

## Guide to Identifying Motifs from Your ChIP-chip Data

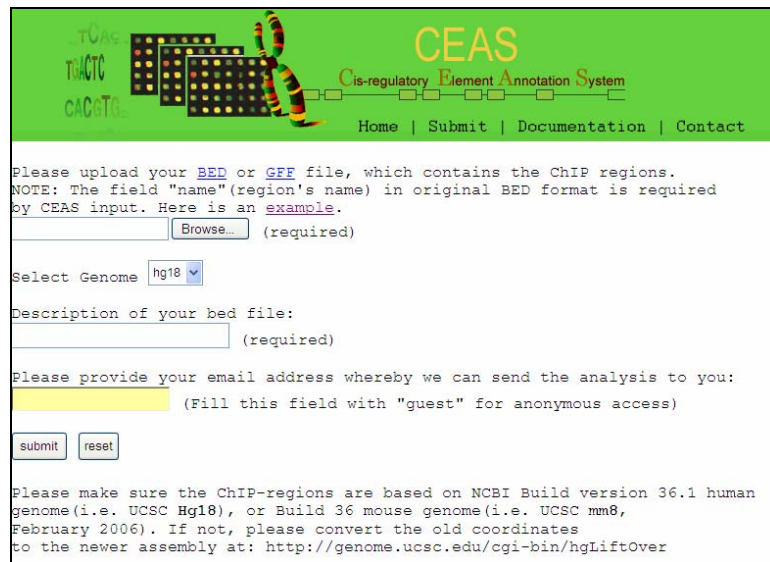
### Outline

This document describes how to use the Cis-regulatory Element Annotation System (CEAS) to identify transcription factor motifs, map nearby genes, evaluate GC content and evolutionary conservation, and mine sequences for qPCR validation in NimbleGen ChIP-chip data. Be advised that CEAS works only with current human (hg18) and mouse (mm8) genome builds.

---

### Step 1. Submit Your ChIP-chip Data for Analysis

1. Access the CEAS web site: <http://ceas.cbi.pku.edu.cn/>
2. At the top of the window, click the **Submit** link.



The screenshot shows the CEAS web interface. At the top, there is a green header with the CEAS logo and navigation links: Home | Submit | Documentation | Contact. Below the header, there is a form for submitting ChIP-chip data. The form includes a text input field for the BED or GFF file, a 'Browse...' button, a 'Select Genome' dropdown menu set to 'hg18', a text input field for the description of the bed file, and a text input field for the email address. There are 'submit' and 'reset' buttons at the bottom of the form. A note at the bottom of the form states: 'Please make sure the ChIP-regions are based on NCBI Build version 36.1 human genome (i.e. UCSC Hg18), or Build 36 mouse genome (i.e. UCSC mm8, February 2006). If not, please convert the old coordinates to the newer assembly at: <http://genome.ucsc.edu/cgi-bin/hgLiftOver>'

3. Click **Browse** and upload the following file from your data DVD:  
SignalMap\_GFF\_files\Peaks\arrayID\_ratio\_peaks.gff
4. Select your genome of interest (hg18 or mm8) in the *Select Genome* field.
5. Enter the name of the *arrayID\_ratio\_peaks.gff* file in the *Description of your bed file* field.
6. Enter the email address to which to send the results in the *Please provide your email address whereby we can send the analysis to you* field.  
**Note:** The email will contain a URL link to the results.
7. Click the **submit** button.

---

## Step 2. Learn More about CEAS Analysis Results

A CEAS analysis provides four types of results:

- Transcription Factor Motif Identification
- Nearby Gene Mapping
- Conservation Plot
- Region Sequence

These types of results are summarized below. For a more detailed description, click the **Documentation** link at the top of the CEAS window.

### Transcription Factor Motif Identification

CEAS analysis results identify enriched sequence motifs within ChIP-chip regions bound by a transcription factor. From approximately 800 eukaryotic motifs, CEAS counts the number of hits for each motif within the ChIP-chip region and throughout the entire genome. Significantly enriched motifs ( $> 2$ -fold,  $p$ -value  $< 1E-10$ ), sequence, and sequence logo corresponding to the ChIP-chip region are reported.

### Nearby Gene Mapping

Promoter reports provided by NimbleGen with your ChIP-chip data include initial analyses that localize the binding of your transcription factor by mapping peaks to transcription start site. CEAS also reports ChIP-chip peaks relative to the nearest RefSeq genes in both upstream and downstream directions on both strands within 300kb. Peaks within a gene are reported as residing in the 5' or 3' UTR, coding exon, or intron. CEAS also provides statistics for gene mapping of all your ChIP-chip regions.

### Conservation Plot

To interrogate the level of conservation across species, CEAS utilizes phastCons information from the UCSC GoldenPath genome resource. This resource assigns a conservation score to every nucleotide in the human genome based on a phylogenetic algorithm. CEAS provides a downloadable PDF file (.pdf) for each ChIP-chip region, which allows for quick identification of many regions enriched by a transcription factor.

### Region Sequence

CEAS provides sequence in FASTA format to download for use in qPCR primer design for validation of ChIP-chip data. For repeat-masked sequence, the character "N" is used.

---

## References

Ji X, Li W, Song J, Wei L, and Liu XS. 2006. CEAS: cis-regulatory element annotation system. *Nucleic Acids Res.* 34:W551-554.

---

## Contact Information

Technical Support	(877) NimbleGen / (877) 646-2534 ext. 4 techsupport@nimblegen.com
Customer Services	(877) NimbleGen / (877) 646-2534 ext. 3 customer-services@nimblegen.com
Sales Support (for billing information)	(877) NimbleGen / (877) 646-2534 ext. 1 billing@nimblegen.com