- Announcements
- Quiz
- Pre-lab Lecture
  - Today in Lab: M2D3
  - Shape effects in DNA gels
  - Diagnostic digest preview
  - Introductory statistics

why? class-wide
WT + ref. mutants

Announcements

- Quiz, then load gel, and then pre-lab lecture
  - don't add loading dye to entire reaction!
- Purpose of shifting K_D for a sensor
  - what would happen if you used fluo-3 w/dendrites?

    everything bright /always on

  - what about purpose of shifting cooperativity?

    steep vs gradual change useful in different biological contexts

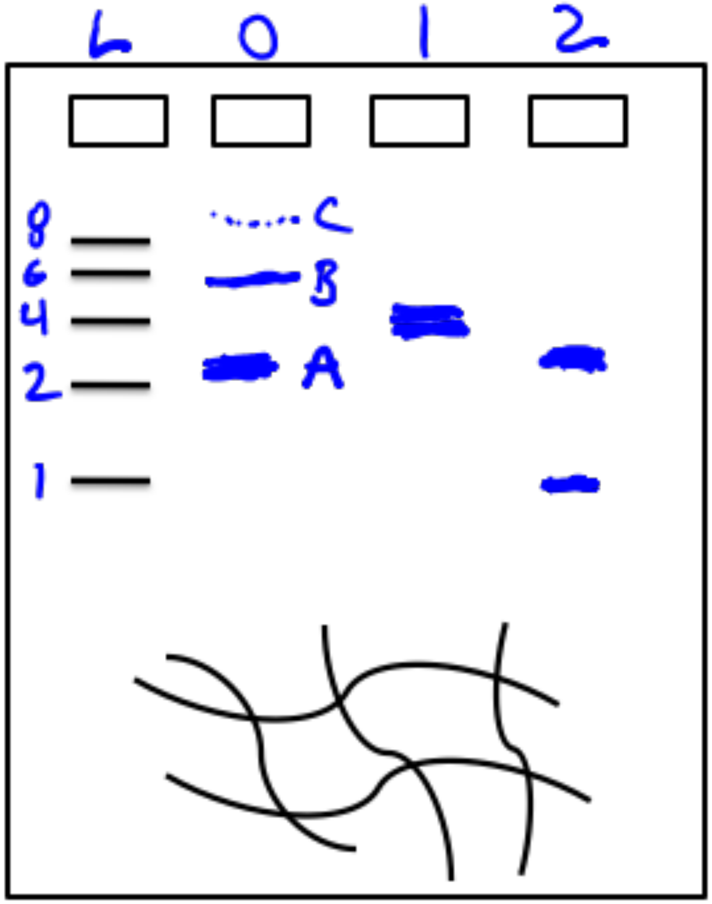
Today in Lab: M2D3

- Set up gel of digested SDM product
  - ~~mark your area with colored tape~~
  - runs 45 min, we will photograph and post it
- Meanwhile…
  - pre-lab lecture
  - label tubes two ways (sticker top, sticker/marker side)
  - start FNT! M2D4 tends to run long – prepare
  - practice analysis optional for now
- Briefly analyze gel for DNA size and amount
- Finally, bacterial transformation – be gentle!
  - includes 30 min incubation step

Polymerase error rates

- *Taq* polymerase ~ 1 in 10^5 # bp
  - Standard version has no proofreading capability
    ↳ exonucleases
- *Pfu* polymerase ~ 1 in 10^6
  - Standard version requires longer extension times

DNA EP: shape-dependence



e.g. 4Kbp

Plasmid versus linear samples

1-cut: linear DNA runs w/ ladder

2-cut: DNA must sum to whole
eg., 3+1 Kbp

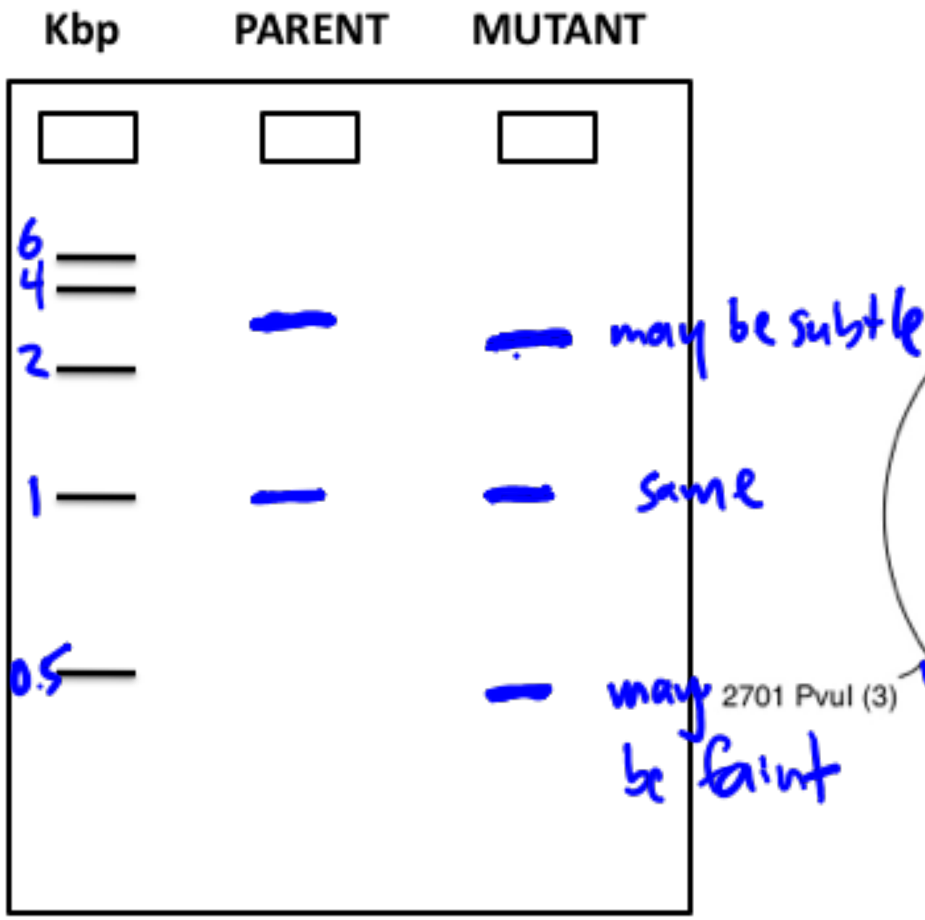
uncut: A) supercoiled - faster

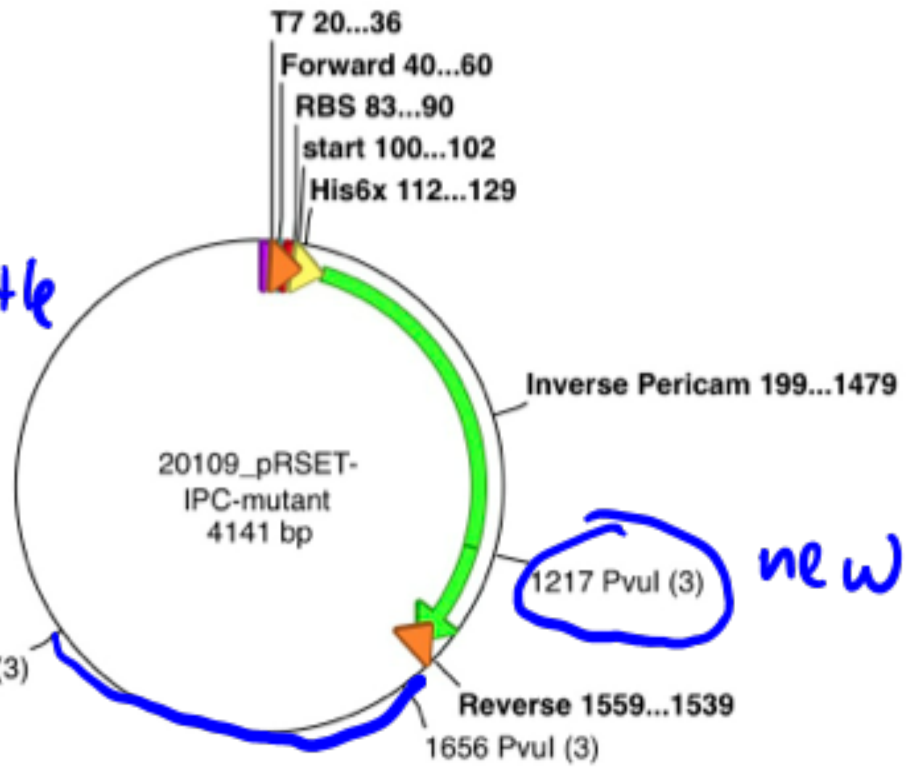
B) relaxed or nicked circular - slower

C) maybe high MW concatemers

Remember to wear nitrile gloves.

Diagnostic DNA gels





Statistics review: basics

- Essential concepts: standard deviation (*s*), mean ($\bar{x}$), sample size *n*, degrees of freedom *DOF*
- Normal (Gaussian) distribution

1 *s* includes
68 %
of the data

x-axis: measured value (e.g., intensity)

y-axis: # or fraction of samples w/ that value

Confidence intervals (CI): principle

- Example: $\bar{x}$ = 60 (sample/measured mean) and 95% CI calculated to be ± 3 from real data
- Meaning: 95% of the time our population/true mean μ lies in the range, here 60 ± 3 as n
  - subtly different from 95% likely that the range 60 ± 3 contains the population (true) mean μ, which we can't say
- 90% CI: $\mu = \bar{x} \pm a$ where $a < 3$, $a > 3$, or $a = 3$?

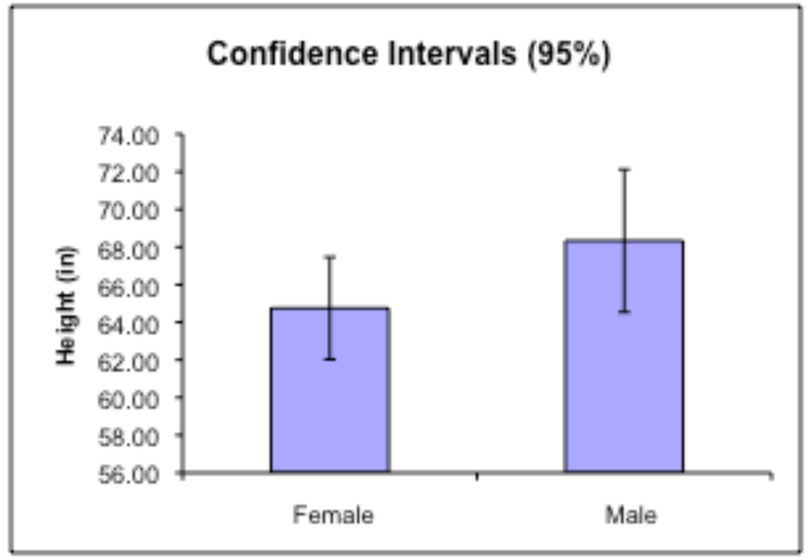
trade-off between precision and confidence

- Consider betting example
- What about n? as n increases for given C.I., more precise

Calculating confidence intervals (CI)

per a population

$$\mu = \bar{x} \pm \frac{t\,s}{\sqrt{n}}$$



- *t* is tabulated by DOF vs CI%
  - DOF = *n* - 1 why? Σ errors $(x_n - \bar{x}) = 0$ → constraint
- In Excel, us *TINV* function
  - input *p*-value = (100-CI)/100

if C.I. = 95%
$p = 0.05$

Introduction to t-test

- Every statistical test
  - Has assumptions
  - Asks a specific question
  - Requires human interpretation (GIGO)
- Some t-test assumptions
  - normal distribution (cf. Mann-Whitney test)
  - equal variances (type 2 in Excel; type 3 unequal)
- Posing a question are mean male and female heights different at a confidence level of 95%?

Calculating t-test significance

$$t_{calc} = \frac{\bar{x}_1 - \bar{x}_2}{s} \sqrt{\frac{n_1 n_2}{n_1 + n_2}}$$

pooled s

DOF $= n_1 + n_2 - 2$

t_{table} listed by DOF vs CL (confidence level)

- If t_{calc} > t_{table} difference *is* significant at that CL
- In Excel, us *TTEST* function
- Excel returns *p*-value → confidence level (CL)
- 1-tailed vs. 2-tailed test

e.g., $p = 0.01 \rightarrow CL = 99\%$

1 - one-sided; hypothesis in advance

2 - full distribution; no a priori hypothesis

Practice assignment for today

- Female heights: 61, 65, 61, 68, 65, 63, 61, 67, 60, 63, 64
- Male heights: 72, 72, 70, 65, 72, 69
- Calculate 95% CI for each mean
- Plot means on bar graph with CI error bars
- Apply t-test to the means
  - for multiple comparisons, ANOVA is better
  - comparing many means requires correction
  - remember, $p = 0.05$ means 1 in 20 false positives!