2019 Library Design

* library\_2019$combinatorial
  + we want to investigate all combinations of W (aromatic) and D (acidic) for 1 to 11 spaces
    - W, D (21 sequences)
    - WW, DD, WD, DW (22 sequences)
    - 211 sequences or 2,048 sequences
    - 4094 sequences total
  + Glink­: we used G as a filler to link the activity module to DNA binding domain
    - Glycine is proposed to be neutral to functionality
  + We want to use this portion of the library for broad analysis and machine learning
* library\_2019$W\_first and D\_first
  + we want to compare aromatics followed by acidics and vise versa and the effects on functionality.
  + 10 sequences each for a total of 20 sequences
    - GGGGGGGGGGGGGWWWWWDDDDDD
    - GGGGGGGGGGGGGDDDDDDDWWWW
* library\_2019$set\_g\_spacing
  + we start with one aromatic and one acidic and progressively space them out with G
    - GGGGGGGGGGGGGWD
    - GGGGGGGGGGGGWGD
    - Ect all the way out to 9 G spaces
  + we did the same thing for 2 aromatic and 2 acidic ect all the way out to 5 aromatics and acids
  + we are trying to understand optimal distance between aromatic and acidic residues to create functional tADs
  + 50 sequences in total
* Library\_2019$g\_in and g\_out
  + We are investigating if tADs perform better with one large functional region or two smaller functional regions.
    - 40 sequences total
  + G\_out
    - WDWDWDWDGDWDWDWDWD
    - WDWDWDWDGGWDWDWDWD
    - WDWDWDWGGGWDWDWDWD
  + G\_in
    - GWDWDWDWDWDWDWDWDWG
    - GGDWDWDWDWDWDWDWDGG
    - GGGWDWDWDWDWDWDWGGG
* Library\_2019$terminal\_wd and internal\_wd
  + We are trying to see if the position of the activity module is essential for functionality
    - Is it better to have the activity module close to the DBD or give it more spatial freedom by putting them at the end terminus?
  + Terminal\_wd
    - GWDWDWDWDWDWDWDWDWDW
    - GGWDWDWDWDWDWDWDWDWD
    - NOTE: these sequences double as the sequences for length
  + Length: we are trying to find the essential length for tADs and we are extending the active module out to 20 spaces
  + Internal\_wd
    - WDWDWDWDWDWDWDWDWDWG
    - WDWDWDWDWDWDWDWDWDGG
    - WDWDWDWDWDWDWDWDWGGG
* Library\_2019$oneR, twoR, threeR
  + We know basics are detrimental to functionality, but we have several sequences with basics that maintain functionality.
    - We are trying to better understand the rules for when R is the most detrimental
  + One way to describe these sets is ‘R scanning’ where we do a flythrough in which R gets substituted at every position (in both the G linker and the activity module)
  + RGGGGGGGGGWDWDWDWD
  + GRGGGGGGGGWDWDWDWD
  + GGRGGGGGGGWDWDWDWD
  + Ect