Protein Naming Utility

http://www.jcvi.org/pn-utility

The PNU is a web based database for storing and applying naming rules to identify and correct syntactically incorrect protein names or to replace synonyms with their preferred name.

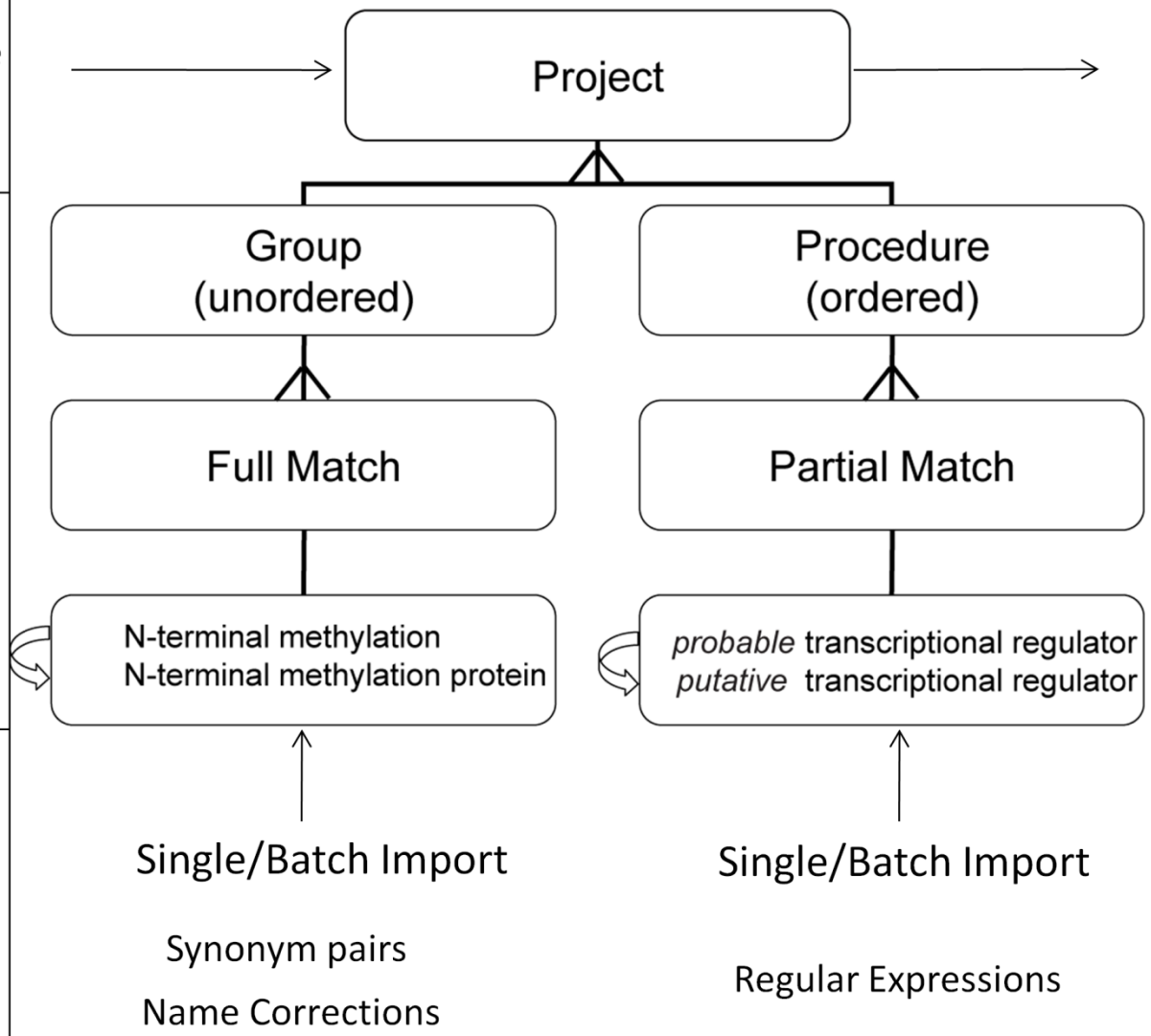
Protein Naming Utility
conceptual overview

Correct Names

Input Names
- GenBank Table File
- List of names

Organize Rules

Enter Rules



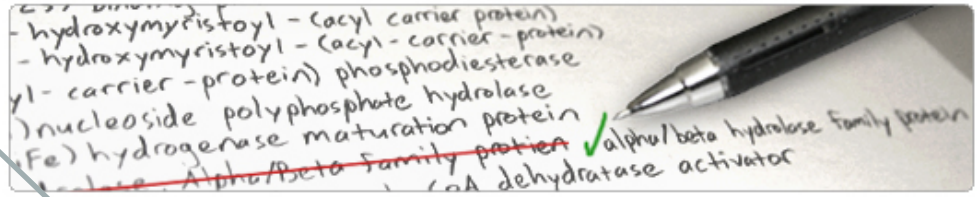
Report

Corrected Names
- GenBank Table File
- List of names

J. Craig Venter INSTITUTE

The Protein Naming Utility

JCVI Home | New Account | Log In



- Correct Names
- Enter Naming Rules
- Download

PROTEIN NAMING UTILITY

Create new PNU account to enter rules

About

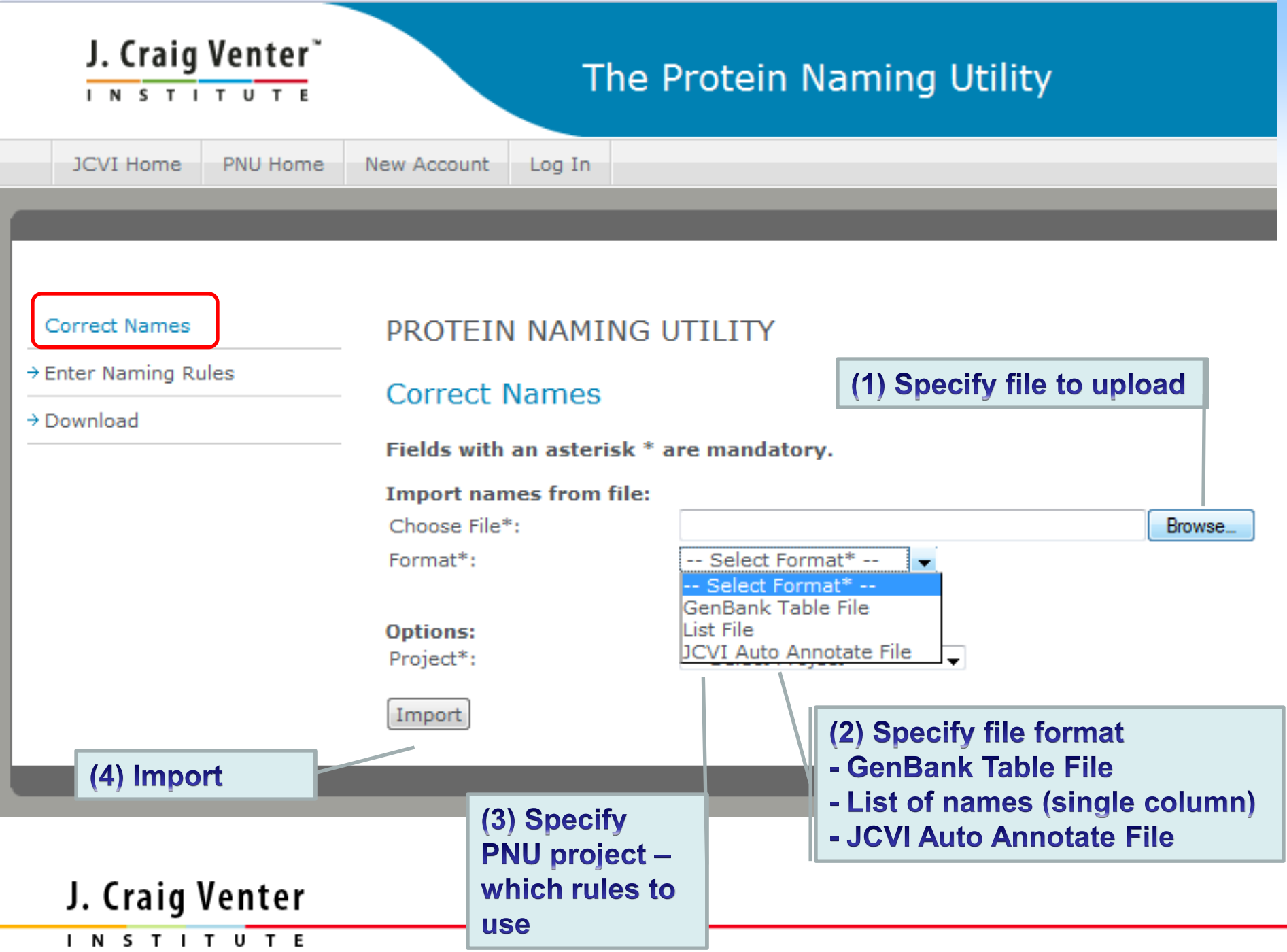
Correct a set of input names using JCVI rules - Accessible without an account

Generation of syntactically co…notation processes. Proteins are often …names. To address the need to generate high quality protein names, and capture our significant experience correcting protein names manually, we have developed the Protein Naming Utility (PNU). The PNU is a web based database for storing and applying naming rules to iden…ein names or to replace synonyms with their preferred name. The PNU allo…s of naming rules, optionally building upon the growing body of rules genera…Since communities often enforce disparate conventions for naming proteins, the PNU supports grouping rules into user-managed collections. Users can check their protein names against a selected PNU rule collection, generating both statistics and corrected names. The PNU can also

Enter Rules – Login required

Download Rules – MySQL dump & Schema

J. Craig Venter INSTITUTE


J. Craig Venter INSTITUTE
The Protein Naming Utility
JCVI Home
PNU Home
New Account
Log In
Correct Names
Enter Naming Rules
Download
PROTEIN NAMING UTILITY
Correct Names
Fields with an asterisk * are mandatory.
Import names from file:
Choose File*:
Browse...
Format*:
-- Select Format* --
-- Select Format* --
GenBank Table File
List File
JCVI Auto Annotate File
Options:
Project*:
Import
(1) Specify file to upload
(2) Specify file format
- GenBank Table File
- List of names (single column)
- JCVI Auto Annotate File
(3) Specify PNU project – which rules to use
(4) Import
J. Craig Venter INSTITUTE

J. Craig Venter INSTITUTE

The Protein Naming Utility

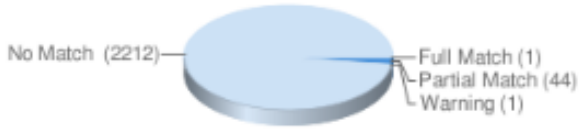
Search

JCVI Home | PNU Home | New Account | Log In

- Correct Names
  - Report
- Enter Naming Rules

Report

found 46 matches in 2258 unique names in bcb6. Executed in 0.25 seconds.



(1) Statistics

(5) Check names to correct

Full Match Suggestions check all | uncheck all

#Names	Gene Product Name	Suggested Name	Select	Log
3	D-alanyl-D-alanine carboxypeptidase	serine-type D-Ala-D-Ala carboxypeptidase	☑	Details

(2) Full Match Suggestions

Partial Match Suggestions check all | uncheck all

#Names	Gene Product Name	Suggested Name	Select	Log
2	aspartyl-tRNA synthetase	aspartyl-tRNA ligase	☐	details
2	cardiolipin synthetase	cardiolipin ligase	☐	details
2	threonyl-tRNA synthetase	threonyl-tRNA ligase	☐	details
2	histidyl-tRNA synthetase	histidyl-tRNA ligase	☐	details
2	methionyl-tRNA synthetase	methionyl-tRNA ligase	☐	details
2	tyrosyl-tRNA synthetase	tyrosyl-tRNA ligase	☐	details

(3) Partial Match Suggestions

(4) Warnings

Warnings

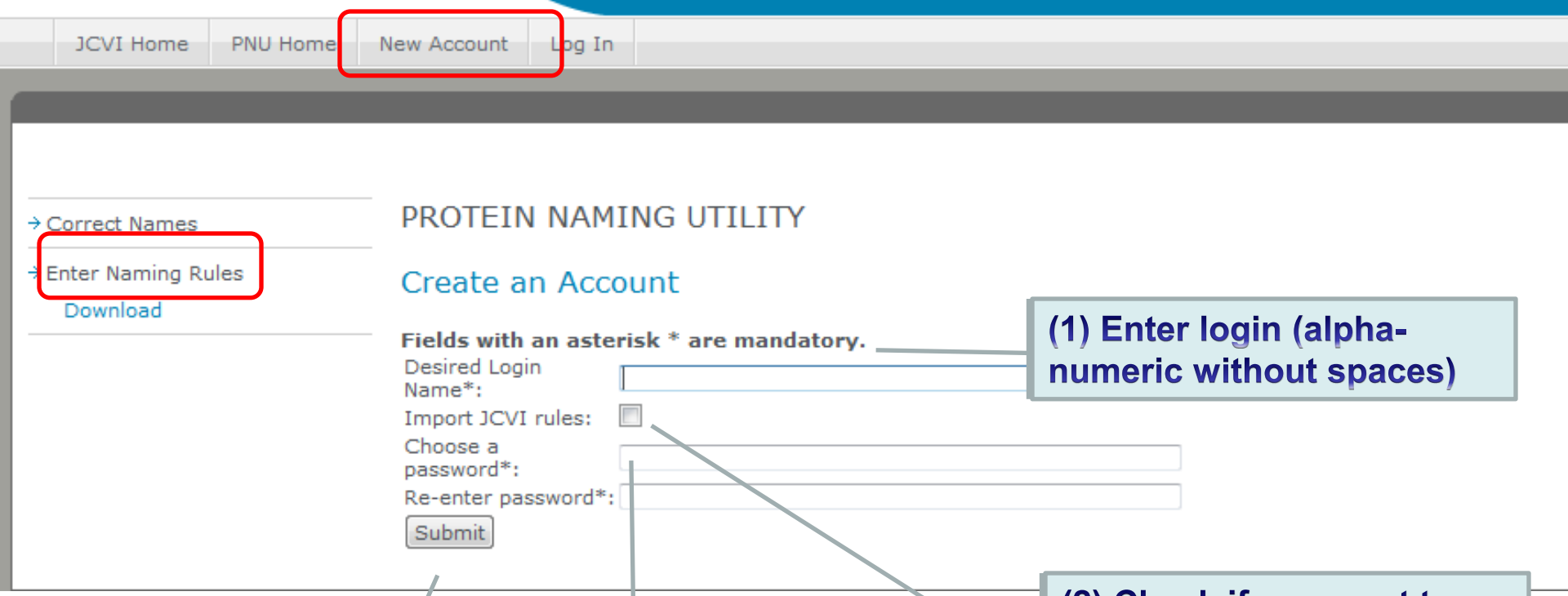
#Names	Gene Product Name	Enter Suggested Name	Log
3	TPP riboswitch (THI element)		details

J. Craig Venter INSTITUTE

Download New File

(6) Correct names in file

(5) Enter alt. names for warnings

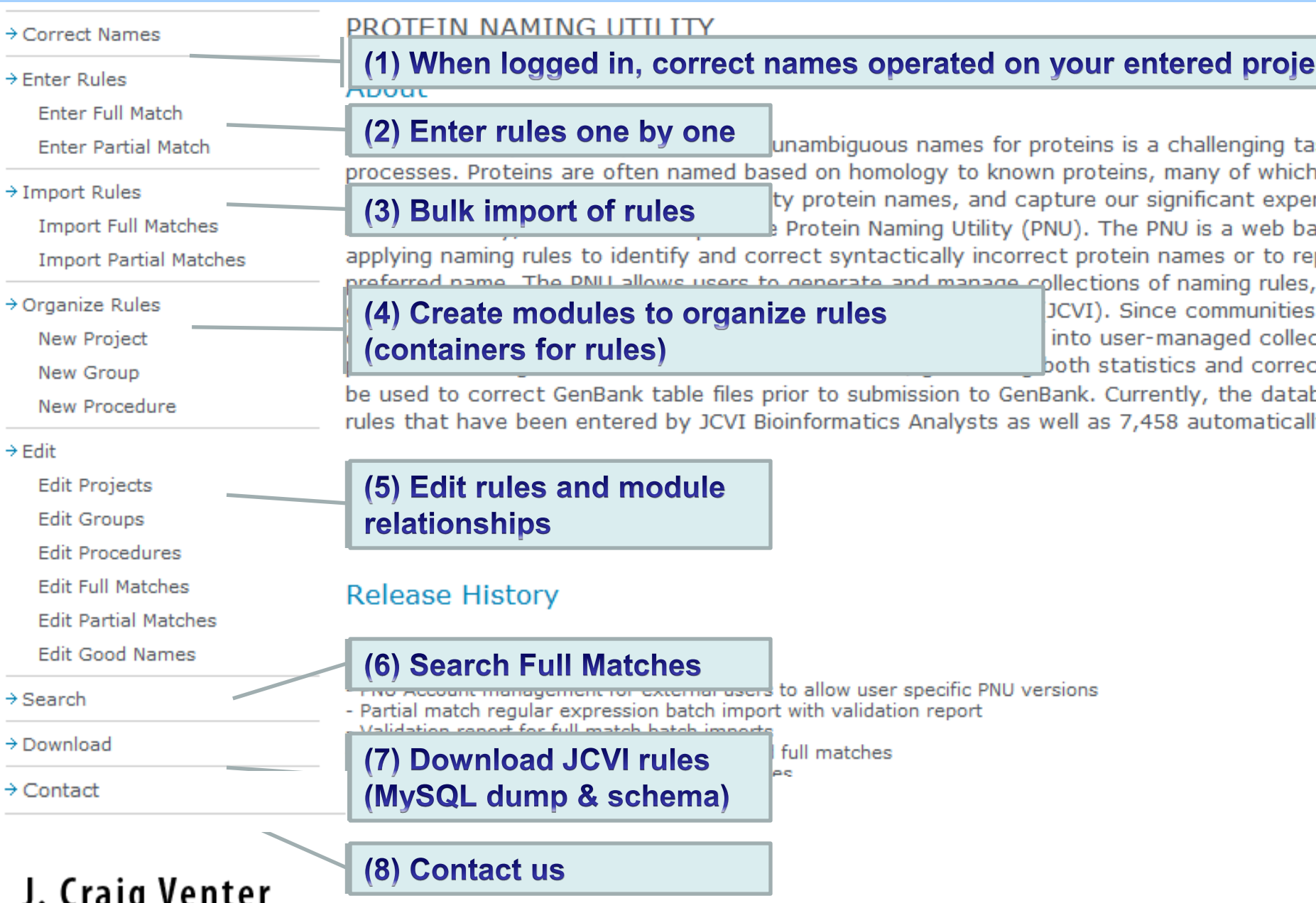


(1) Enter login (alpha-numeric without spaces)

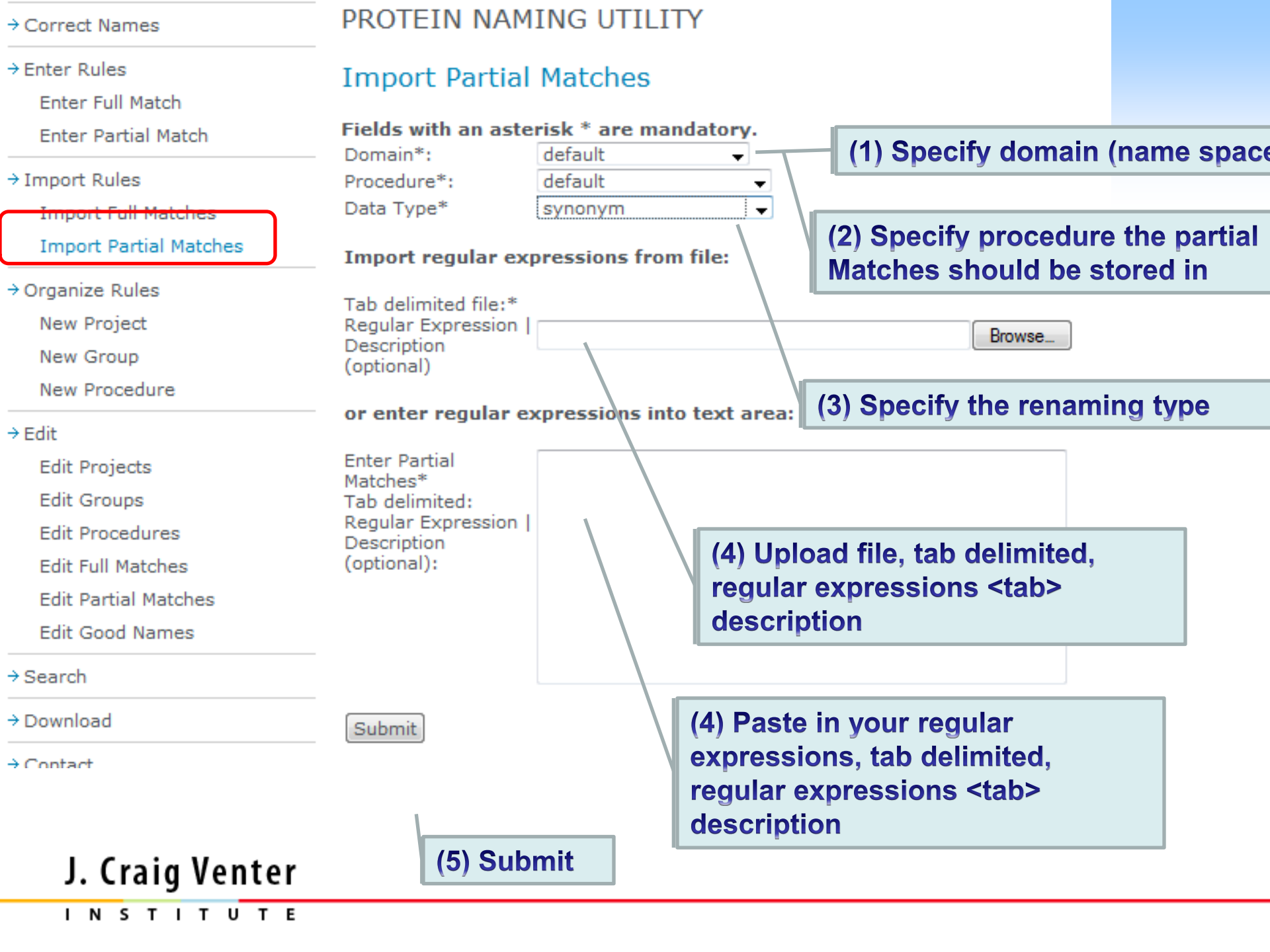
(2) Check if you want to build on JCVI rules – leave blank if you want to populate an empty system

(3) Enter user name and password

(4) Create account



J. Craig Venter
INSTITUTE


Correct Names
Enter Rules
Enter Full Match
Enter Partial Match
Import Rules
Import Full Matches
Import Partial Matches
Organize Rules
New Project
New Group
New Procedure
Edit
Edit Projects
Edit Groups
Edit Procedures
Edit Full Matches
Edit Partial Matches
Edit Good Names
Search
Download
Contact
PROTEIN NAMING UTILITY
Import Partial Matches
Fields with an asterisk * are mandatory.
Domain*:
default
Procedure*:
default
Data Type*
synonym
Import regular expressions from file:
Tab delimited file:*
Regular Expression | Description (optional)
Browse...
or enter regular expressions into text area:
Enter Partial Matches*
Tab delimited:
Regular Expression | Description (optional):
Submit
(1) Specify domain (name space
(2) Specify procedure the partial Matches should be stored in
(3) Specify the renaming type
(4) Upload file, tab delimited, regular expressions <tab> description
(4) Paste in your regular expressions, tab delimited, regular expressions <tab> description
(5) Submit
J. Craig Venter
INSTITUTE

PROTEIN NAMING UTILITY

Validate Partial Matches

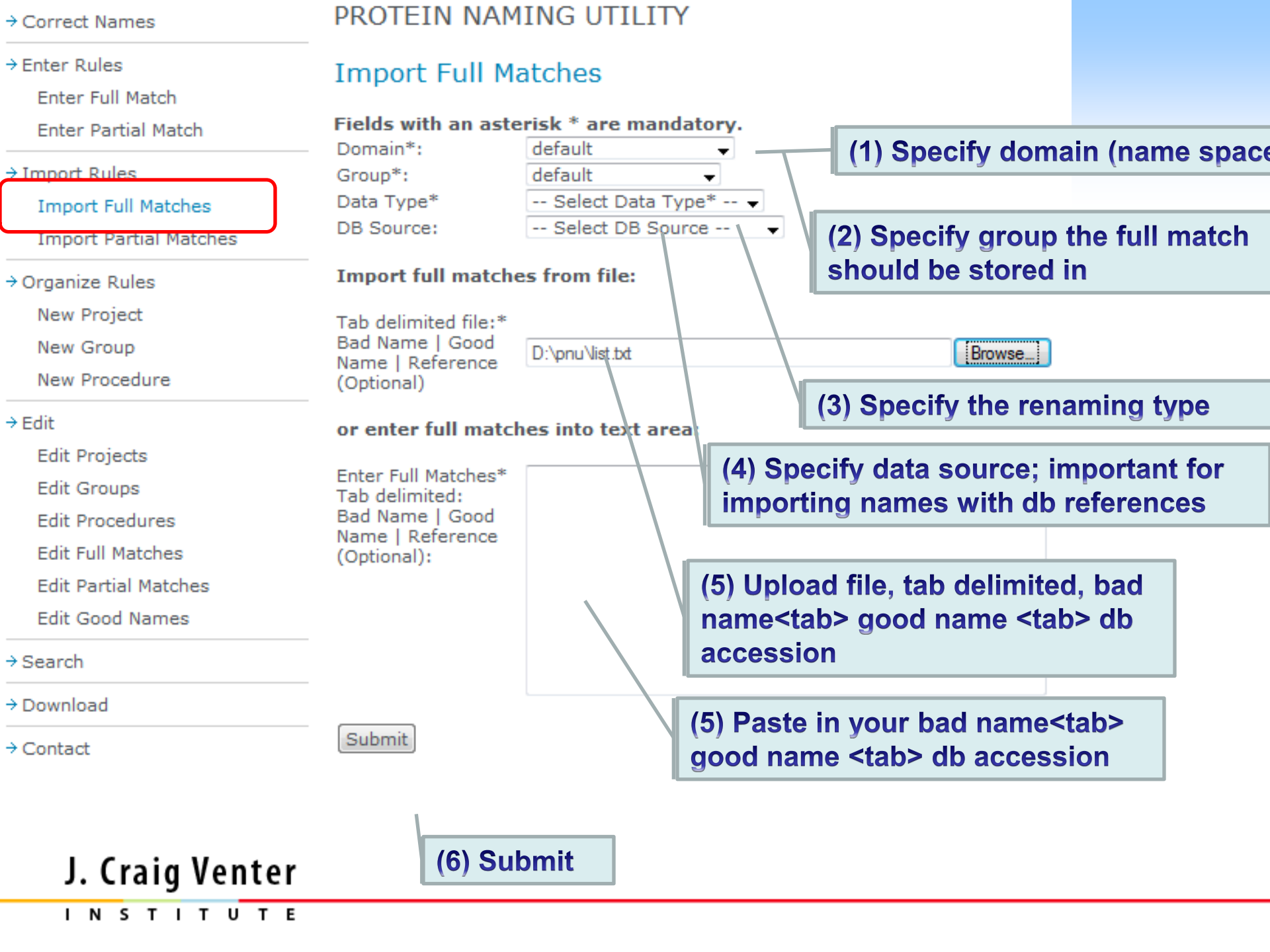
Regular Expression	Description	Status	
/NADPH:adrenodoxin oxidoreductase, \(adrenodoxin reductase\) \(ar\) \(ferredoxin reductase\)\(ferredoxin--nadp\(\+\) reductase\)/NADPH:adrenodoxin oxidoreductase/		✓	Match value = /NADPH:adrenodoxin oxidoreductase, \(adrenodoxin reductase\) \(ar\) \(ferredoxin reductase\)\(ferredoxin--nadp\(\+\) reductase\)/, Replace value = NADPH:adrenodoxin oxidoreductase
/^tRNA$/tRNA-Xxx/		✓	Match value = /^tRNA$/, Replace value = tRNA-Xxx
/^conserved protein.*/conserved hypothetical protein/i		✓	Match value = /^conserved protein.*/i, Replace value = conserved hypothetical protein
/^ABC transporter, CydDC cysteine exporter (CydDC-E) family, permease/ATP-binding protein CydC\/ATP-binding\/permease protein cydC/i		✓	Match value = /^ABC transporter, CydDC cysteine exporter (CydDC-E) family, permease/i, Replace value = ATP-binding protein CydC\/ATP-binding \/permease protein cydC
/^ABC transporter, CydDC cysteine exporter (CydDC-E) family, permease/ATP-binding protein CydD\/ATP-binding\/permease protein cydD/i		✓	Match value = /^ABC transporter, CydDC cysteine exporter (CydDC-E) family, permease/i, Replace value = ATP-binding protein CydD\/ATP-binding \/permease protein cydD

(●) Cancel Batch Upload ( ) Submit Batch Upload

Continue

(1) Check the validation report and check submit or cancel.

(2) Confirm


Correct Names
Enter Rules
Enter Full Match
Enter Partial Match
Import Rules
Import Full Matches
Import Partial Matches
Organize Rules
New Project
New Group
New Procedure
Edit
Edit Projects
Edit Groups
Edit Procedures
Edit Full Matches
Edit Partial Matches
Edit Good Names
Search
Download
Contact
PROTEIN NAMING UTILITY
Import Full Matches
Fields with an asterisk * are mandatory.
Domain*:
default
Group*:
default
Data Type*
-- Select Data Type* --
DB Source:
-- Select DB Source --
Import full matches from file:
Tab delimited file:*
Bad Name | Good Name | Reference (Optional)
D:\pnu\list.txt
Browse...
or enter full matches into text area:
Enter Full Matches*
Tab delimited:
Bad Name | Good Name | Reference (Optional):
Submit
(1) Specify domain (name space
(2) Specify group the full match should be stored in
(3) Specify the renaming type
(4) Specify data source; important for importing names with db references
(5) Upload file, tab delimited, bad name<tab> good name <tab> db accession
(5) Paste in your bad name<tab> good name <tab> db accession
(6) Submit
J. Craig Venter
INSTITUTE

- Correct Names
- Enter Rules
  - Enter Full Match
  - Enter Partial Match
- Import Rules
  - Import Full Matches
  - Validate Full Matches
  - Import Partial Matches
- Organize Rules
  - New Project
  - New Group
  - New Procedure
- Edit
  - Edit Projects
  - Edit Groups
  - Edit Procedures
  - Edit Full Matches
  - Edit Partial Matches
  - Edit Good Names
- Search

Validate Full Matches

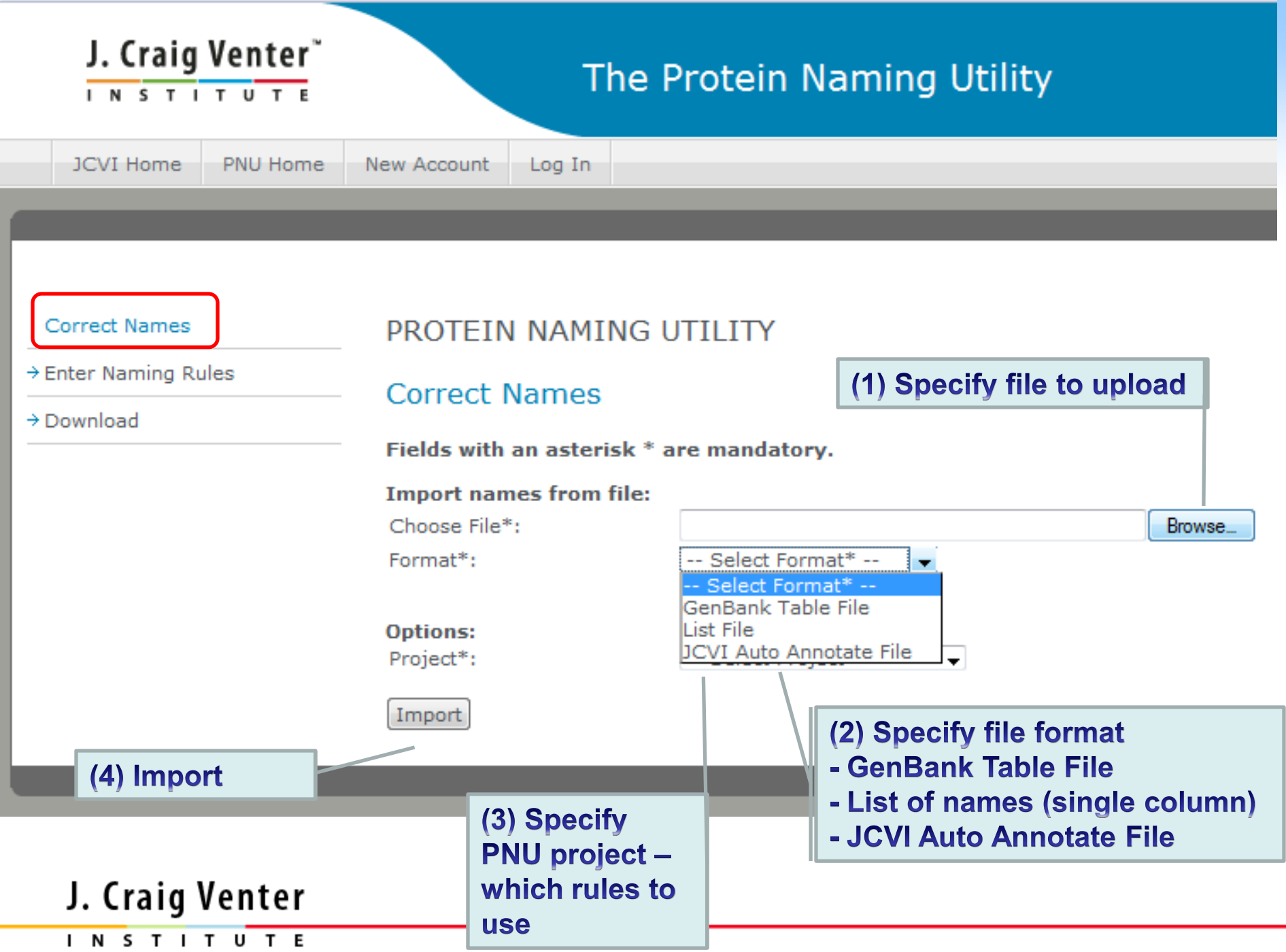
Bad Name	Good Name	Reference	Status	
DNA polymerase III, beta chain	DNA polymerase III, beta subunit		✓	valid
DNA gyrase subunit B	DNA gyrase, B subunit		✓	valid
DNA gyrase subunit A	DNA gyrase, A subunit		✓	valid
DNA-directed RNA polymerase omega chain	DNA-directed RNA polymerase, omega subunit		✓	valid
v-type sodium ATP synthase subunit D	V-type sodium ATPase, D subunit		✓	valid
v-type sodium ATP synthase subunit B	V-type sodium ATPase, B subunit		✓	valid
v-type sodium ATP synthase subunit A	V-type sodium ATPase, catalytic A subunit		✓	valid
v-type sodium ATP synthase subunit G	V-type sodium ATPase, G subunit		✓	valid
v-type sodium ATP synthase subunit C	V-type sodium ATPase, C subunit		✓	valid
v-type sodium ATP synthase subunit E	V-type sodium ATPase, E subunit		✓	valid
v-type sodium ATP synthase subunit K	V-type sodium ATPase, K subunit		✓	valid
V-type sodium ATP synthase subunit I	V-type sodium ATPase, I subunit		✓	valid

(●) Cancel Batch Upload ( ) Submit Batch Upload

Continue

(1) Check the validation report and check submit or cancel

(2) Confirm


J. Craig Venter INSTITUTE
The Protein Naming Utility
JCVI Home
PNU Home
New Account
Log In
Correct Names
Enter Naming Rules
Download
PROTEIN NAMING UTILITY
Correct Names
Fields with an asterisk * are mandatory.
Import names from file:
Choose File*:
Browse...
Format*:
-- Select Format* --
-- Select Format* --
GenBank Table File
List File
JCVI Auto Annotate File
Options:
Project*:
Import
(1) Specify file to upload
(2) Specify file format
- GenBank Table File
- List of names (single column)
- JCVI Auto Annotate File
(3) Specify PNU project – which rules to use
(4) Import
J. Craig Venter INSTITUTE