Proposal (Draft 1)

* Needs to be included in the proposal:
  + A specific hypothesis to be tested by aims (experimental)
  + Subsections:
    - 3 specific aims
    - Background and Significance
    - Preliminary Data
    - Research (experimental) designs and methods
      * Include possible trip-ups
    - Biblio
* Background
  + What is REST?
  + Why are we interested in REST?
  + What is its historical role in cancer?
  + What is its historical role in general?
  + How can we study REST differently than other people?
  + What can we learn from data integration?
  + What is context mediated changes
  + What are transcription factors/histone modifications/methylation etc
  + What are the cell lines being used and why?
  + Where else are we looking at
  + What other data (other rest bound cancer shit) are we looking at?
  + Why should we study REST and context mediated changes and why?
  + What is RESTs function in cancer and other diseases
  + How will having an greater understanding of it, though my methology better human kind?

“The transcriptional repressor element 1–silencing transcription factor (REST) has been implicated in both oncogenic and tumor-suppressor functions” -- ([Reddy, Greco et al. 2009](#_ENREF_6))

* Prelim data:
  + See stuff from genome bio papers, rest binding been shown to corr with hist mods,
  + Rest is known to have many hist mod changing fns assoc
  + Only has significant changes to some hist mods (see older papers) (+fig2 gpdyz)
  + Only studied really in promoter (true?)
  + Known to bind some genes that are important for other cancers fig1 gpdyz
* AIM 1 (Epigenetics of cancer, with a focus on REST)
  + Hypothesis:
    - Many epigenetic modifications are dysregulated in cancerous tissues (we know this through studies of hist mods, tfs and methylations)
    - REST has been shown to be dysreg in various cancers? At certain genes this is true? It is dysregulated in T cell lymphoma in a context-specific manner?
    - REST is responsible for some of the dysregulation in this kind of cancer (thought because it has high binding in this cell line and in other cancer tissues) 🡪 HOW? (what tf are bound in other cancers that are known to be bound/controlled by rest are the bound here?)
  + Data
    - TFs
      * CBP
      * CTCF
      * REST
    - Hist mods
      * H3k4me3
      * H3k27me3
    - RNA Polymerases
      * 2
      * 3
    - Transcriptomics
      * RNA-seq
      * miRNA array
    - methylation
  + Questions
    - How is rest dysregulated in cancer?
    - What are the epigenetic differences between the two cell lines???
    - How is rest correlated with other stuff dysreg in cancer?
    - What are rest bound sites that are bound only in cancer or in noncancerous tissues?
    - Group of REST interacting cancer genes, deal in jurkat and cd4+?
    - How does REST have both oncogenic and tumor suppressor roles? ([Coulson, Fiskerstrand et al. 1999](#_ENREF_2); [Zhao, Gjoerup et al. 2003](#_ENREF_12); [Westbrook, Martin et al. 2005](#_ENREF_11); [Majumder 2006](#_ENREF_5); [Su, Gopalakrishnan et al. 2006](#_ENREF_9); [Weissman 2008](#_ENREF_10); [Reddy, Greco et al. 2009](#_ENREF_6))
  + Methods:
    - Compile all data that exists in both cell lines
    - Find things that are significantly different in one or the other (stats)
    - Find subgroups of genes that are of interest
    - Stratify into rest bound/rest unbound/ and rest not pred to bind expr matched
    - Find things that are different in the two cell lines that correlate with rest bound and not others, (using them as controls) [make sure to use randomization to really check stuff]
    - What are the epigenetic differences between the two cell lines???
    - Use covariance or something to think about if things are going on at once, or something, are these things being indep reg through diff fucntions of rest or through one thing which is getting passed down.
    - Separate sites that are diff regulated and look at other transcription factor binding sites in RE1s (MEME/TRANSFAC)
    - What are the function of these REST bound in one or the other cell genes (esp in Jurkats)
    - Since rest interacts with RNA pol, is there something funky going on there, with RNA pol or RNApol 3, in cancer cells specifically and not noncancer? Can we show an interaction or something. Is there an interaction only in cancer cells, that does not exist in noncancer or something?
    - Find genes that are dysreg in other cancers, and mod by REST what are their functions? And how are they dysreg in T cell lymphoma
* AIM 2 (REST context mediated function and analysis)
  + Hypothesis:
    - REST regulates in different tissues differentially. How is this happening. Hypothesize that context specific regulation going on, with histone mods, methylation, and other TFs, seq data (loc, features (MEME)) influencing RESTs accesss to the genome.
    - These context specific information, help to control if REST down or up regulates different genes.
    - REST has many cofactors not all of which are known? What new functions are correlated with REST (can we find them, and in AIM 3 confirm that they are REST dependant) [tcell and neuronal?]
    - RNAi has been shown to interact with stuff. Also variant isoforms of genes with poison codons, have been shown to regulate protein quantity through NMD. This is going on and helps to influence REST binding and ultimate control over the genome. (what other functions does REST have with non transcribed genes?) Does REST binding influence miRNA landscape etc? RNApol3 access etc?
    - The landscape of the genome and of transcription factors is not one dimensional. Rest must interact in a landscape of many other transcription factors, and the binding of those might play a role in RESTs function and activation/repression
    - Are these epigenetic modifications that interact with the function of REST etc different in T cells (where REST is repressed) vs neuronal tissues, and what epigenetic modifications help to control REST binding and regulation of transcription in both cell lines (can we ascribe this differential activity to REST or to other transcription factors that are binding in these other tissues (corr with MEME)
    - What is the role of cofactors of REST? Information on coREST in neuronal tissue, how does its cofactor help to regulate the binding and activity of gene networks etc?
  + Data (Cd4+ T cells)
    - HDACs + HATs (9)
    - Methylation information
    - TFs
      * HGMN1
      * STAT6
      * P300
    - RNAseq
    - DNAseq
    - miRNA array
    - Histone modifications
      * 18 acetylations
      * 20 methylations
    - Pols 2 + 3
    - REST chipseq
  + Data (Forebrain)
    - Histone Modifications
      * H3K4me3
      * H3K4me1
      * H3K27me3
    - RNAseq
    - DNAseq
    - P300
    - Pax6
    - REST
    - coREST
  + Questions
    - How is REST differentially bound in neuronal vs other tissues and how does it propagate into activity/epigenetics of those genes(this is already extensively studied I believe)
    - How does REST influence the epigenetic state at each binding site? [both in CD4 and nueral tissues]
    - What sequence features are associated with each group, location/sequence (esp in ncRE1s) or nearby
    - How does the epigenetic state influence REST binding?
    - Are there other transcription factor interactions that can be predicted by the sequence that REST is binding to
    - Look at cRE1s and ncRE1s (better ways of stratifying groups?)
    - What are the functions of genes that have REST activating vs repressing them?
    - How does REST binding effect methylation?
    - What is the interplay between REST and RNAi (production and with RNAi downregulating REST or cobinders etc)?
    - Does rest have differential effects at different parts of the genome ( mostly the promoter region has been studied thus far, what is going on in the exons or introns or intergenic regions for example?)
    - Are there changes in the histone modifications at regions where there are miRNA and bound REST? (maybe not this? Dunno how to test, esp since not really a good way to know if miRNA bound)
    - What role does coREST play to control RESTs function, and how does it influence histone modifications etc.
    - Does coREST have its own functions outside of REST, what is going on at coREST bound sites? (what are coREST bound sites, I do not think there is any coRESTseq data in exsistance)
    - Since we have so much data what else can we look at (prolly not something that will make the cut, but are there any interesting patterns/networks/functions of genes that are highly transcribed without any histone modifications for transcription, what about stuff that is lowly transcribed but hist mods would suggest it should be more highly transcribed?)
    - We have information about transcription of REST and REST isoforms, does the amount of noncoding isofrom of REST correlate in some way with number of bound sites as normalized to the number of reads or something, since it has been reprted the noncoding isoforms regulate coding isoforms and protein levels?
  + Methods:
    - Find rest binding sites in both tissues, find sites that are only bound in neuronal and sites that are only bound in cd4+ t cell. Separate out these sites, what are their functions, what are their expression, is this statistically significant. What is going on in the genes that are bound in both cell lines, do they have different epigenetic modifications recruited? Are their activities statistically different?
    - Stratify into groups of REST pred (RE1 site) REST bound, REST pred no binding (ctrl 1), and expression matched genes from neither of the first two groups (ctrl 2). Ctrl 1 will help separate out function of rest from sites that have sequence features that recruit other TFs or whatever that are similar. Ctrl 2 will help separate out changes due to rest binding vs stuff that is correlated with whatever the expression is at that level. (learn more about how this is done) (see genome bio paper deyou) Look at ctrl 1 vs bound to see how epigenetic state affects rest binding [build histone modification occupancy profiles across all groups, cluster and compare (**kolmogorov-smirnov tests**?) quantify rel. betwn hist mods and REST, highest REST dependency🡪 linear/ other model, statistical analysis)]
    - find seq features in each group corr? [also genes with different expression levels have diff clustering groups, genes with different functions cluster differently, or have diff hist mods?]
    - PCA and other things like that (**Baysian inference network**), how can you explain the majority of the variation in group 1 through comb. Patterns of mods/tfs/etc (what allows REST to access genome?)
    - MEME TRANSFAC
    - ?
    - DAVID/GO/IPA groups [clusters]
    - Look at stratified groups, corr methylation? If there are methylation changes perhaps there is a methyl transferase (dnmt) somewhere in big protein bundle that is REST and stuff, that has not been IDed)
    - PCA or something SVM etc to find miRNAs that play a large role in REST binding etc. Also how does REST binding influence Pol3 access to genome etc (group 1 stuff) literature search for RESt related miRNAs or miRNAs in dbs that bind to REST and perhaps influence its production
    - Epigenetic mods etc different at diff parts of the genome? Rest binding to different kinds of sites at different genomic locations? Strat into groups by location, and see. Corr? Keep ctrls 1 and 2 from before but look in spec loc instead of genome wide.
    - ?
    - like before sep into REST and coREST bound from group 1. Keep cntrls. Also find genes that are being reg only by coREST, function? GO/IPA/DAVID also what is the RE1 status of those? What about the deal with binding with other TFs TRANSFAC/MEME, more.?
    - Functional analysis GO/DAVID/IPA
    - Find number of rest bound sites in both tissues, find production of transcript.
    - **Search for RE1 sites with mismatches in non cRE1 or ncRE1 regions**
    - **Look at REST on epxr up, and expr matched group. Perhaps other TFs here or something?**
    - **Search for mods that have high corr with RE1s or REST, but low corr with the other, maybe are groups that if only RE1, are being bound by other TFs, if only REST have other co-shit going on or something?**
    - **Find hist mods that corr with enhancers? Are RE1 epigen landscape diff there?**
    - **What about groups that have RE1 sites in their promoter and elsewhere, in the gb or something? Extra silenced?**
  + Trip ups
    - Finding the right scheme for analyzing the data, might take a while, but am hardy, will figure it out! ☺
* AIM 3 (Biological Validation of aspects of AIM2)
  + Hypothesis
    - The changes in epigenetic modifications as observed in aim 2 are real, reproducible, and dependent on REST binding
  + Materials
    - Antibody against REST
  + Questions
    - Are the changes shown in AIM 2 really attributable to REST binding?
  + Methods
    - Transfect cd4+ t cells with REST (this may be a problem primary cells/primary cell lines sometimes are) (if cd4+ t cells do not work, transfect jurkat cells ) discussion of how this changes interpretation, to be inserted here
    - There is a human siRNA for REST used in Reddy – construction of siRNA vectors ([Greco, Smirnov et al. 2007](#_ENREF_3)) stable transfection ([Corcoran, Trzaska et al. 2008](#_ENREF_1))
    - Quantify REST expression using western blots
    - Knock down REST in cells (control with mutant siRNA and nontransfectants)
    - Verify knockdown with qpcr and westerns (see Greco and reddy)
    - Take the cells that have been kd for REST and do ChIP-qPCR assay at loci identified in AIM 2 as being significantly effected by REST binding
    - Crosslink using formaldahyde and do ChIP([Shang, Hu et al. 2000](#_ENREF_7); [Kwon, Garcia-Bassets et al. 2007](#_ENREF_4)) esp shang! (control with IGg I think ChIP, and no ab chip, see what lit has to say. qPCR [read about]
    - Those loci from above (use statistical power test to determine how many (read ch 7 or something of stats book)
    - Find how different mods are effected by rest knock down. (why not do a knock out?)
    - Choose top lets say 3 mods that can trace a lot of their effect to rest mediated stuff, chip them.
    - Use some loci from groups 2 and 3 as controls (stat power testing to det how many)
    - ?
    - **find activating isoform 4 of REST in T cells (not done said Zheng 2009)?**
    - **Double ChIP 🡪 define how REST and histone modification correlated at molecular level**
    - **Calculate fpkm for miRNAs that are highly expressed and bound in t cells after KD**
    - overexpression of REST as well? (prolly not)
  + trip ups?
    - There have been problems in the past transfecting? (right word) cd 4+ t cells with sirnas. 🡪 found some ideas online, nanotubes?
    - Chip can be a finicky assay 🡪 chip seq forums

“**REST maintains self-renewal and pluripotency in mouse ES cells through suppression of the microRNA miR-21”(**[**Singh, Kagalwala et al. 2008**](#_ENREF_8)**)**

expression needs to be compared to something in aim 3 perhaps a housekeeping gene like cyclophilin (used in Greenway 2007) or 16S rRNA

How would this experiment be different if instead of using a KO or KD, used a dominant negative form of REST, as seen in Greenway 2007?

Use drugs to inhibit function of found epigenetic modifiers (such as trichostatin A (TSA) (Greenway 2007))

Corcoran, K. E., K. A. Trzaska, et al. (2008). "Mesenchymal stem cells in early entry of breast cancer into bone marrow." Plos One **3**(6): e2563.

Coulson, J. M., C. E. Fiskerstrand, et al. (1999). "Arginine Vasopressin Promoter Regulation Is Mediated by a Neuron-restrictive Silencer Element in Small Cell Lung Cancer." Cancer Research **59**(20): 5123-5127.

Greco, S. J., S. V. Smirnov, et al. (2007). "Synergy between the RE-1 Silencer of Transcription and NFŒ∫B in the Repression of the Neurotransmitter Gene TAC1 in Human Mesenchymal Stem Cells." Journal of Biological Chemistry **282**(41): 30039-30050.

Kwon, Y.-S., I. Garcia-Bassets, et al. (2007). "Sensitive ChIP-DSL technology reveals an extensive estrogen receptor Œ±-binding program on human gene promoters." Proceedings of the National Academy of Sciences **104**(12): 4852-4857.

Majumder, S. (2006). "REST in good times and bad: roles in tumor suppressor and oncogenic activities." Cell Cycle **5**(17): 1929-1935.

Reddy, B. Y., S. J. Greco, et al. (2009). "RE-1‚Äìsilencing transcription factor shows tumor-suppressor functions and negatively regulates the oncogenic TAC1 in breast cancer cells." Proceedings of the National Academy of Sciences **106**(11): 4408-4413.

Shang, Y., X. Hu, et al. (2000). "Cofactor dynamics and sufficiency in estrogen receptor-regulated transcription." Cell **103**(6): 843-852.

Singh, S. K., M. N. Kagalwala, et al. (2008). "REST maintains self-renewal and pluripotency of embryonic stem cells." Nature **453**(7192): 223-227.

Su, X., V. Gopalakrishnan, et al. (2006). "Abnormal expression of REST/NRSF and Myc in neural stem/progenitor cells causes cerebellar tumors by blocking neuronal differentiation." Molecular and cellular biology **26**(5): 1666.

Weissman, A. M. (2008). "How much REST is enough?" Cancer Cell **13**(5): 381-383.

Westbrook, T. F., E. S. Martin, et al. (2005). "A genetic screen for candidate tumor suppressors identifies REST." Cell **121**(6): 837-848.

Zhao, J. J., O. V. Gjoerup, et al. (2003). "Human mammary epithelial cell transformation through the activation of phosphatidylinositol 3-kinase." Cancer Cell **3**(5): 483-495.