

How to add gene & ORF names to velocity templates

Pierre-Andre Michel <Pierre-Andre.Michel@sib.swiss>

ven 18/03/2022 14:41

A:Vincenzo Daponte <Vincenzo.Daponte@unige.ch>;

Cc:calipho-dev <calipho-dev@sib.swiss>;

Hello Vincenzo,

Below you'll find examples and elements you need to implement the gene names for RDF.

If you have questions about the API you can ask Valentine during my holidays.

When you're done, please push your changes in sub-branches under develop branch, as discusses one branch by option.

1. The relevant template to extend is overview.ttl.vm
2. Class Overview.java has a method called geneNames() that you can use to retrieve the names as a list of EntityName.java classes
3. Each gene name object (see EntityName.java) gives you:
 - a. the recommended name at the top level
 - b. the alternate gene names + ORF names in the getSynonyms()
4. I attach a list of protein entry examples for which you have both rec / alt and ORF names
5. the json structure of an overview which reflects the properties / methods you can use in velocity templates

Cheers,
Pam

Entry examples

protein accession	name type	name
NX_A2IDD5	recommended gene name	CCDC78
NX_A2IDD5	alternative gene name	C16orf25
NX_A2IDD5	ORF	JFP10
NX_A4D1E9	recommended gene name	GTPBP10
NX_A4D1E9	alternative gene name	OBGH2
NX_A4D1E9	ORF	UG0751c10
NX_A6QL64	recommended gene name	ANKRD36
NX_A6QL64	alternative gene name	KIAA1641
NX_A6QL64	alternative gene name	ANKRD36A
NX_A6QL64	ORF	UNQ2430/PRO499
NX_O75121	recommended gene name	MFAP3L
NX_O75121	alternative gene name	KIAA0626
NX_O75121	ORF	HSD-39
NX_O75121	ORF	HSD39

Call to overview with json response elements of interest

https://api.nextprot.org/entry/NX_075121/overview.json

```
{
  properties: {},
  uniqueName: "NX_075121",
  overview:
  {
    ◦ history: {...},
    ◦ families: [ ... ],
    ◦ proteinNames: [...],
    ◦ geneNames: <= list of gene names (1 recommended + 0-N synonyms by gene)
      [
        {
          ■ clazz: "GENE_NAMES",
          ■ type: "gene name",
          ■ qualifier: null,
          ■ id: "PR_1168821",
          ■ category: "gene name",
          ■ name: "MFAP3L",
          ■ parentId: null,
          ■ mainEntityName: null,
          ■ main: true, <= main = true means "recommended"
          ■ composedName: "gene name",
          ■ value: "MFAP3L"
          ■
          ■ synonyms: <= ORF names and alternative names are in synonyms
            [
              {
                ■ clazz: "GENE_NAMES",
                ■ type: "open reading frame",
                ■ qualifier: null,
                ■ id: "PR_1168823",
                ■ category: "ORF", <= ORF name
                ■ name: "HSD-39",
                ■ parentId: "PR_1168821",
                ■ mainEntityName: null,
                ■ synonyms: [ ],
                ■ otherRecommendedEntityNames: [ ],
                ■ main: false,
                ■ composedName: "open reading frame",
                ■ value: "HSD-39"
              },
              {
                ■ clazz: "GENE_NAMES",
                ■ type: "open reading frame",
                ■ qualifier: null,
                ■ id: "PR_1168820",
                ■ category: "ORF", <= other ORF name
                ■ name: "HSD39",
                ■ parentId: "PR_1168821",
                ■ mainEntityName: null,
                ■ synonyms: [ ],
                ■ otherRecommendedEntityNames: [ ],
                ■ main: false,
                ■ composedName: "open reading frame",
                ■ value: "HSD39"
              },
              {
                ■ clazz: "GENE_NAMES",
```

```
    ■ type: "gene name",
    ■ qualifier: null,
    ■ id: "PR_1168822",
    ■ category: "gene name", <= alternative gene name
    ■ name: "KIAA0626",
    ■ parentId: "PR_1168821",
    ■ mainEntityName: null,
    ■ synonyms: [ ],
    ■ otherRecommendedEntityNames: [ ],
    ■ main: false,
    ■ composedName: "gene name",
    ■ value: "KIAA0626"
  }...
```

1.