

# Analysis of editing sites characteristics in RNA-Seq data

Irina Shchukina

Alexandr Kanapin, Anastasia Samsonova  
Department of Oncology, University of Oxford

# Types of RNA-editing

A → G (ADAR proteins)

C → U (APOBEC)

Global goal: framework development

Current goal: to find any site's features that can help to separate editing events from noise

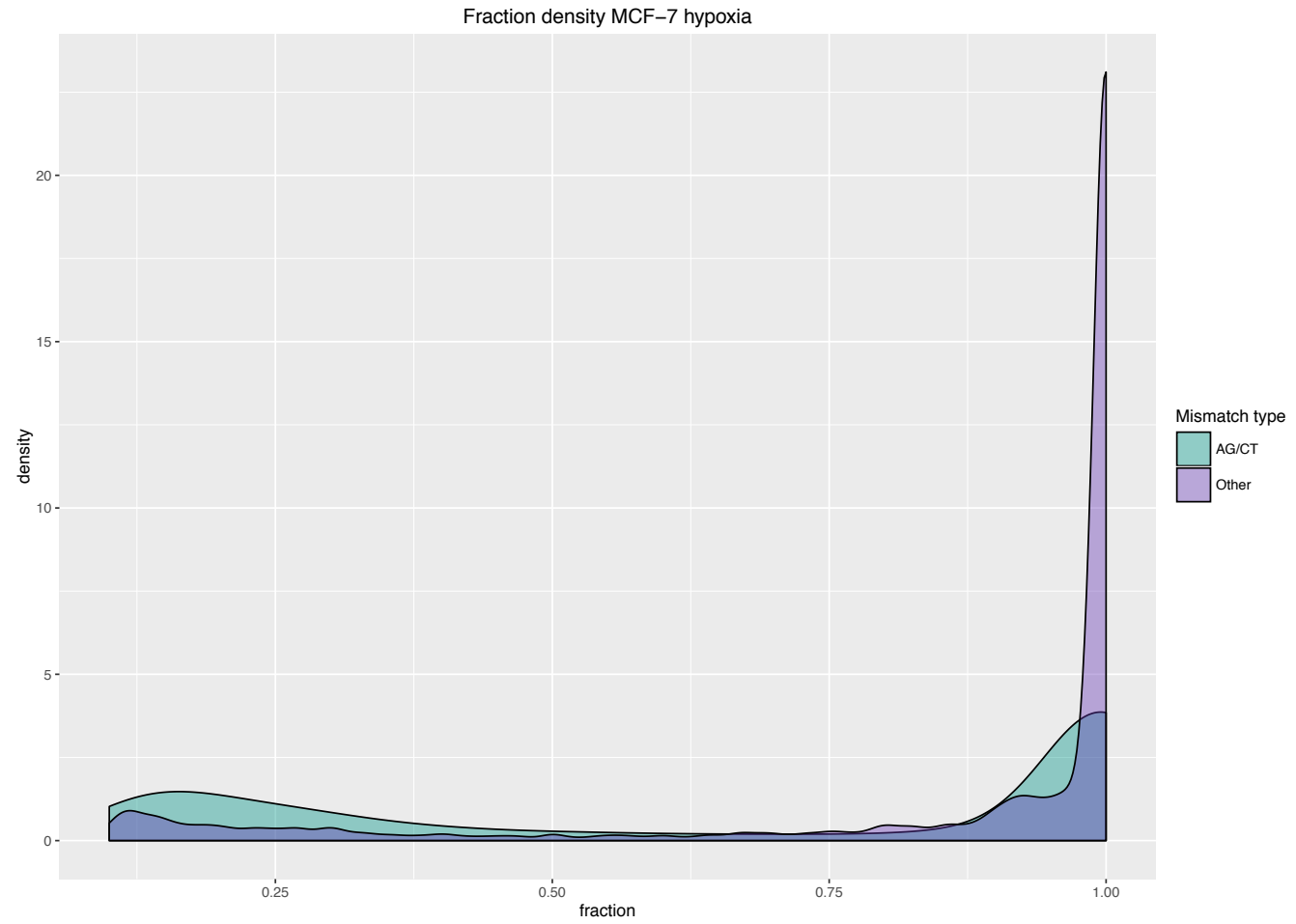
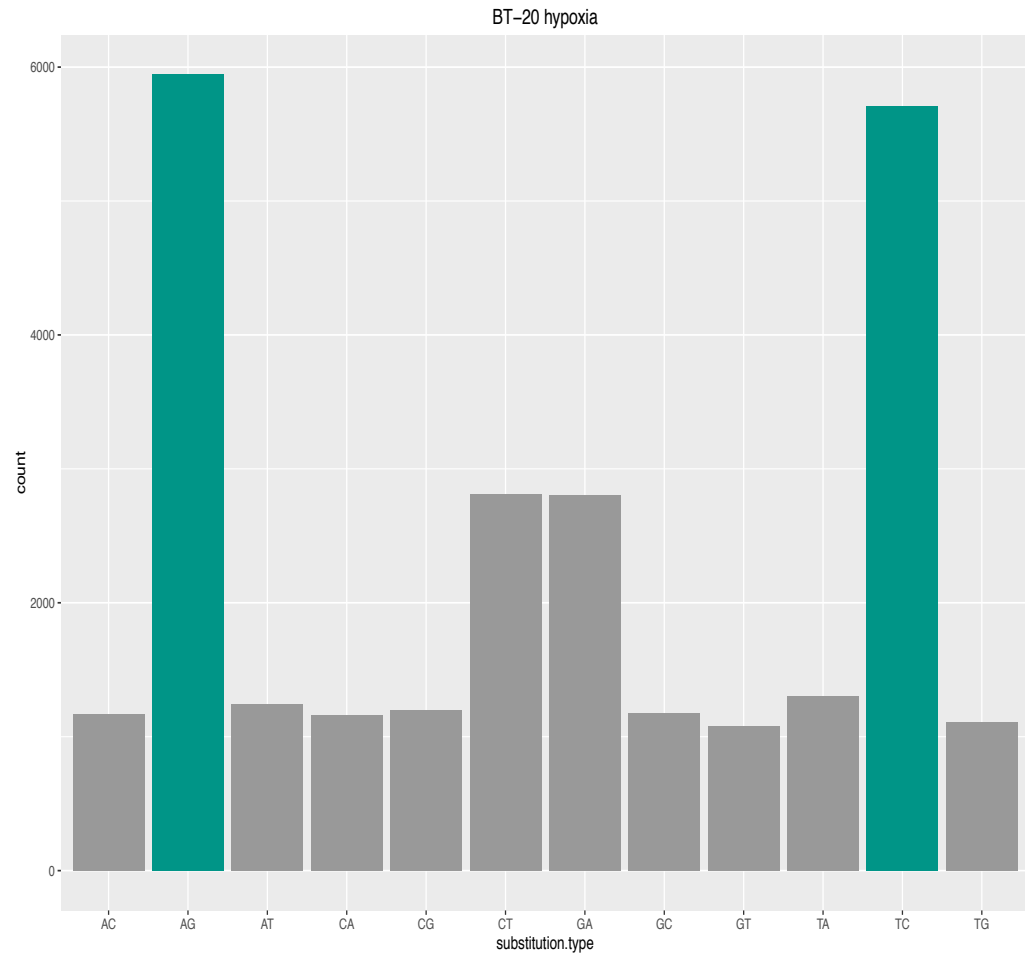
# Features to explore

- Different quality metrics
- Fraction
- Local GC-content
- Motif
- Secondary structure

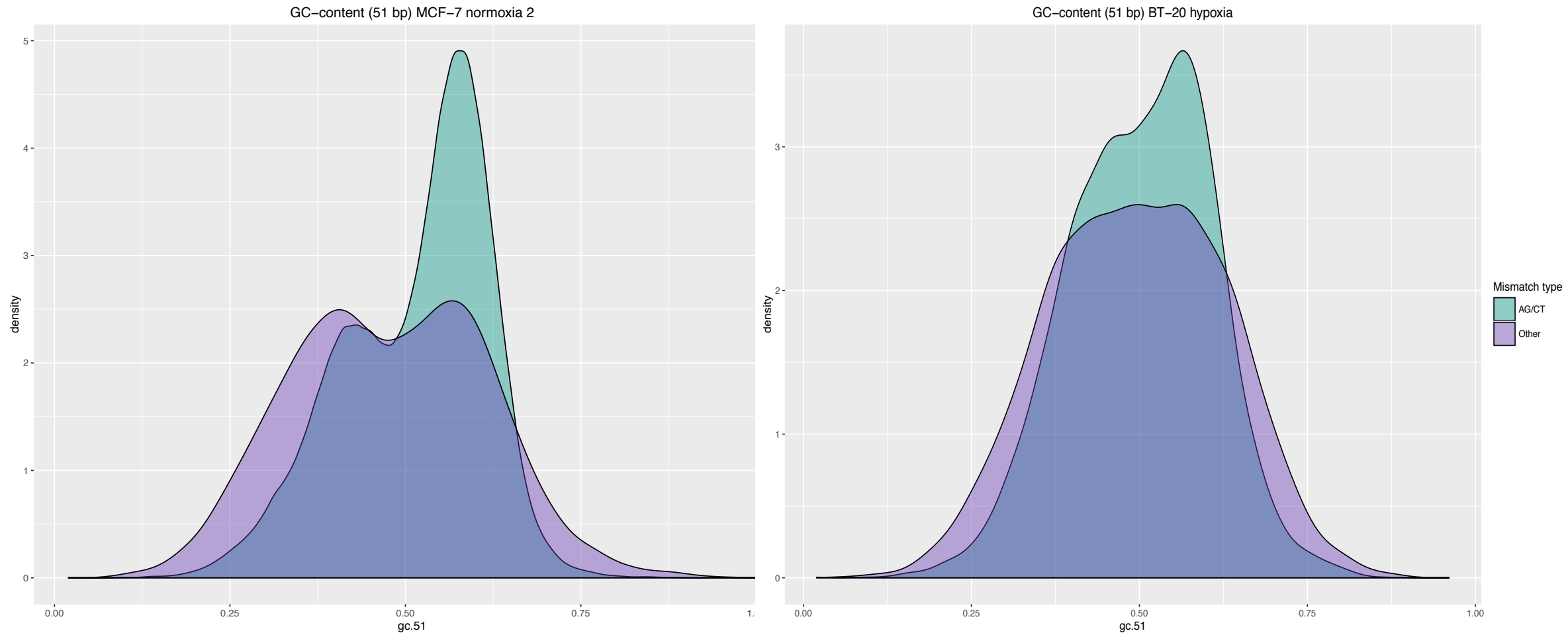
# Data

- MCF7 cell line RNA-Seq (2 conditions x 3 biological replicates)
- BT20 (2 conditions) and K562 (1 condition) cell lines RNA-Seq
- Existing databases (RADAR and DARNED)

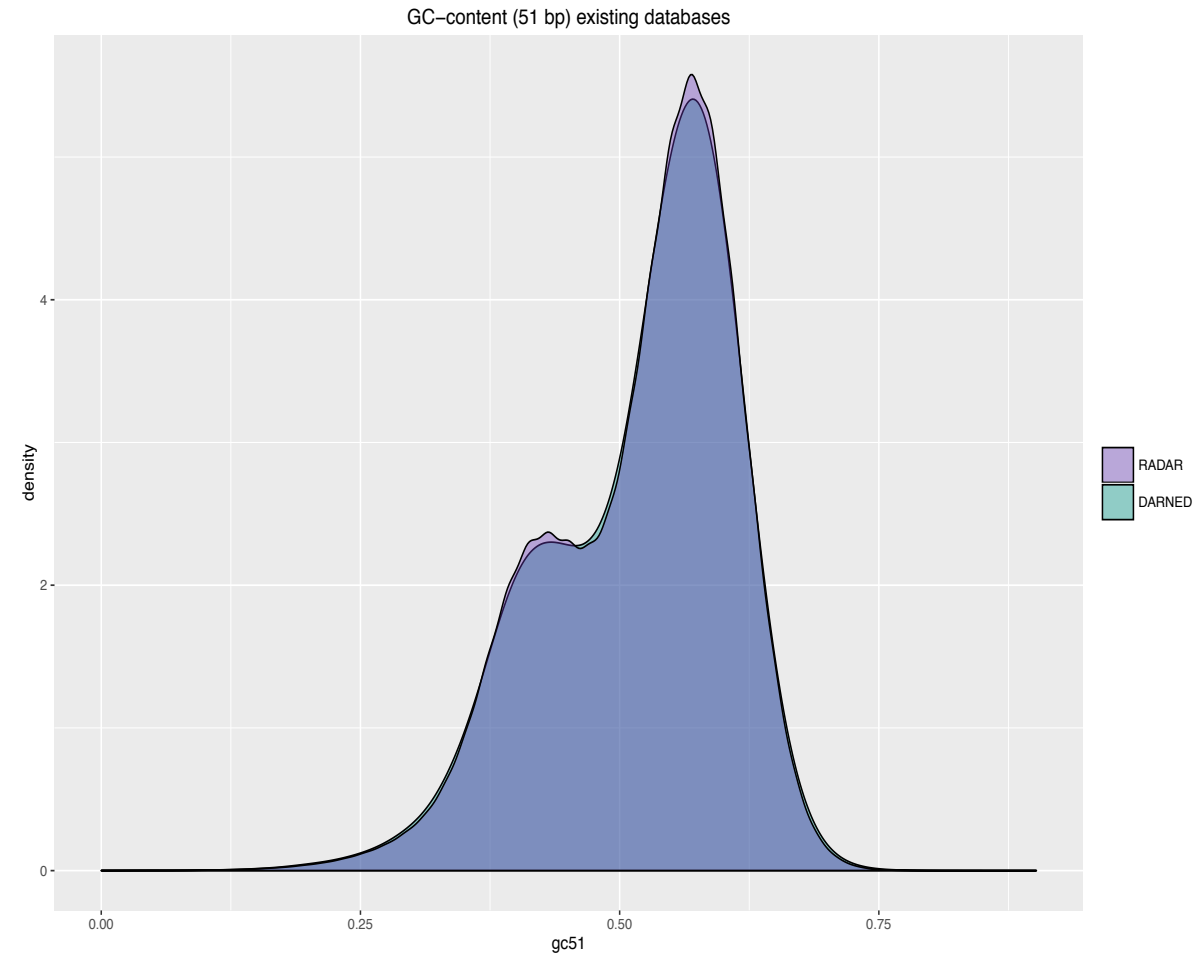
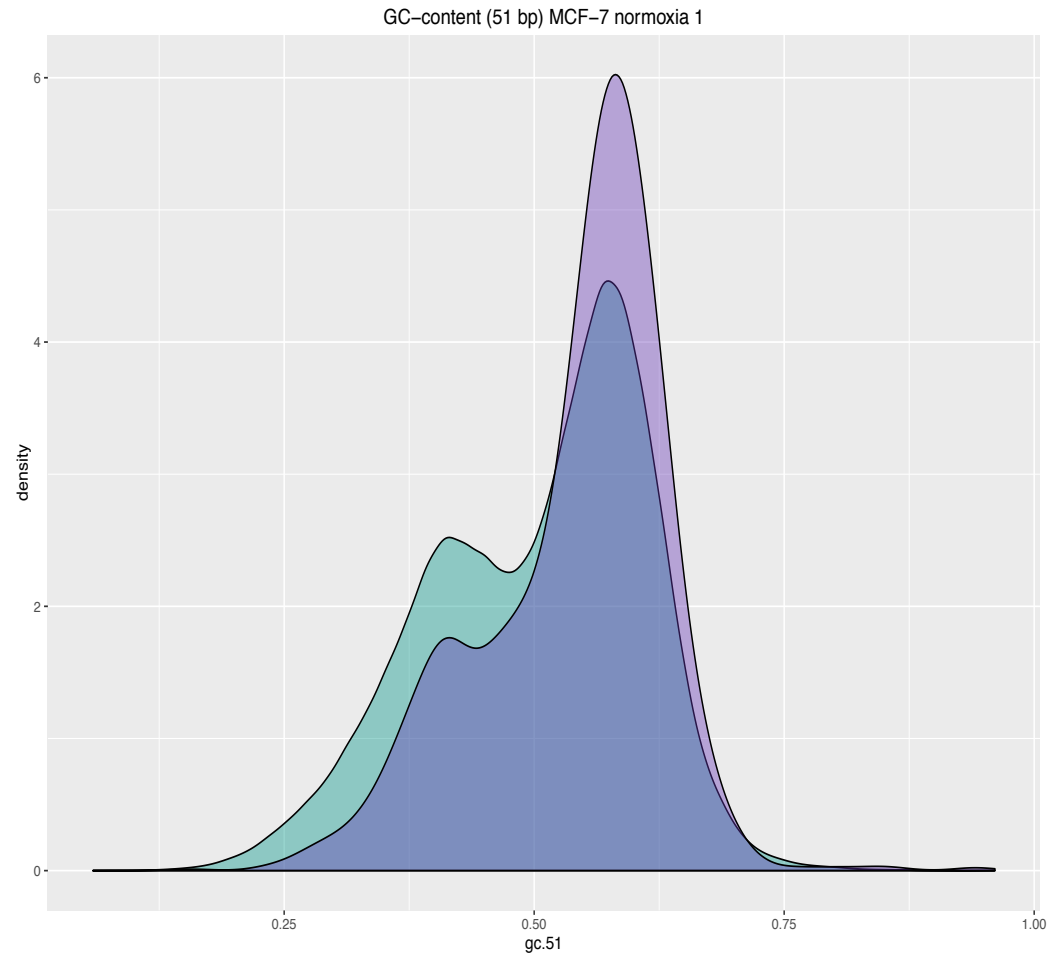
# Fraction (BT-20)



# Local GC-content



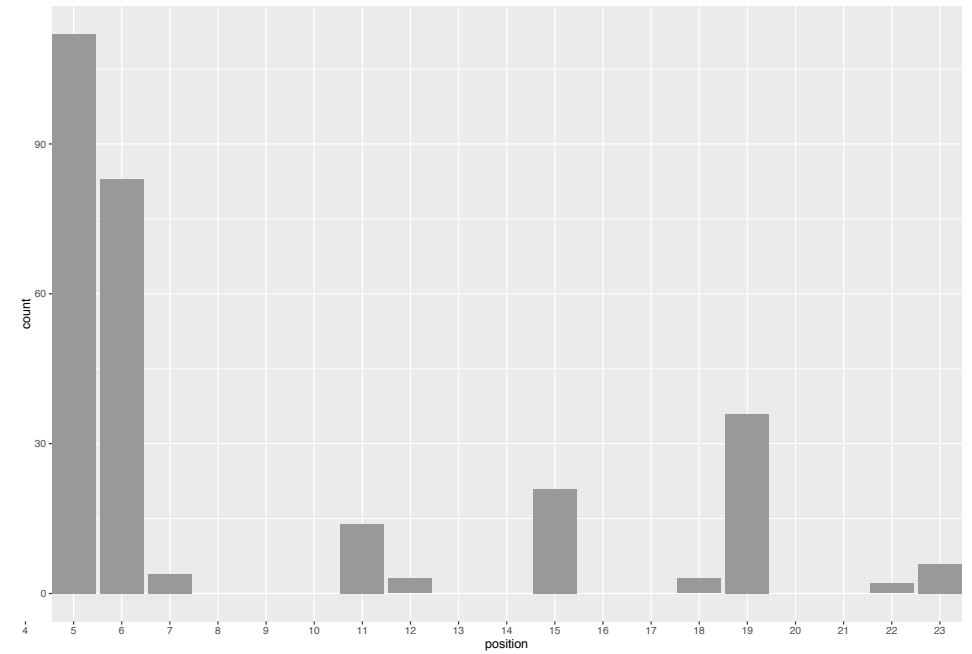
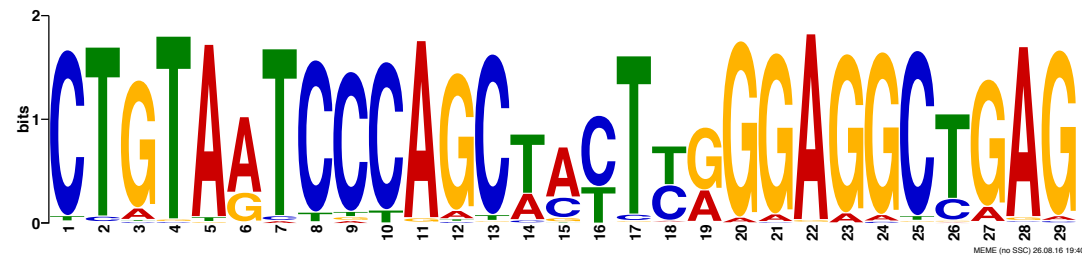
# Local GC-content



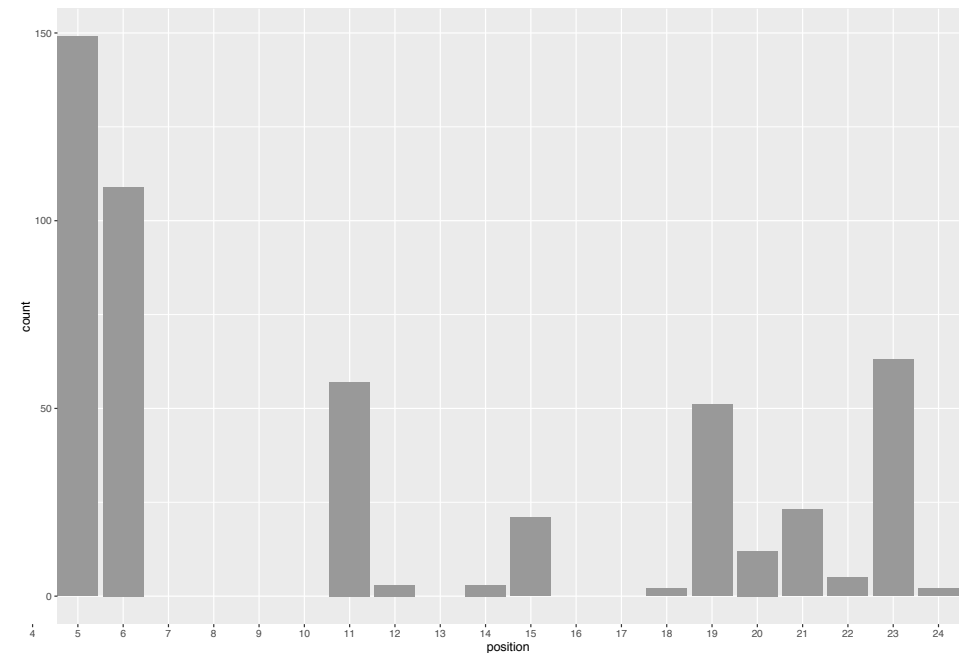
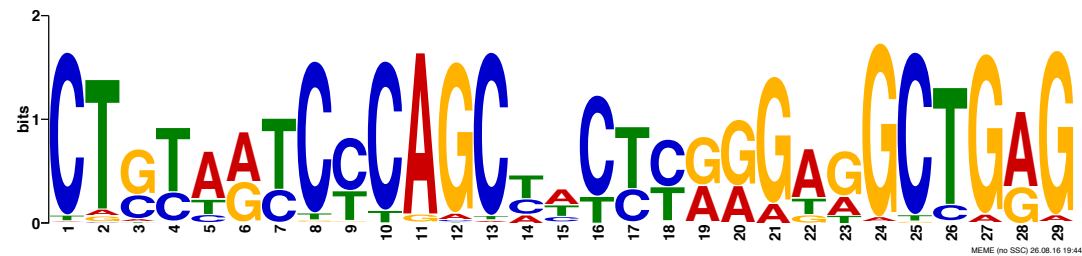


# MEME motifs

Normoxia, E-value: 1.9e-2786



Hypoxia, E-value: 3.1e-3673



# Plan

- Secondary structure
- Structural entropy

THANK YOU!