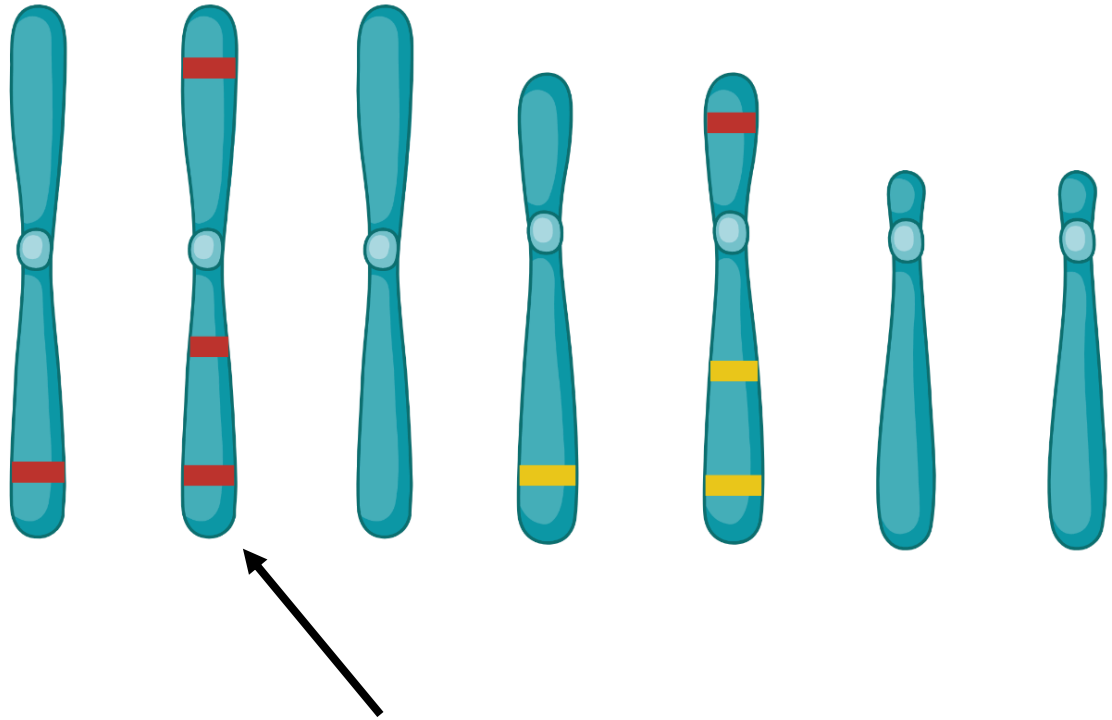
Monte Carlo methods

Named after the resort in Monaco well known for its casinos

for next week please install the packages for R markdown

You will eventually be faced with questions where you can't find a clear statistical test that will be appropriate to your data.

Are these beetles randomly distributed, aggregating or avoiding each other?



Does chromosome 2 have more "red" genes than I would expect by chance?

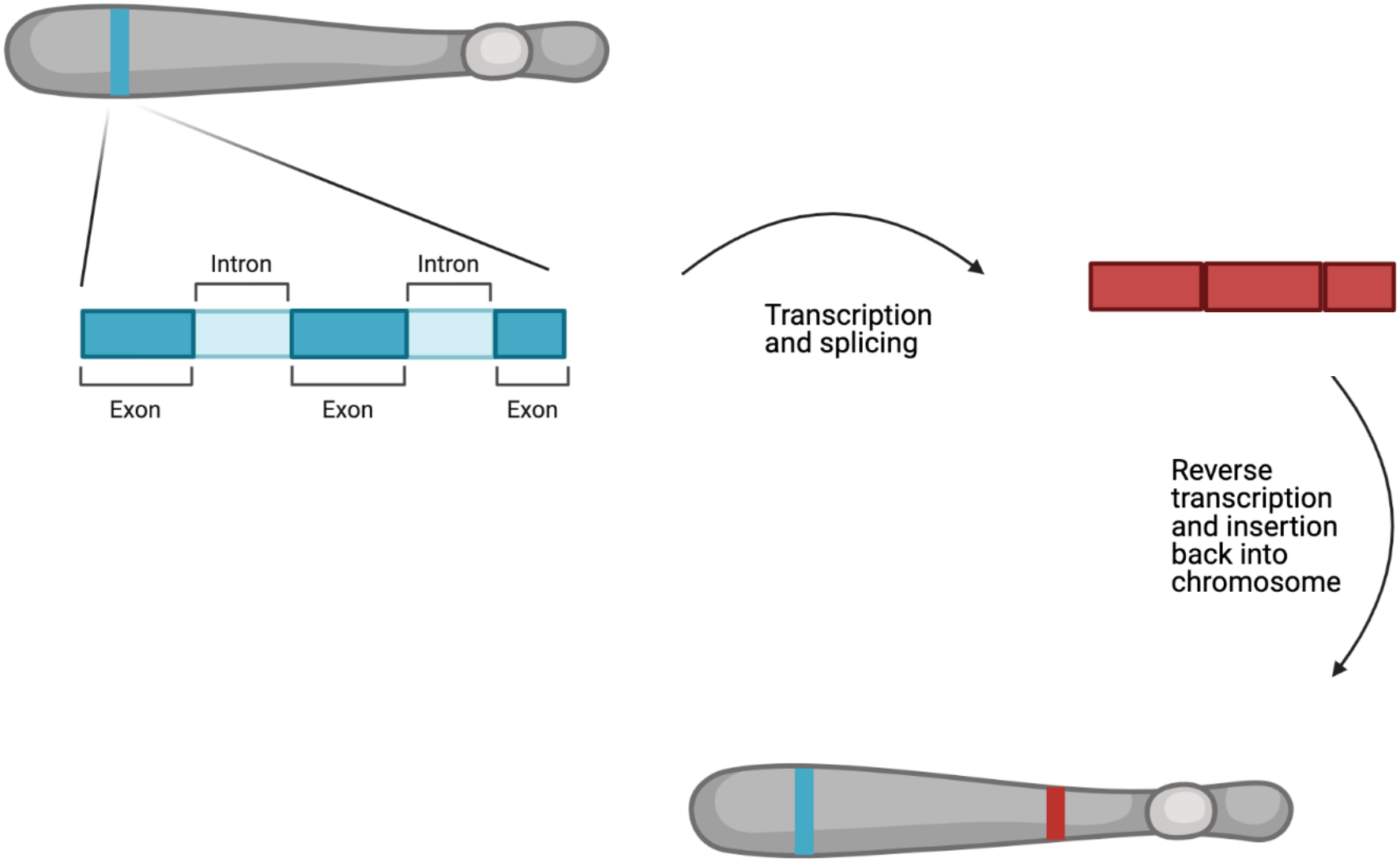
You will eventually be faced with questions where you can’t find a clear statistical test that will be appropriate to your data.

From last week although we didn’t discuss it in depth we mentioned that this is the approach to see if an f-statistic with an ANOVA was significant. We know we can’t really depend on the ANOVA when we are using data from multiple species so we simulate data and calculate a statistic many times and compare our empirically derived f-statistic to this simulated one to assess significance.

Monte Carlo methods offer an approach to answer difficult often complex questions.

Lets look at an example with retrogenes.

Lets look at an example with retrogenes.



If some genes have variation that is beneficial to one sex but harmful to the other fitness could be increased in two ways.

1) make a duplicate of a gene and express one version in males and one in females
2) make a duplicate of a gene and move it from an autosome to an X or Y chromosome

Additionally in many species the sex chromosomes do not get expressed during meiosis if you have a gene that can benefit males during spermatogenesis moving this type of gene from a sex chromosome to an autosome will be beneficial.

From this we might hypothesize that we should see more retrogenes on sex chromosomes and possibly more parental genes on sex chromosomes than we would expect by chance.

From this we might hypothesize that we should see more retrogenes on sex chromosomes and possibly more parental genes on sex chromosomes than we would expect by chance.

1) Size of the chromosome
2) Number of genes on each chromosome

We identified retrogenes in a Tribolium beetle that has a recent fusion of chromosome 2 (lg2) to the sex chromosome. This means that the X in this species now include everything that is in lg2 as well as lgX.

	lg1	lg2	lgX	lg4	lg5	lg6	lg7	lg8	lg9	lg10	Total
GeneNumber	3254	2941	3705	1507	1543	1287	994	684	325	212	16512
PhysicalSize	45	41	38	30	20	15	14	11	12	9	235
Parents	22	45	39	9	8	7	8	3	1	0	142
Daughters	24	40	35	15	7	6	5	4	5	1	142

Do we have more parent and daughter retrogenes than expected by chance?

1. First decide what we want to test
    a) the number of parents on lg2
    b) the number of daughters on lg2

2. Create an object to store a null distribution of these values

3. Simulate an expected number of parents and and an expected number of daughters on lg 2.
    a) when simulating account for chromosome size and gene number

4. Repeat step 3 thousands of times.

5. Compare our observation to the null distribution to calculate a p-value for our observation.

R