

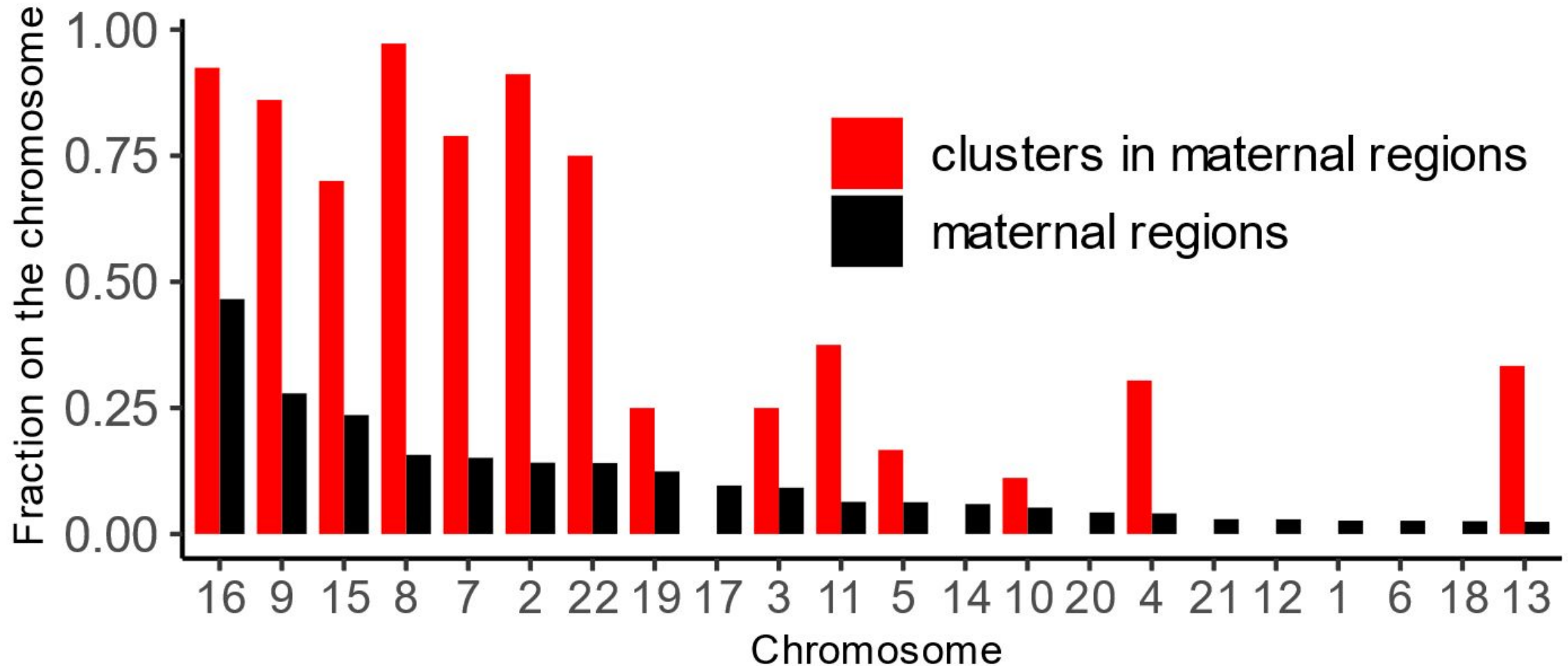




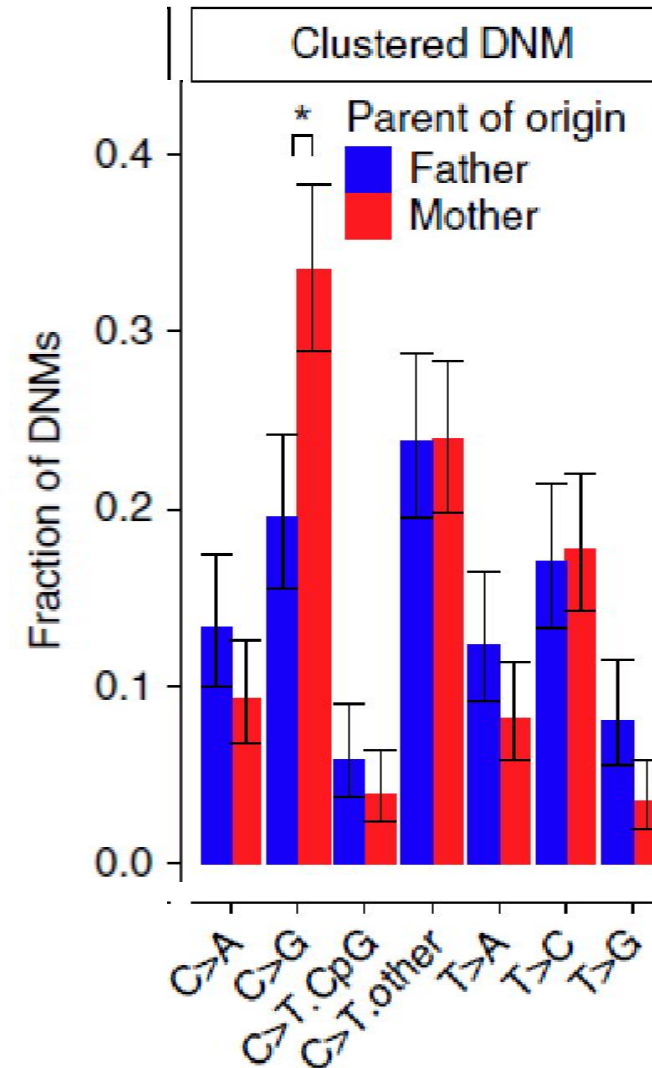
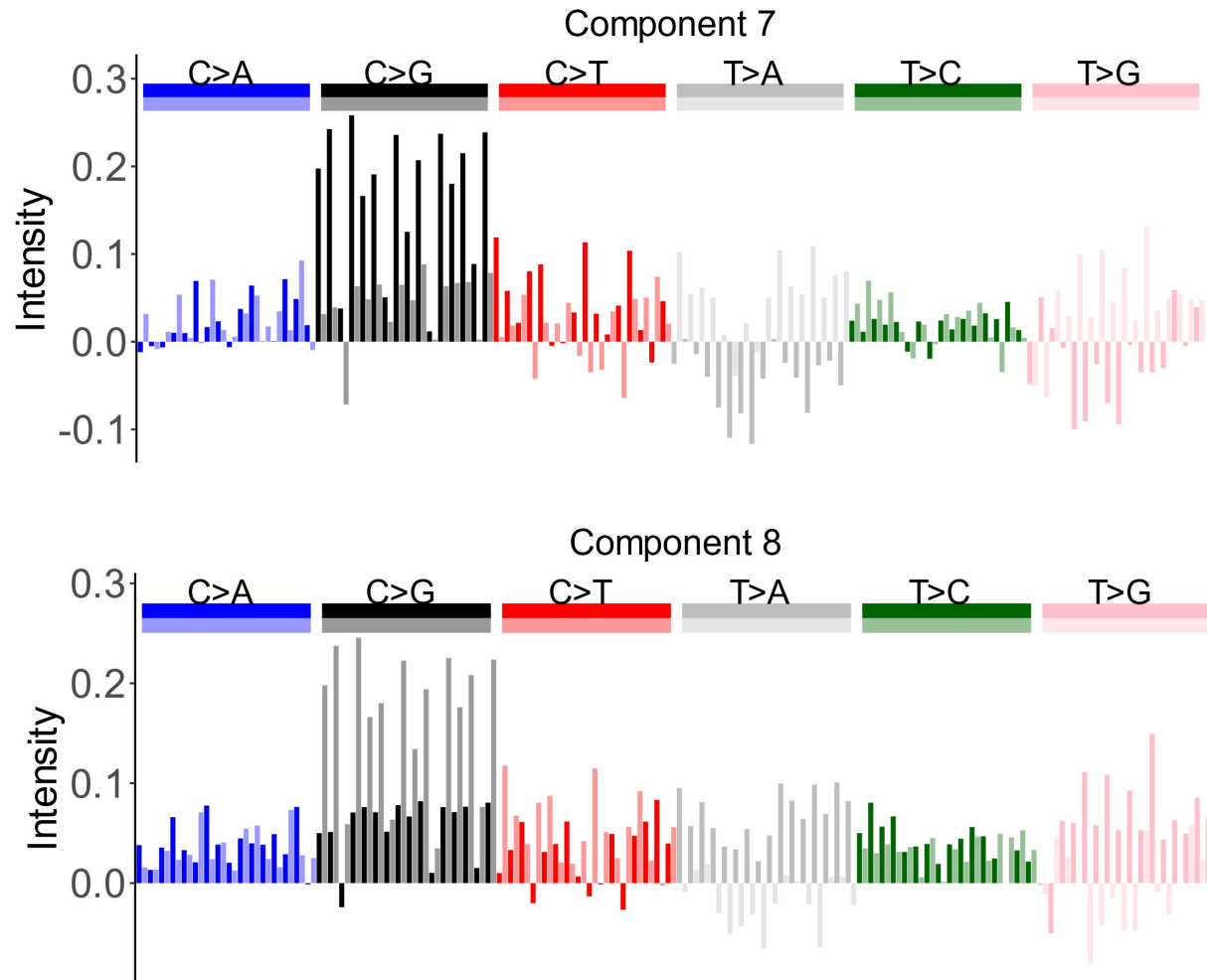
# Process 7/8 – maternal signature



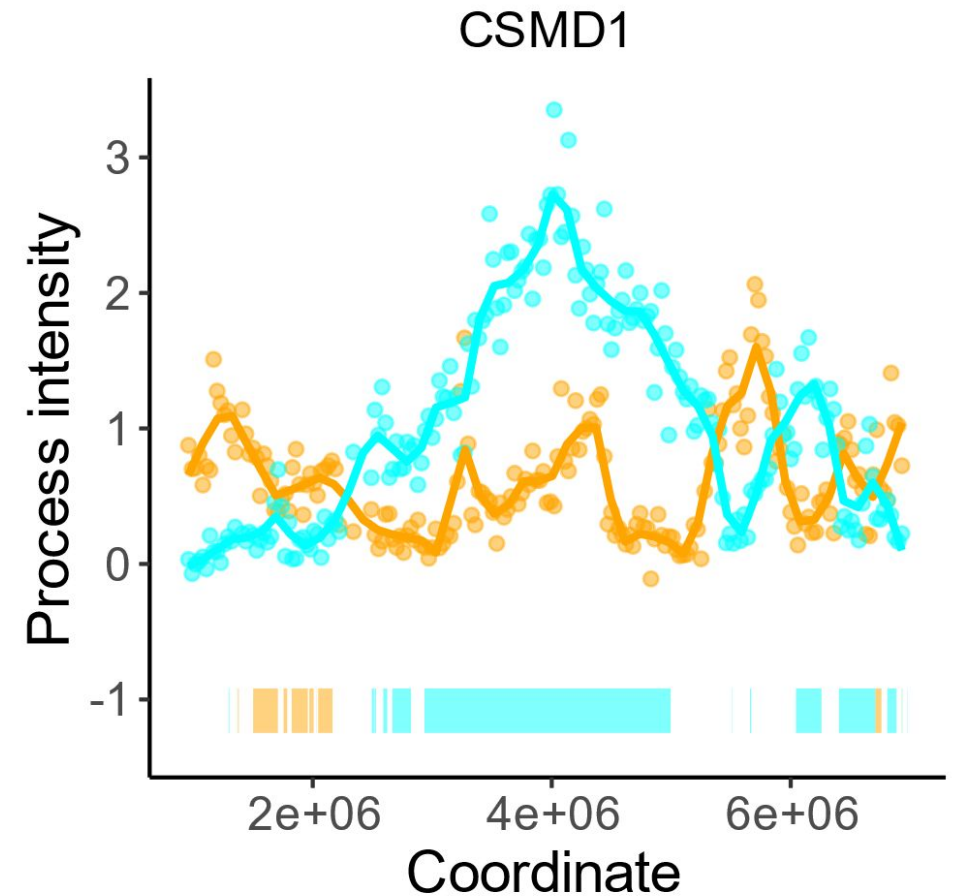
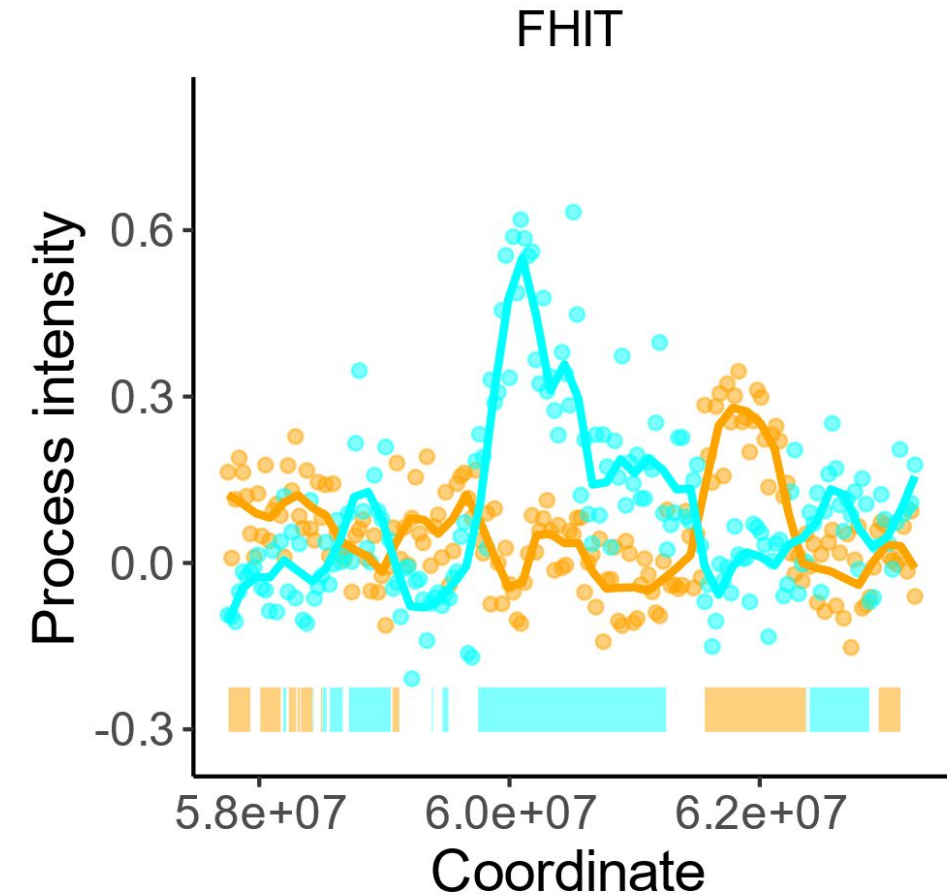
We are able to predict genomic regions susceptible to maternal clusters



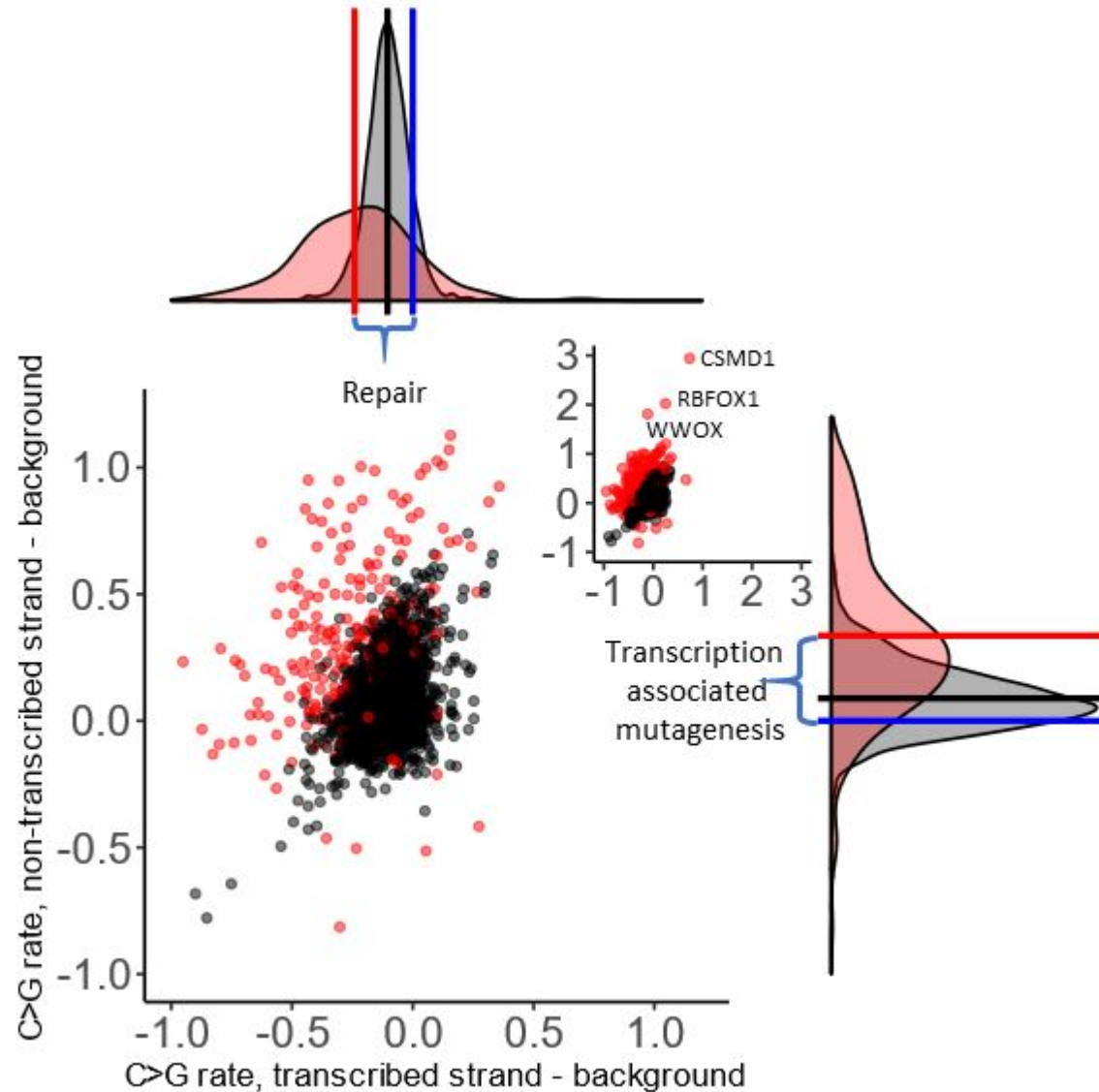
# Process 7/8– maternal signature



# Process 7/8 asymmetric in respect to transcription



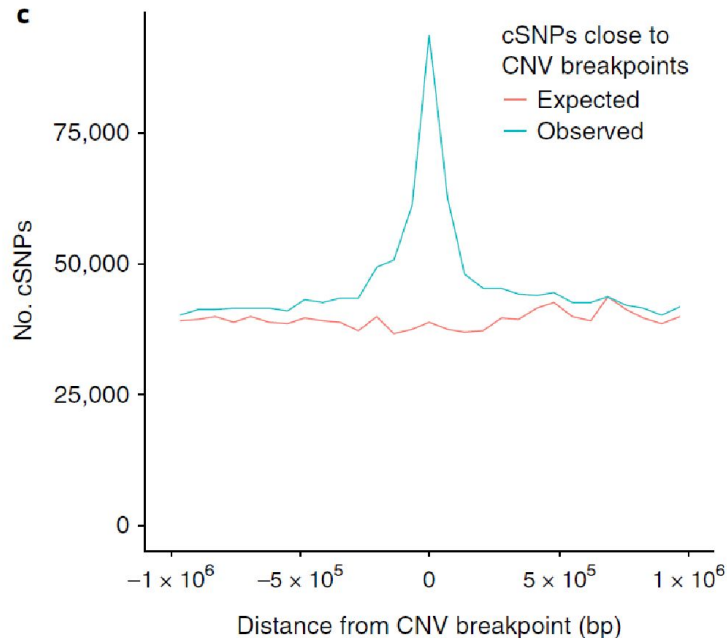
# Process 7/8 asymmetric in respect to transcription



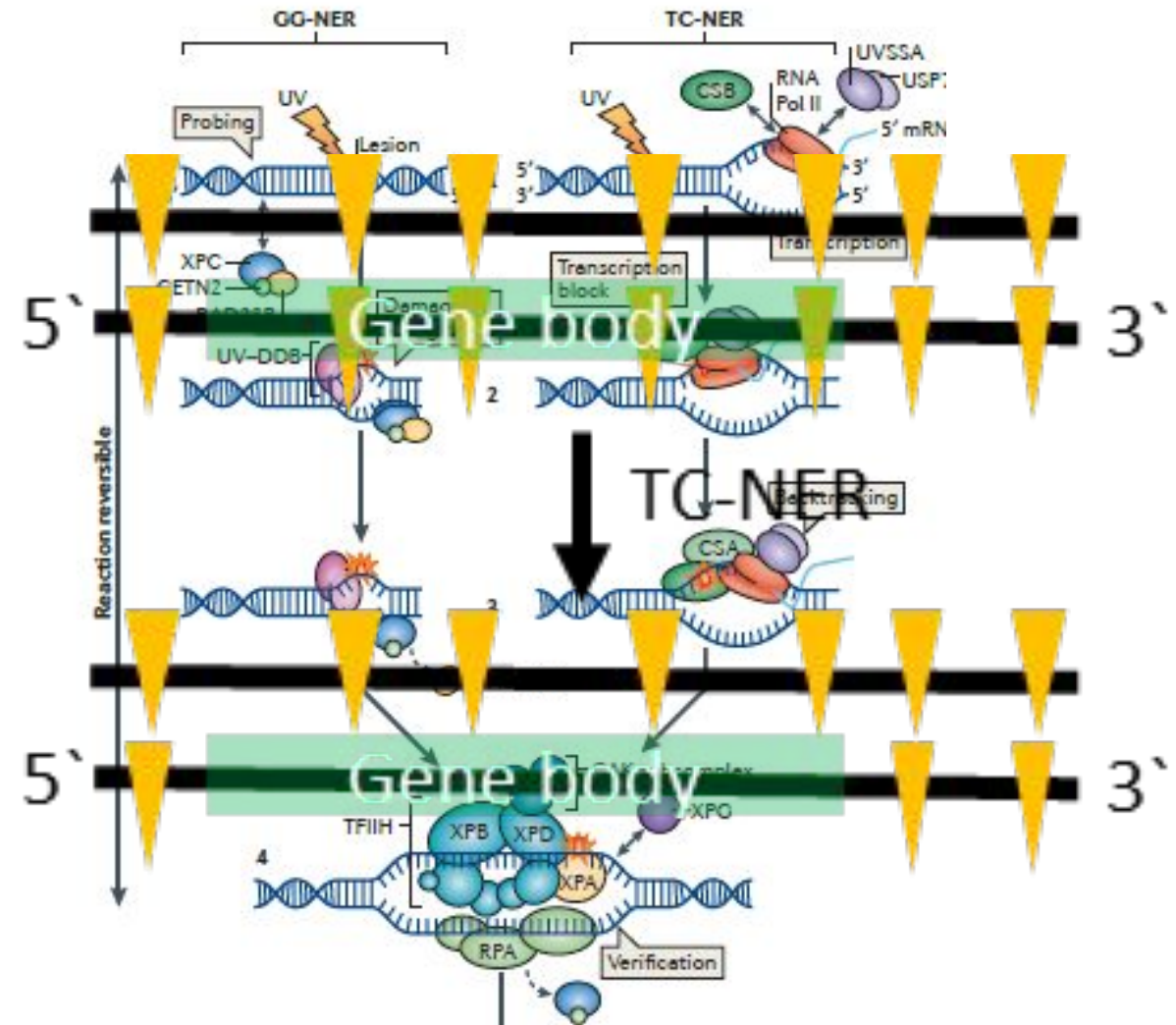


# Mechanistic insights to maternal signature

I was proven wrong



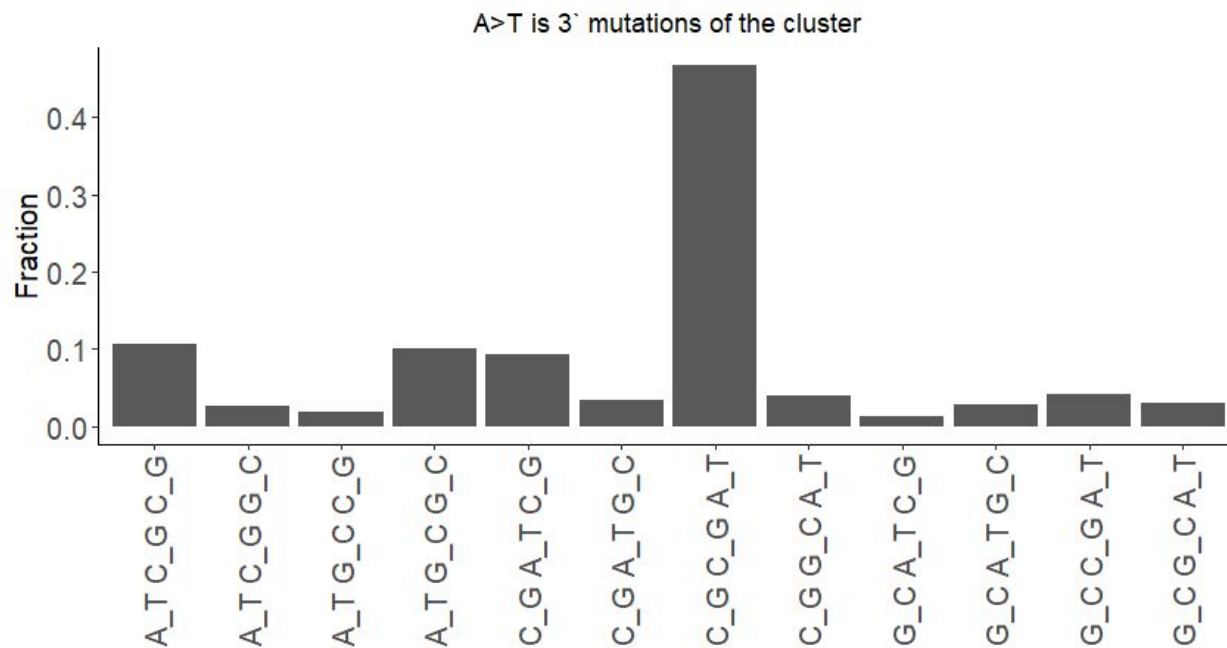
Goldman\*, Seplyarskiy\*, Wong\* et al, 2018 Nature Genetics



Marteijn et al 2014, Nat. Rev Mol. Cell. Bio

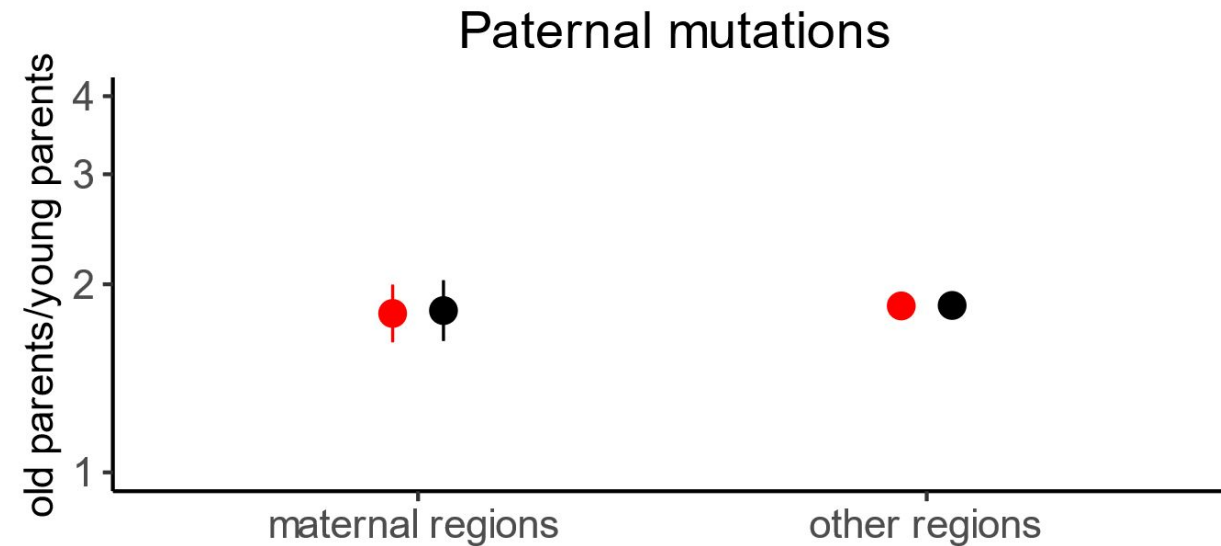
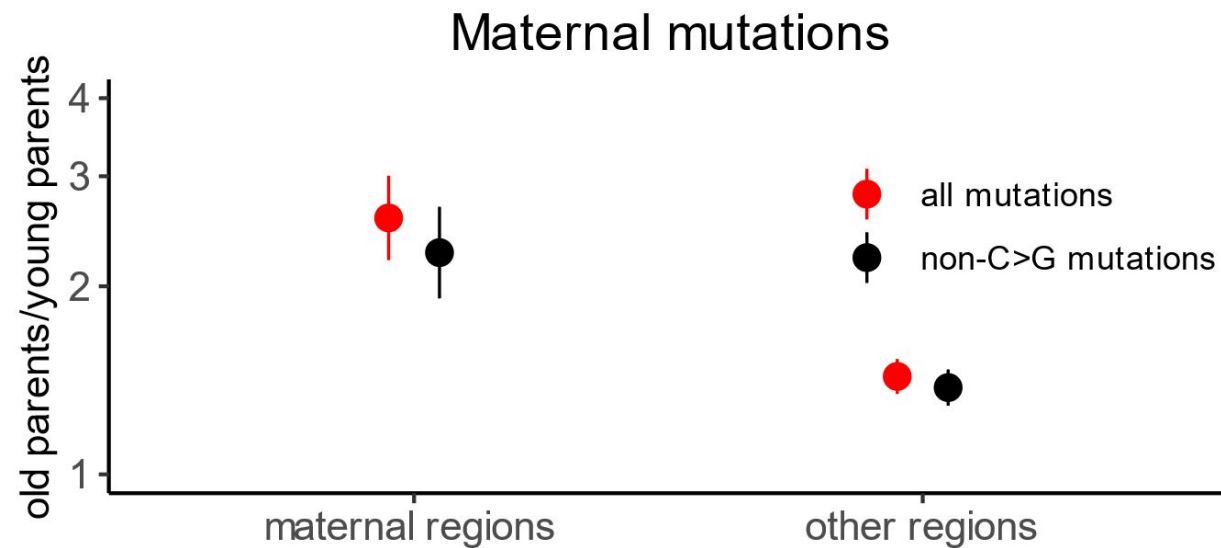
# C>G mutations created by secondary mechanism (wrong already)

Clustered mutation contain few C>G substitutions,  
But started with different mutation type





# Third of maternal age effect localized in the regions with high level of the process 7/8



# Insights about maternal signature

- Maternal signature sensitive to the direction of transcription and replication (no evidence for DNA breaks)
- Characteristic scale of the signature about 20Mb, but highest intensity achieved on non-transcribed strand of long genes
- We find evidence for transcription associated mutagenesis
- Maternal signature is by product of the activity of error-prone polymerase

# Acknowledgements



Signature extraction

Ruslan Soldatov

Colaborators:

J. Goldmann, P.Kharchenko,  
C. Gilissen, W. Wong



Supervision

Shamil Sunyaev

TOPMed population working  
group