A Comprehensive Bioinformatics Analysis of HDAC-interacting Proteins

Skyler Seto, Xiaotao Qu

H. Lee Moffitt Cancer Center & Research Institute, Tampa, Florida

H. LEE MOFFITT Cancer Center & Research Institute
A National Cancer Institute Comprehensive Cancer Center
At the University of South Florida

Background

A key event in the regulation of eukaryotic gene expression is the post-translational modification of histones, which converts regions of chromosomes into transcriptionally active or inactive chromatin. The most common post-translational modification of histones is the acetylation of ε-amino groups on conserved lysine residues in the histones' amino-terminal tail domains. Hyperacetylation of histones generally correlates with transcriptionally active chromatin, perhaps by increasing the accessibility of transcription factors to nucleosomal DNA, while hypoacetylation of histones correlates with transcriptional silencing. The long-term goal of our research program is to understand the underlying mechanisms of histone deacetylation, with particular emphasis on a detailed dissection of the mammalian histone deacetylase (HDAC) enzymes. These studies will ultimately provide important insights critical for a thorough understanding of the intricate mechanisms operating to orchestrate gene expression in mammalian cells.

In humans, HDACs are divided into three categories: the Class I RPD3-like proteins (HDAC1, HDAC2, HDAC3, and HDAC8); the Class II HDA1-like proteins (HDAC4, HDAC5, HDAC6, HDAC7, HDAC9, and HDAC10); and the Class III SIR2-like proteins (SIRT1-7). The Class II enzymes can be further divided into two subclasses: IIa (HDAC4, HDAC5, HDAC7, and HDAC9) and IIb (HDAC6 and HDAC10). HDAC11, the most recently discovered human HDAC, uniquely shares sequence homology to the catalytic domains of both Class I and II HDAC enzymes. The Class III proteins do not exhibit any sequence similarity to HDAC family members from Class I or II and differ from other HDACs in that they require the cofactor NAD for activity.

Studies suggest that many proteins that have previously been described as protein modification enzymes may represent only single members of protein complexes and function properly only together with other factors. Numerous data suggest that HDACs may also require additional factors for their function. First, while immunopurified HDAC complexes deacetylate histones efficiently, bacterially- or baculovirally- expressed and highly purified recombinant HDACs do not deacetylate histones at all. Although this could be interpreted as due to technical difficulties, an alternative and more likely explanation is that each HDAC serves only as one component of the histone deacetylase system. Perhaps the best indication that each HDAC is involved in protein-protein interactions with other cellular factors is the fact that we and others have detected, through immunoprecipitation of whole cell extracts, a number of proteins that bind specifically to each HDAC. This project outlines bioinformatics methods that we have undertaken to identify possible cellular proteins that interact with HDACs. The data obtained in this study allows us now to complement the findings using a variety of laboratory-based techniques to place the interacting proteins into their cellular context with classical biochemical methods. In the future, we will verify these protein-protein interactions, and the results from these studies will lay the groundwork for generating new testable hypotheses that can be addressed by more specific experiments. Knowledge of these interactions will provide a basic understanding of the mechanism of each HDAC and add to our understanding of gene regulation by histone deacetylation.

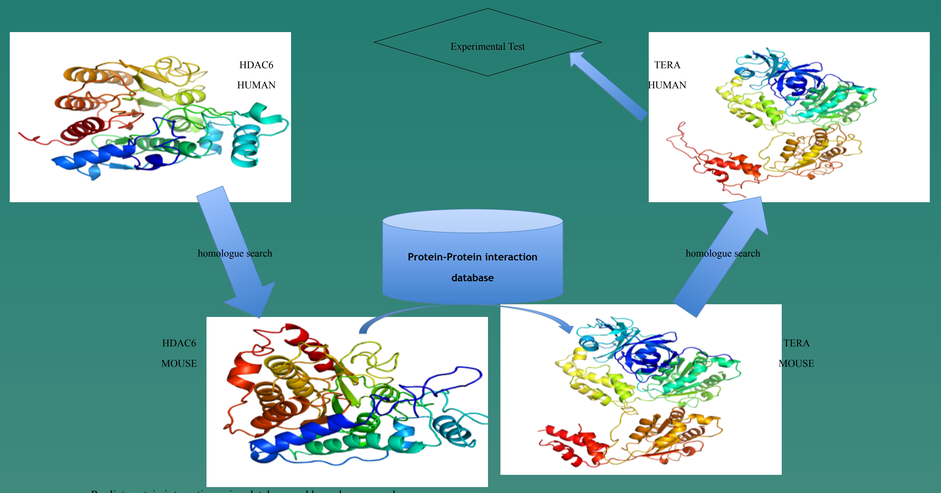
Abstract

Histone deacetylases (HDACs) are enzymes that regulate the functions of histone and nonhistone proteins by catalyzing the removal of acetyl groups from lysine residues. They control gene expression by altering chromatin organization and dynamics. In addition, HDACs function as key regulators of many cell processes including intracellular trafficking, autophagy and organelle turnover, ciliary disassembly, cell matrix interactions and signaling , cell migration, cell death , and cell-cell Interactions.

Many proteins that have previously been described as protein modification enzymes represent only single members of protein complexes and only function properly together with other factors. Increasing evidence suggests that HDACs may also require additional factors for their various functions. Using publicly available databases and bioinformatics tools such as BLAST, we were able to identify putative interacting proteins of eighteen human HDACs by constructing a "round trip" between human HDACs and their interacting proteins through homologues in other species that are experimentally verified to interact. A database was developed by combining many data sources containing protein-protein interaction information, and a fully automatic method was developed to query, annotate and identify potential interacting proteins, which can be applied to any protein of interest.

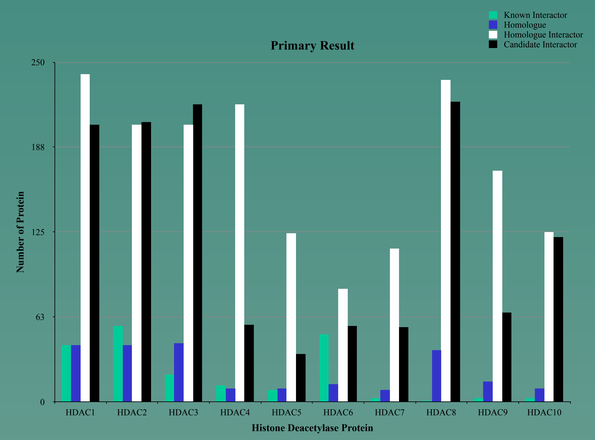
Our findings that HDACs associate with a diverse number of different proteins lead to many questions. For example, do proteins that interact with HDACs stimulate or antagonize HDAC activities? Do HDACs reciprocally alter the functions of these HDAC-binding proteins? Work is now underway to further analyze the proteins that associate with HDACs and to determine how HDACs modulate the functions of these proteins. Knowledge of these interactions will provide a basic understanding of the mechanism of each HDAC and add to our understanding of how HDACs regulate many important biological processes.

Method

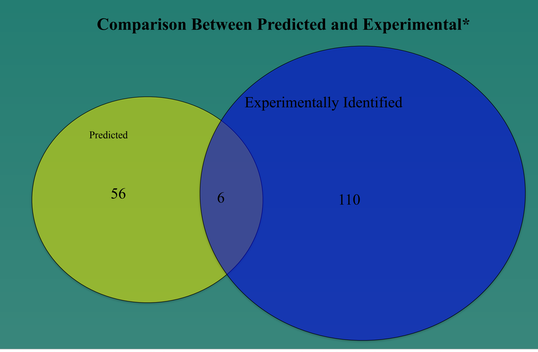


Predict protein interaction using database and homologue search

Result 1



Result 2



Experimental results was from pull-down experiments using cytoplasmic fraction from HeLa cells
Note: the experimental results contain proteins that do not directly interact with HDAC10 while our prediction method only predicts those proteins that can form direct contact with HDAC protein.

Sample Results of Potential Candidates

	First Round Sequence Alignment					Second Round Sequence Alignment					
HDAC	Sequence Identity (%)	Alignment Coverage (%)		Homolog	Interactor	Sequence Identity (%)	Alignment Coverage (%)		Candidate	Score
		HDAC10	Homolog				Interactor	Candidate		
Q13547	58.5	99	99	Q569C4	P63102	69.8	95	92	P62258	11.54
	51.2	54	59	Q9UBN7	P68366	93.5	100	100	Q9BQE3	11.18
	51.2	54	59	Q9UBN7	Q71U36	94.4	96	96	P68366	11.15
	51.2	54	59	Q9UBN7	P68366	39.4	97	95	Q9BUF5	10.77
	51.2	54	59	Q9UBN7	P68366	38.7	97	95	P68371	10.76
	51.2	54	59	Q9UBN7	P68366	38.5	97	95	Q9BVA1	10.76

Examples of our results correspond to experimental results ordered by our draft score. The first round sequence alignment stands for the alignment between HDAC10 (Q13547) and its homologues in non-human species. The second round sequence alignment stands for the alignment between non-human integrators and their human homologues. The alignment converge indicates how much of the two sequence can be aligned together, a 50% cut-off was used to ensure selectivity.
The score was calculated based on the sequence identity and alignment coverage for both first round and second round alignment.

Conclusion

The identification and characterization of proteins in a cell with which a given protein interacts is often helpful for understanding the function and mechanisms of action of that protein. This was certainly the case for HDACs. For this reason, we initiated a data mining method (knowledge discovery from databases) to identify protein-protein interactions crucial in cancer. The project is designed to predict proteins that could interact with human histone deacetylase (HDAC) enzymes by searching the public databases for crosss-species homologs of HDACs and proteins known to interact with these homologs. Our results not only confirmed many HDAC-interacting proteins previously identified by biochemical and cell biology methods, but also suggested hundreds of potential HDAC-interacting partners that escaped conventional detections. We are currently working on completing a fully automatic method to