* **Abstract**
* **Introduction**
* **Methods**
  + **Model Structure**
  + **Subsampling Methods**
    - *All subsamples are taken from the 1018 successfully genotyped hair clusters, from the original pool of ~2000 hair clusters.* In order to avoid confounding number of successful samples with subsampling type, subsamples were taken from the successful subset of hair clusters.
    - *Subsampling methodology ranges from representative and random to non-representative and uniform with respect to site-sessions.* Both strategies are used in genetic CMR studies and were chosen as such.
    - *‘SimpleRandom’ was most representative of the actual data, and “Spread.two’ was most uniform with respect to site-sessions.* See accompanying figure explaining the relationship between these subsampling types.
    - *In ‘Subsampling trials’, subsamples varied only by the subsampling treatments described above and n was fixed at 450; in ‘Size trials’, subsamples varied by size and by either “SimpleRandom” or “Spread.one” subsampling type.* Former used to identify differences between subsampling types, and latter analyzed two with distinct differences.
  + **Simulation Structure**
    - *Each simulation, regardless of trial type, involved a subsampling step and a model fitting step, using five spatial MR models with differing covariates.* After a subsample and fitting of five models, density estimates as well as model objects were saved for later analysis.
    - *We used parallel processing to cut down on processing time.* Parallel processing allowed the fitting of a single model to take about 25% of the time.
    - *Simulations moved stepwise through the four subsampling types perpetually for 6 weeks.*The simulations occurred perpetually for 6 weeks on a (Insert Sony VAIO specs) computer, where the fitting of a model took an average of X about minutes.
* **Results**
  + **Subsampling**
  + **Sample Size**
* **Discussion**
* **Conclusion**

**General Thoughts:**

Overall - take home message, because bears tend to leave many samples at a given site, simplerandom and stratified samples contain lots of redundant samples, where spread one and spread two do not contain lots of redundant samples. As such, it makes sense that g0bTRUE is more negative for simplerandom and stratified, because that means recapture probability is lower. (g0 + g0btrue = recapture prob, eg .06 + -.013 = .47 recap).

g0 is allowed to vary up or down based on the constraint that it must be between 0 and 1, naturally. This is different from traditional distance sampling models, where the detection function have an intercept that is less than 1. In our case, sigma was not manipulated, so it varied based on sex in all cases, which is what we’d expect biologically – note to self, look at the values for sigmaSEX

Later SECR models are different somewhat because they don’t account for movement implicitly, it gets rolled up into detection probability, which makes it resilient to things that move a lot or a little

“One of the key features of SCR models is that the point locations are latent, or unobserved, and we only obtain imperfect information about the point locations by observing individuals at trap or observation locations. Thus, the realized locations of individuals represent a type of “thinned” point process, where the thinning mechanism is not random but, rather, biased by the observation mechanism” – scrbook ch1 pg 16

In models that don’t have t as a covariate, all of the biological time response gets wrapped in in the behavior covariate (since, by the end, most of them have been detected at least once prior). Same with the model w/o time, behavior wraps up the time response.

Less n in general, less accurate estimate of an animals AC? (good example in royle book ch 5.3.2)

Its already a non-random subsample to begin with, so even the full models are missing some of the original captures that biologically happened but are not in the full dataset

TODO : estimate vs size

plan for simulation - spit out secr simulation, then repeat rows and whatnot in order to simulate the effects of redundant samples. also can look at data sets with behavior effects and without, and then look at size and spread vs simple random

also, lots of studies argue that with TOO high n, there is a chance of erroneously identifying an individual when there is none, so maybe true n is actually lower. Maybe it would never actually be asymptotic, if the error rate was high enough.