Ledidi: Designing genomic edits that induce functional activity

Jacob Schreiber*, Yang Young Lu*, William Stafford Noble
Department of Genome Science
University of Washington
*contributed equally to this work

jmschreiber91

@jmschrei

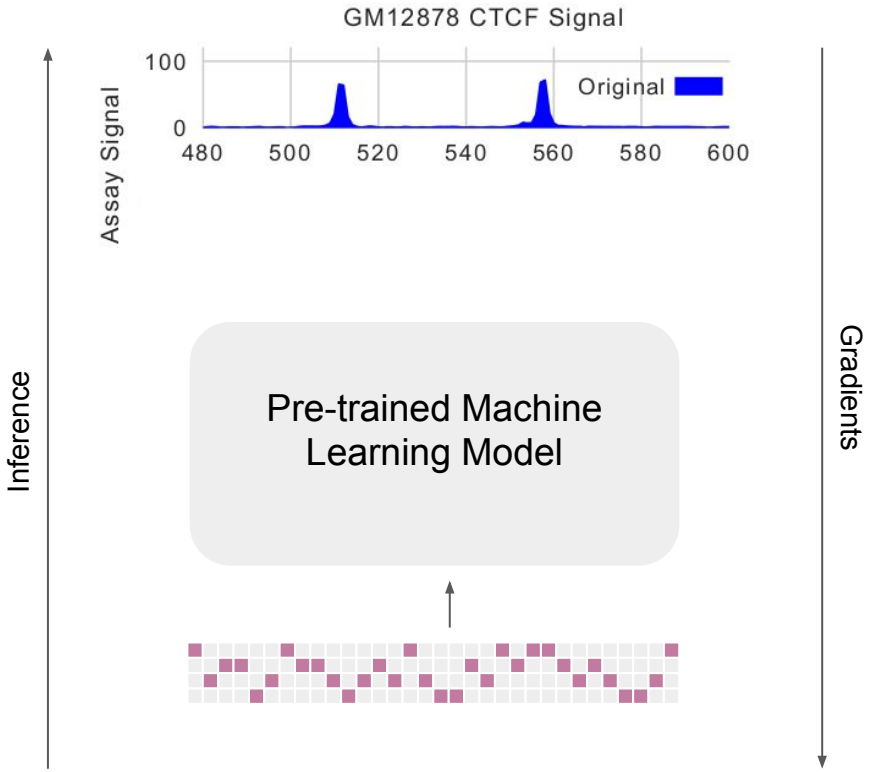
@jmschreiber91

Genome editing technologies show great promise but face challenges

1. How do you choose a set of edits that achieves a desired effect?

2. How do you ensure that these edits do not have unintended consequences?

Ledidi turns any machine learning model into a sequence editor



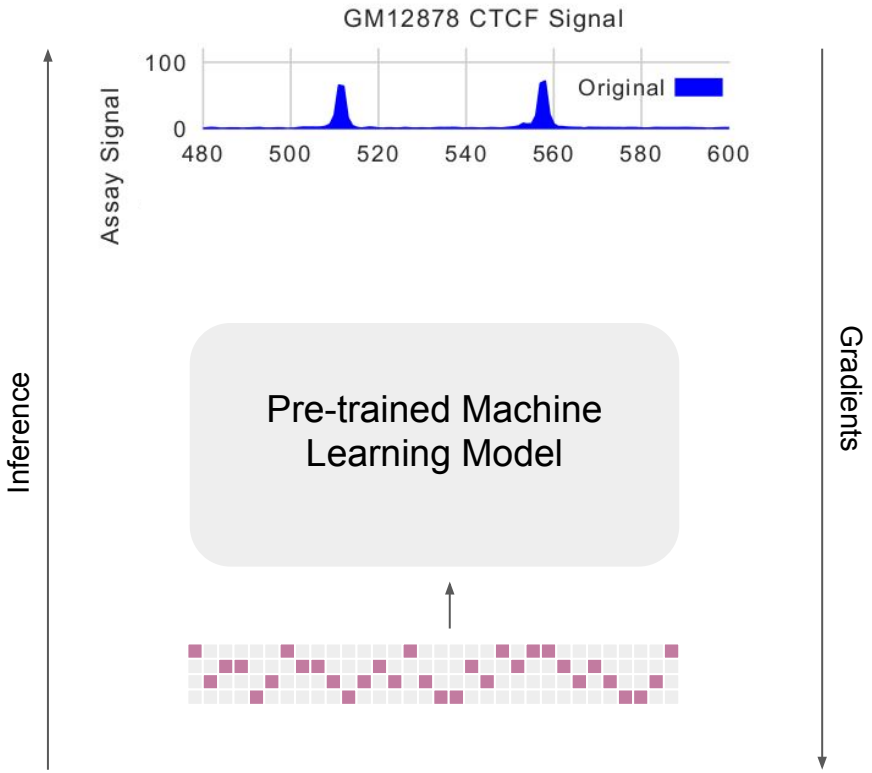
Ledidi optimizes the following mixture objective:

$$\min_X ||X - X_0||_1 + \lambda ||f(X) - \hat{y}||_2^2$$

the edited sequence
the original sequence

the predicted output
the desired output

Ledidi turns any machine learning model into a sequence editor



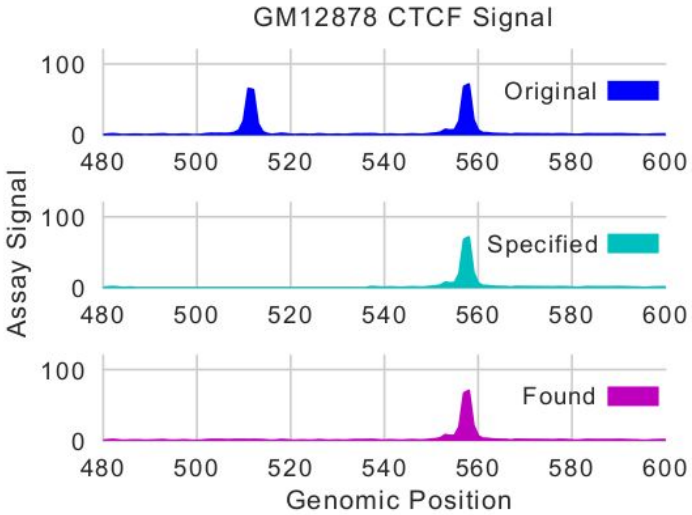
Ledidi optimizes the following mixture objective:

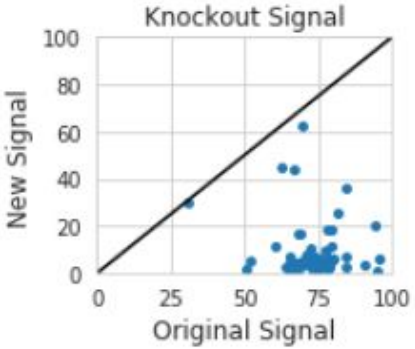
$$\min_X ||X - X_0||_1 + \lambda||f(X) - \hat{y}||_2^2$$

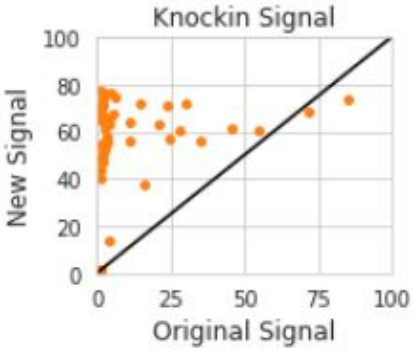
Directly optimizing discrete sequence is difficult so Ledidi uses the **Gumbel-softmax** reparameterization trick to replace X with a continuous representation, W.

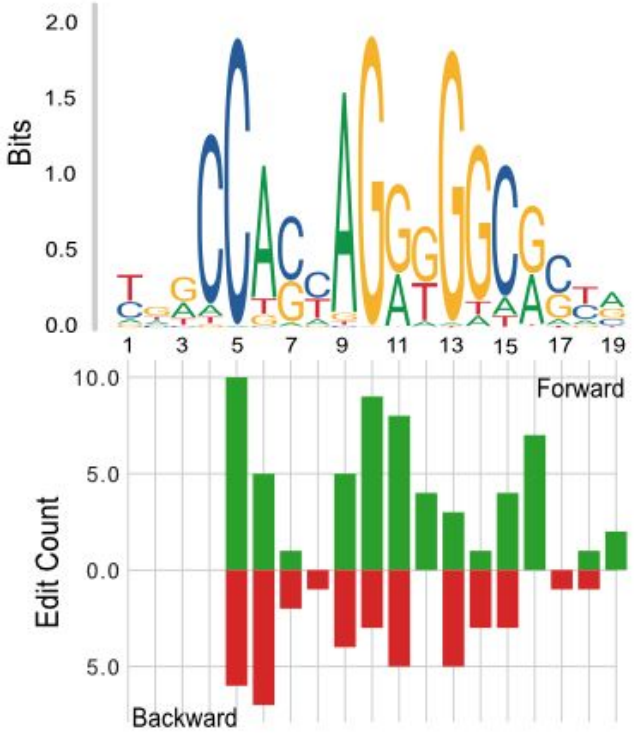
$$g(W_{ij}) = \frac{(\exp(W_{ij} + z)/\tau)}{\sum_{j'=1}^{d} \exp((W_{ij'} + z)/\tau)}$$

We demonstrate that Ledidi can design functional edits using Basenji



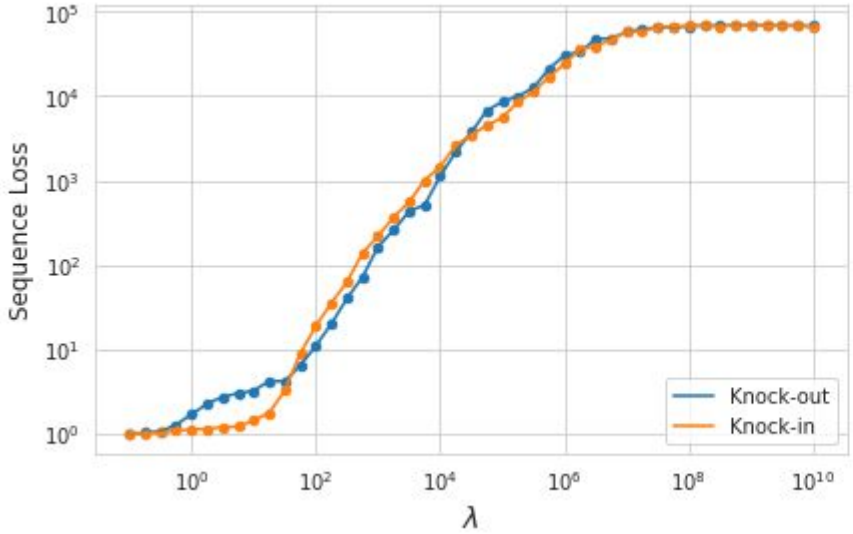


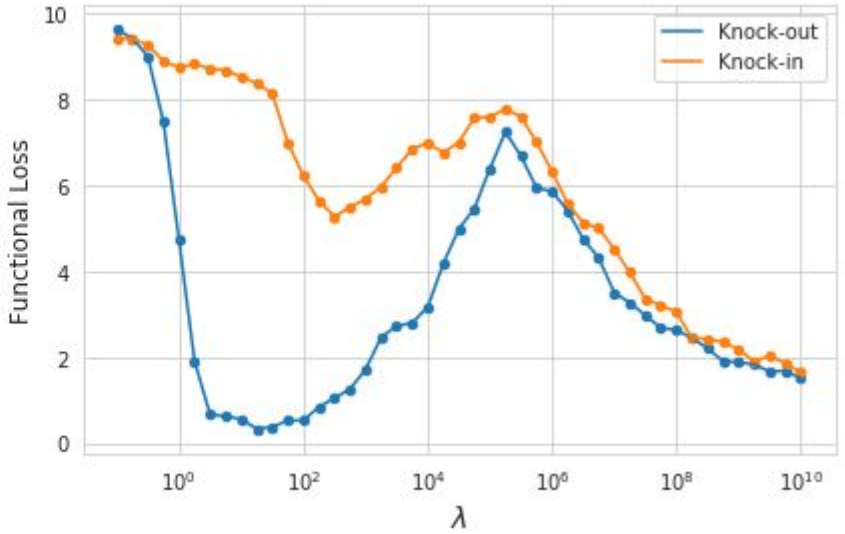




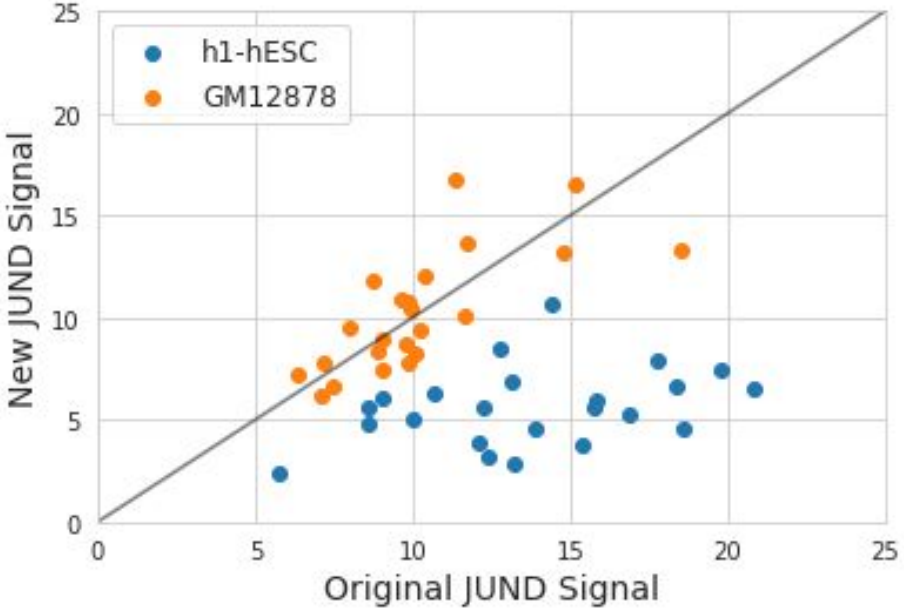
Changing the regularization strength affects the found edits

$$\min_X ||X - X_0||_1 + \lambda ||f(X) - \hat{y}||_2^2$$





Ledidi can design edits that induce cell-type-specific JUND binding



Acknowledgements

Yang Lu

William Noble

eScience Institute
ADVANCING DATA-INTENSIVE DISCOVERY IN ALL FIELDS

NSF National Science Foundation
WHERE DISCOVERIES BEGIN