Research proposal for Dr Mario Suvà MEDIC Award 2010: completion of the following research project in Boston, at MGH, in the Laboratory of Dr Bernstein and Dr Louis.

The laboratories and a grant from Harvard Stem Cell Institute will provide full support for the experimental cost. The request is that CHF 100,000 from the MEDIC award would be use to cover my salary for one year.

Research Abstract:

Glioblastoma (GBM) is a high-grade infiltrating brain tumor that remains incurable despite aggressive multimodal therapy. GBM has recently been shown to contain a primitive population of cancer stem cells (CSC) with exclusive tumor-initiating potential. These cells show high chemotherapy and radiation resistance, highlighting the urgent need for alternative therapeutic strategies. Importantly, CSC share the same genetic alterations with the more differentiated cancer cells to which they give rise and it is believed that their defining properties are rooted in their epigenetic state. To identify a semi-comprehensive set of genes that drive brain tumor stem cells, we have applied chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) to deeply profile a panel of histone modifications in GBM CSC and matched non-stem GBM cells from different individuals, as well as in human neural stem cells and primary human astrocytes. Our preliminary data reveals aberrant activation of a set of transcription factors (TFs) in GBM CSC and our initial functional experiments reveal a minimal core of TFs that is sufficient to confer CSC properties to GBM cells *in vitro*. We propose to refine a network model of the interactions between these factors, to identify their target gene repertoire and to perform *in vivo* experiments to demonstrate their acquired CSC features. Over the long-term, we predict that these innovative and systematic approaches have high potential to significantly impact our understanding and treatment of malignant brain tumors.

Research Proposal: Background and rationale

Glioblastoma (GBM) is a high-grade brain tumor that is among the most lethal of all cancers. GBMs contain a primitive subpopulation of cancer stem cells (CSC) that are believed to be the tumor-driving force in this disease and to play a major role in tumor recurrence after therapeutic intervention[1]. Different cell surface markers, including CD133[2], SSEA1[3], CD44[4] and integrin α6[5] have been suggested to enrich for CSC in GBM and a few signalling pathways, including Sonic-Hedgehog[6], Bone Morphogenic Proteins[7], HIF-1α[8] and TGF-β[4], have been implicated in GBM CSC maintenance. However, a comprehensive epigenomic characterization of GBM CSC is needed to identify the core network and key features defining CSC in GBM and to distinguish them from genetically identical but more differentiated cancer cells, as well as from non-malignant counterparts.

The chromatin landscape ('epigenome') of a given cell type can be mapped across the genome comprehensively and at high-resolution through the use of next-generation sequencing ChIP-seq[9,10]. Signature chromatin structures can be used to identify promoters, transcripts, enhancers, silencers and repressive chromatin domains[11-13]. They also inform on the regulatory state assumed by a given DNA element in a given cell type; for example, promoters that are active, repressed or 'poised' can be distinguished by the combined patterns of H3K4 and H3K27 tri-methylation[14].

We acquired high resolution ChIP-seq profiles for H3K4 and H3K27 tri-methylation for GBM CSC, and compared them against matched non-stem GBM cancer cells as well as non-malignant neural stem cells and primary human astrocytes (two candidate cells of origin for GBM). Importantly, the relevance of our culture model for expansion of CSC and non-stem populations of GBM has been extensively studied and their respective *in vivo* tumor-initiating potential precisely characterized[15]. Our preliminary results suggest that GBM CSC have a transcriptional regulatory network that is distinct from their genetically identical non-stem counterparts as well as from non-malignant CNS cells. In particular, we have found a set of developmental transcription factor (TF) genes that switch from a bivalent (H3K4me3/H3K27me3) to an H3K4me3-specific state in GBM CSC, suggesting that they are specifically activated in the tumor stem cell subpopulation. Moreover, an analysis of H3K4me1 patterns indicates that the corresponding TF loci are enriched for CSC-specific enhancers, which may collectively activate the TF module. These novel and compelling preliminary data form the basis of our proposal to understand the role of this transcriptional network in GBM tumorigenesis (**Figure**).

Specific aims: aim 1. Refine a network model of the transcriptional regulators and functional genomic elements that govern gene expression in GBM CSCs.

Our preliminary data support the feasibility of GBM CSC chromatin analysis, and suggest that aberrant TF expression is a prominent feature distinguishing these cells from other cancer cells as well as from their non-tumor counterparts. The initial TF set included 21 genes, encompassing several TFs with established roles in gliomagenesis as well as many novel candidates. Because TFs are highly likely to control a set of downstream targets, we will investigate regulatory relationships encompassing our glioma TFs by searching for TF consensus motifs in promoters and distal regulatory elements activated in CSC. This screen will be performed using standard computational methods and motif databases[16]. Binding motifs matching our GBM TFs will help us understand the downstream network of these regulators. We will use chromatin profiles generated for epitopes associated with enhancers and other distal elements (H3K4me1, H3K27Ac) to annotate such elements and compare their activities between our different cell types[17]. Based on our preliminary studies, we anticipate that this combination of promoter, gene and distal elements annotations will provide a powerful means for deriving the cis-connectivity of the GBM regulatory network. In addition, the chromatin state analyses will be complemented by mRNA expression profiling, using whole-transcriptome sequencing (RNA-seq). We expect to detect additional genes highly correlated with TFs that represent putative downstream targets. We plan to integrate gene expression profiles into our chromatin-focused study with the goal of collating a high confidence set of GBM CSC regulators for network derivation.

Our experimental approach has thus far provided an unprecedented richness of annotations that reveal the locations and activities of promoters, transcripts and regulatory elements, and have highlighted a small group of

novel TFs as exclusively activated in GBM CSC. We believe that refining this model and deriving its connectivity network will allow us to identify the core network of regulators that are key for GBM CSC maintenance.

Specific aims: aim 2. Validate the transcription factor network by *in vitro* and *in vivo* functional validation, and identify its target genes repertoire

Based on our preliminary data, a set of experiments was conducted to identify TFs among our initial set that would be critical for CSC. The experimental approach (ongoing) consisted of systematic re-introduction of all candidate TFs into non-stem glioma cells and assessing for appearance of CSC features within these cells. This approach led thus far to the identification of a set of 2 TFs (SOX2 and a POU domain TF novel to gliomas) whose co-expression by lentiviral system is sufficient to endow differentiated glioma cells with *in vitro* characteristics of CSC, including the capacity to grow in serum-free conditions, form gliomaspheres and express CD133. We have not yet completed the experimental approach with the whole set of TFs and we anticipate the possibility of complementing our initial system in different ways including addition of a third or possibly fourth TF, use of dose-responsive inducible vector systems and addition of small molecules - parameters that have proven key for improved reprogramming of somatic cells[18]. Our working hypothesis is that a minimal core of TFs, containing at least these 2 TFs is critical for CSC maintenance in GBM. The inclusion of both a well-known (SOX2) and a yet un-described glioma factor supports the functional relevance and novelty of our candidate regulators.

To further demonstrate acquisition by GBM cells of CSC features, we propose to perform an additional set of functional experiments including: a) *in vitro* differentiation assays into neuronal and astrocytic lineage; b) *in vivo* orthotopic xenotransplantation assays in immunocompromised mice, in limiting dilutions; c) genome-wide chromatin marks profiling and mRNA expression profiling (see figure). We will also perform experimental perturbation of the network through lentiviral-mediated RNA interference in GBM CSC. Phenotypic effects of specific TF silencing on CSC will be tested through evaluation of viability, proliferation, clonogenicity, cell surface markers, and tumor-initiating potential. Finally, we will supplement these analyses by generating ChIP-Seq maps for our minimal TF core in GBM CSC to identify their target genes repertoire.

While the identification as well as the functional and phenotypic characterization of CSC in GBM have been extensively studied and reported[2-7], no work has yet identified their core defining regulator network in a genome-wide unbiased approach. We believe that our chromatin state analysis complemented by the proposed functional experiments will allow us to make significant progress in this direction and to thereby achieve a better biological understanding of this devastating malignancy.

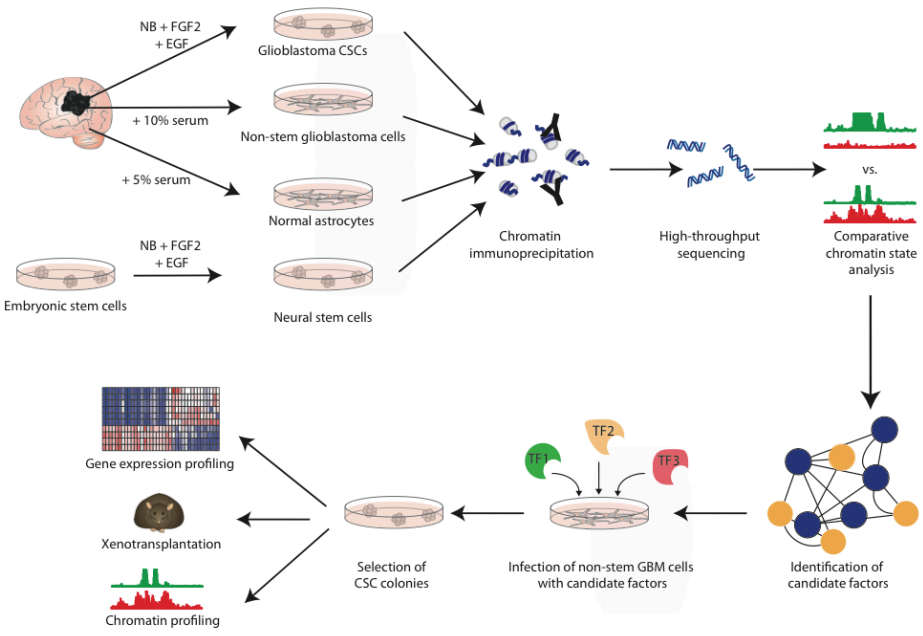

Figure legend: **Top panel:** surgically removed GBM samples were cultivated as gliomaspheres in stem cell conditions and as adherent non-stem cultures in medium supplemented with FCS. Primary human astrocytes as well as ES-derived neural stem cells were expanded in respective media. Chromatin immunoprecipitation for H3K4me3 and H3K27me3 was performed for each cell type followed by high-throughput sequencing. Comparative genome-wide analysis of local enrichment was performed. **Bottom panel:** we propose to derive a refined network of transcription factors regulating GBM CSC. We will validate the model by adding selected transcription factors (TF1-3) to adherent non-stem GBM cultures and wait for gliomaspheres to grow out. These colonies will then be submitted to gene expression profiling, orthotopic xenotransplantation and chromatin profiling to test their similarity to primary CSC cultures. **Abbreviations:** NB: neurobasal medium; FGF2: basic fibroblast growth factor; EGF: epidermal growth factor; TF: transcription factor; FCS: fetal calf serum; CSC: cancer stem cells.

Milestones : *Months 1-6:* refine TF lists and complete their cloning into lentiviral systems; refine minimal TF combinations sufficient to confer *in vitro* CSC features, perform *in vivo* serial xenotranplantation assays. Systematic and stable knock-down of TFs in GBM CSC by lentiviral shRNA. *Months 7-12:* acquire ChIP-seq and RNA-seq data to identify the target genes repertoire of selected TFs, data analysis and integration into a final model, which will be evaluated for potential to suggest diagnostic or therapeutic strategies.

Key risks and plans to mitigate:

1. Identifying the minimal core of TFs providing CSC features to cancer cells may be complicated by parameters such as TF numbers and combinations, relative expression levels and culture conditions. However, our initial data provide us with a limited number of candidates and suggest that our settings will enable accurate selection of TF combinations; if necessary, expression levels will be matched to CSC populations using inducible, dose-responsive lentiviral systems, and alternate combinations will be attempted; different protocols for cell passaging, as well as addition of small molecules blocking pathways preventing reprogramming[18] will be attempted; readout will be based on acquired spherogenic potential, induction of CSC markers including CD133 and SSEA-1, as well as the capacity of a given TF combination to induce the TF network; to help identify the minimal core, additional TFs will be chosen among those non-induced, until the whole network and/or the phenocopy is completed.

2. Robustness and flexibility of the network: GBM is a highly refractory tumor, and it is possible that the regulatory network of GBM CSCs could be re-organized in response to perturbation of key components by RNAi. We may find resistant clones emerge within GBM CSC cultures upon intervention or in the *in vivo* xenotransplantation assays. We will explore simultaneous targeting of multiple nodes or pathways by combinations of shRNA vectors.

References

1. Bao, S.*, et al.* Glioma stem cells promote radioresistance by preferential activation of the DNA damage response. *Nature* **444**, 756-760 (2006).
2. Singh, S.K.*, et al.* Identification of human brain tumour initiating cells. *Nature* **432**, 396-401 (2004).
3. Son, M.J., Woolard, K., Nam, D.H., Lee, J. & Fine, H.A. SSEA-1 is an enrichment marker for tumor-initiating cells in human glioblastoma. *Cell Stem Cell* **4**, 440-452 (2009).
4. Anido, J.*, et al.* TGF-beta Receptor Inhibitors Target the CD44(high)/Id1(high) Glioma-Initiating Cell Population in Human Glioblastoma. *Cancer Cell* **18**, 655-668 (2010).
5. Lathia, J.D.*, et al.* Integrin alpha 6 regulates glioblastoma stem cells. *Cell Stem Cell* **6**, 421-432 (2010).
6. Clement, V., Sanchez, P., de Tribolet, N., Radovanovic, I. & Ruiz i Altaba, A. HEDGEHOG-GLI1 signaling regulates human glioma growth, cancer stem cell self-renewal, and tumorigenicity. *Curr Biol* **17**, 165-172 (2007).
7. Piccirillo, S.G.*, et al.* Bone morphogenetic proteins inhibit the tumorigenic potential of human brain tumour-initiating cells. *Nature* **444**, 761-765 (2006).
8. Li, Z.*, et al.* Hypoxia-inducible factors regulate tumorigenic capacity of glioma stem cells. *Cancer Cell* **15**, 501-513 (2009).
9. Barski, A.*, et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823-837 (2007).
10. Mikkelsen, T.S.*, et al.* Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448**, 553-560 (2007).
11. Bernstein, B.E.*, et al.* Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**, 169-181 (2005).
12. Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R. & Young, R.A. A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130**, 77-88 (2007).
13. Heintzman, N.D.*, et al.* Histone modifications at human enhancers reflect global cell-type-specific gene expression. *Nature* (2009).
14. Bernstein, B.E.*, et al.* A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315-326 (2006).
15. Wakimoto, H.*, et al.* Human glioblastoma-derived cancer stem cells: establishment of invasive glioma models and treatment with oncolytic herpes simplex virus vectors. *Cancer research* **69**, 3472-3481 (2009).
16. Matys, V.*, et al.* TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res* **31**, 374-378 (2003).
17. Ernst, J.*, et al.* Mapping and analysis of chromatin state dynamics in nine human cell types. *Nature* **473**, 43-49 (2011).

18. Maherali, N. & Hochedlinger, K. Guidelines and techniques for the generation of induced pluripotent stem cells. *Cell stem cell* **3**, 595-605 (2008).