Potential outline for a completely *in silico* paper.

1. Intro / Motivation
   1. Talk about experimental diversity (or lack of) in expression datasets
   2. Potential to discover new biology by activating paths that were previously dormant (who knows what happens when those paths are activated?)
   3. Review Bayesian 1.0 paper as background; explain how aijs can be the foundation for detecting always-inactive genes
      1. Unless we find we need to base this on the Bayesian 2.0 methods. In which case we’d have to get Bayesian 2.0 published before even this *in silico* paper.
   4. Maybe mention the need for two somewhat separate metrics: inactive vs. active, and not-changing-state (or maybe this belongs in Methods)
2. Methods  
   1. Metric for determining inactive vs. active (Ln of S-statistic)
      1. Theoretical definition (S-statistic is the probability the gene is always-inactive, assuming experiments are independent and aijs are calibrated)
      2. Perhaps some discussion of the distribution of Ln(S-statistic)
   2. Metric for determining not-changing-state (Bayes Factor and its derivatives)
      1. Theoretical foundation
      2. Perhaps some discussion of the distribution, or what values can be considered strong evidence
   3. Explain procedure for simulating data, if different than in Bayesian 1.0
   4. Explain data source(s) for real data, if different than in Bayesian 1.0
   5. Explain the idea of subset-dropping for validation
3. Results  
   1. On simulated data
      1. How well does the Methods (A) metric do at determining inactive vs. active
      2. How well does the Methods (B) metric do at determining not-changing-state
      3. How well does the combination of metrics do at identifying always-inactive genes
      4. (Maybe) Pick a random operon and use subset-dropping to get rid of all of its active experiments; hopefully show it’s now identified as always-inactive
   2. Subset-dropping on the real data (whatever real data we’re using). The general pattern is, drop a set of experiments; discuss which genes looked changing-state on the full dataset but look always-inactive on the subset; hopefully make a biological case that this makes sense. Some ideas for subsets to think about dropping are listed below
      1. Drop experiments containing arabinose
      2. Drop experiments containing yeast extract
      3. Drop experiments *not* containing yeast extract
      4. Repeat with other nutrients / media conditions
      5. Drop experiments in a given compendium or set of compendiums; or, drop all experiments *except* those in a given compendium. (Point out how in real life it is easy to get a group of experiments with poor diversity.)
   3. Now, talk about our actual results (rather than just validation). What genes/operons/pathways/etc are indicated always-inactive on our full dataset? Is this biologically reasonable? Point to future work to explore these.