

This document has been moved to <https://github.com/stjude/proteinpaint/wiki/Tracks#track-js-on-bed-track-format> And is no longer maintained!

ProteinPaint JSON-BED track format

Example:

https://proteinpaint.stjude.org/?block=on&genome=hg19&bedjfile=JSON-BED_proteinpaint_demo/hg19/bedj/gencode.gz

1. File format

2. Prepare and use a track

3. JSON object for a BED item

["name": STR](#)
["isoform": STR](#)
["strand": STR](#)
["exon": \[\]](#)
["intron": \[\]](#)
["utr5": \[\]](#)
["utr3": \[\]](#)
["rnaLen": INT](#)
["cdslen": INT](#)
["codingstart": INT](#)
["codingstop": INT](#)
["startCodonFrame": 1/2](#)
["coding": \[\[start, stop\], ... \]](#)
["description": STR](#)
["color": STR](#)
["exon2color": \[{ start, stop, color }, ... \]](#)
["category": STR](#)
["isoformonly": STR](#)

4. Declaring a track as a JSON object

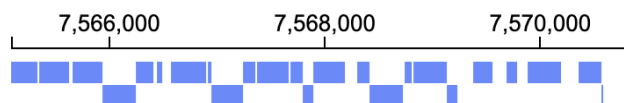
1. File format

A track file requires at least 3 columns, separated by **tab**:

1. Chromosome
2. Start position (0-based)
3. Stop position (not including the ending base)
4. Stringified JSON object, ***this is optional***.

Each line is a genomic feature. Using a file with only the first 3 columns will produce a “barebone” rendering of the data.

https://proteinpaint.stjude.org/?genome=hg19&block=1&bedurl=test,https://pecan.stjude.cloud/static/prot einpaint_demo/hg19/misc/rmsk.bed3.gz



Using JSON objects at the 4th column allows to richly describe the genomic features, which offers better flexibility compared to fixed columns of BED format and its variants. See section 3 about JSON object specification.

2. Prepare and use a track

```
$ sort -k1,1 -k2,2n [file] > [file.sorted]
$ mv [file.sorted] [file]
$ bgzip [file]
$ tabix -p bed [file].gz
```

To host the track on a web server: put both .gz and .gz.tbi files at the same directory on the web server. Obtain the URL to the .gz file and submit it to the browser.

To host the track on ProteinPaint server: put both .gz and .gz.tbi files at the same directory under the <TP> directory. Obtain the relative path to the .gz file and submit it to the browser.

Alternatively, a [bigbed file](#) could be used as source file and will be parsed into JSON-BED format:

To host the track on a web server: put the bigbed file at the same directory on the web server. Obtain the URL to the bigbed file and submit it to the browser.

To host the track on ProteinPaint server: put bigbed file at the same directory under the <TP> directory. Obtain the relative path to the bigbed file and submit it to the browser.

3. JSON object for a BED item

The content is a string representation of an object (key-value pairs). Examples:

```
{"name":"CTCF1","strand":"+"}  
  
{"name":"MIR6859-1","isoform":"NR_106918","strand":"-","exon":[[17368,17436]],"rname":"68"}
```

Format requirements:

- No line breaks in the JSON text
- Must include braces { }
- Use double quotes for strings
- Don't use quotes for numerical values
- Keys are case-sensitive
- Following names cannot be used as keys and will be ignored:
 - **chr**
 - **start**
 - **stop**
 - **canvas**

See full JSON spec at <http://json.org/>

"name": STR

Value is a string.

For genes, "name" is gene symbol.

When "itemurl_appendname" is specified, "name" is required to enable clicking on an item from track display and trigger a URL .

"isoform": STR

Gene isoform accession, e.g. NM_000546 or ENST00000269305

Both name and isoform can appear in the tooltip when hovering cursor over the track display.

"strand": STR

“+” or “-”.

Unstranded if not provided.

"exon": []

Array of two-number arrays, e.g. `[[665562, 665731], [665277, 665335], [661138, 665184]]`.

Must be present for genes.

All positions are 0-based.

The stop position of an exon is the nucleotide next to the last exonic nucleotide, similar to the UCSC BED format.

- For coding genes: Value should be a union of UTRs and CDS.
- For noncoding genes: Value should be all exons.
- Exons in this array are ordered from 5' to 3'.
- Despite the presence of “utr5”, “utr3”, “coding”, “intron” attributes, the “exon” attribute is still required.

"intron": []

Same format as “exon[]”.

Required for native gene tracks.

The stop position of an intron is the first nucleotide in the exon.

"utr5": []

Same format as “exon[]”.

Required for any 5' UTRs in coding genes.

"utr3": []

Same format as “exon[]”.

Required for any 3' UTRs in coding genes.

"rnalen": INT

Base-pair length of RNA transcript.

Required for all genes.

"cdslength": INT

Base-pair length of coding region length of an mRNA transcript.

Required for coding genes.

Will include nucleotides from incomplete codon as defined by startCodonFrame.

Do not use for noncoding genes.

"codingstart": INT

Genomic position of the smaller boundary of the coding sequence.

Required for coding gene.

Do not use for noncoding gene.

"codingstop": INT

Genomic position of the bigger boundary of the coding sequence.

Required for coding gene.

"startCodonFrame": 1/2

Tells how many nucleotides the “start codon” of this transcript should be shifted for translation.

In the case of [IGKC](#), startCodonFrame=1 means it will borrow 1 nt from the previous IGKJ exons. So the first two nucleotides of IGKC will not be translated when looking at IGKJ alone.

"coding": [[start, stop], ...]

Same as “exon[]”, for coding exons only.

Set “translatecoding” to true in the track object to enable gene translation according to the coding exons as well as coding frame defined by the “coding:[]” attribute. The translation can happen when the browser is at sufficient resolution.

The first element will include nucleotides from incomplete codon as defined by startCodonFrame.

"description": STR

Some text, e.g. gene function.

"color": STR

Rendering color of this item. Use follow color names:

- Names such as “red” “green” listed in https://en.wikipedia.org/wiki/Web_colors
- #FF0000
- rgb(255,0,0)

"exon2color": [{ start, stop, color }, ...]

Optional.

Each element: { start, stop, color }

Start and stop are 0-based.

This will override the item color for the matching exon from “exon[]” array.

"category": STR

Value is string or integer. Value must be a key of the “categories” attribute of the track object.

"isoformonly": STR

Experimental fix to filter bed items by isoform, so that certain bed items will be shown under a specific isoform. Value is an isoform accession.

4. Declaring a track as a JSON object

```
{
  "type": "bedj",
  "name": "gene track",
  "file": "anno/gencode.v24.hg19.gz",
  "stackheight": 14,
  "stackspace": 1
}
```

- **stackheight:20**
 - Height of rows in number of pixels. All rows share the same height.
 - For gene tracks, this height will be the thickness of coding exons, while UTRs and noncoding exons will have height reduced by 4 pixels.
- **stackspace:1**
 - Spacing distance between rows.
- **color:"blue"**
 - Track rendering color for lines, boxes, and text labels.
 - Per-item color defined in the track file will override this setting.
- **onerow:true**
 - Value is "1" for true.
 - Forces all items in the view range to be displayed in the same row, and item names will be hidden.
 - Useful for making compact representation of certain tracks, e.g. chromHMM.
 - Delete this attribute to cancel the effect.
- **categories:{}**
 - List of categories, each item in the track will belong to one category and will be colored accordingly
 - {"1":{"color":"red","label":"type 1"}, "2":{"color":"blue","label":"type 2"}, ... }
- **translatecoding:1**
 - Will translate genes when the resolution is fine enough. This requires the .coding[] attribute in the JSON objects of BED items.
- **itemurl_appendname:** URL

- Allows clicking on an item from the track and open up a URL customized by the name of that item; item's name will be appended to the end of the URL as the value of a parameter.
- Example: given URL of "<http://google.com?query=>". When clicked on an item named "HOX", this URL will be triggered: <http://google.com?query=HOX>
- **hideItemNames:** true
 - Do not show item names in the track
- **filterByName:** "Item1\nItem2"
 - Multiple item names joined by line break. Will only show given items in the track.