## Sequence analysis

Advance Access publication October 12, 2010

# **PICMI:** mapping point mutations on genomes

Loredana Le Pera<sup>1</sup>, Paolo Marcatili<sup>1</sup> and Anna Tramontano<sup>1,2,\*</sup>

<sup>1</sup>Department of Biochemical Sciences and <sup>2</sup>Istituto Pasteur – Fondazione Cenci Bolognetti, Sapienza University of Rome, P.le A. Moro, 5 -00185 Rome

Associate Editor: Alfonso Valencia

## ABSTRACT

**Motivation:** Several international collaborations and local projects are producing extensive catalogues of genomic variations that are supplementing existing collections such as the OMIM catalogue. The flood of this type of data will keep increasing and, especially, it will be relevant to a wider user base, including not only molecular biologists, geneticists and bioinformaticians, but also clinical researchers. Mapping the observed variations, sometimes only described at the amino acid level, on a genome, identifying whether they affect a gene and—if so—whether they also affect different isoforms of the same gene, is a time consuming and often frustrating task.

**Results:** The PICMI server is an easy to use tool for quickly mapping one or more amino acid or nucleotide variations on a genome and its products, including alternatively spliced isoforms.

**Availability:** The server is available at www.biocomputing.it/picmi **Contact:** anna.tramontano@uniromal.it

Received on July 2, 2010; revised on September 20, 2010; accepted on September 21, 2010

## **1 INTRODUCTION**

The availability of novel high-throughput technologies for identifying variations, both pathological and physiological, in sequenced genomes is producing a wealth of data that is readily available to researchers.

These data will continue to be produced at an unprecedented speed not only in projects based on large international collaborations, but also in individual labs and will add to existing collections such as OMIM (Amberger *et al.*, 2009), SwissProt (The UniProt Consortium, 2010) and the related mutation portal SwissVar (Mottaz *et al.*, 2010).

It can be easily foreseen not only that more and more data will be available, but also that the scientists who will need to access and analyze them will not be limited to molecular biologists, geneticists and bioinformaticians, as it has been mostly the case so far, but will include clinical researchers and in the future also medical doctors. This implies that tools to easily access and interpret these data should be provided to the community and that they have to be simple, reliable and user-friendly.

Given one or more variations of interest, one needs to map them back to the corresponding genome, verify in which region they fall and, if they map to a coding region, understand whether they affect, and in which way, one or more of the isoforms of the gene. This task is not made easier by the fact that the version of the genome might have changed since the time of identification of the mutation. Less straightforward is the analysis of an amino acid mutation when the corresponding nucleotide variation is not reported, as is the case for several instances in OMIM (Amberger *et al.*, 2009) and for those in the SwissVar collection (Mottaz *et al.*, 2010).

At present, Ensembl (Hubbard *et al.*, 2009) can be used to retrieve the location of nucleotide variations, by installing the relevant APIs and locally running a perl script. Associated web-based tools such as the one described in McLaren *et al.* (2010) can perform the mapping of nucleotide variations, but not of amino acid variations. For the latter, the corresponding nucleotide variations can only be retrieved, for example using SIFT (Kumar *et al.*, 2009), when they correspond to a known SNP, stored for example in dbSNP (Sherry *et al.*, 2001).

To address this conceptually easy, but technically time consuming and often frustrating problem, we developed the PICMI (Perhaps I Can Map It) server.

The server can map nucleotide variations on the human, mouse, rat and chicken genomes (altogether accounting for more than three quarters of the annotated variations) and on their different versions, report in which region they map and, when they fall in a coding region, provide information on their location on all isoforms of the gene, if any. Notably, the user can also input one or more amino acid variations for proteins in the UniProt database. In this case the system maps them back to the genome and infers, whenever this can be done unambiguously, the corresponding nucleotide variations that are subsequently analysed as described above.

## 2 DESCRIPTION

The server allows the selection of the relevant species and, if more than one genome assembly exists, of the specific version from Ensembl. Multiple nucleotide and amino acid variations can be used as input (Fig. 1).

Nucleotide variations are identified by their position on a chromosome and by the wild-type and mutated nucleotide. The server uses the information on the wild-type nucleotide to identify the correct strand and to verify that the selected base is indeed present in the correct position of the selected version of the specific genome. The VCF 1000 genome format can be selected as input as well by checking the appropriate box.

Unless the input position falls in an intergenic region, the tool will map it with respect to the transcript(s) and report whether it falls upstream, downstream, in the 5' or 3' untranslated region, in a stop-codon, in a skipped exon or in a coding exon. In the last case, the mutation is mapped on all the isoforms of the gene. The variation is assigned to the synonymous, nonsense or missense category and, in the latter case, the system provides the wild-type and mutated amino acid in each of the isoforms.

<sup>\*</sup>To whom correspondence should be addressed.

<sup>©</sup> The Author(s) 2010. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/ by-nc/2.5), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

to acid variations int the species (Homo sapiens ) and the genome version (current version (Ensembli 58) ) at (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot uence, wild type amino acid, mutated amino acid, optional identifier, separated by co icolons or tabs.	o acid variations         it the species (Hemo septens ()) and the genome version (current version (insembl 58) ())         it (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot accession number, position on the Swiss-Prot accession number, position on the Swiss-Prot colors or tabs.         97       8       0       VAR 023064       Or upload a text file:	acid variations         t the species (Home sapiens (Home	acid variations         he species       Here septents       and the genome version       current version (insembl 58)       ()         ane variation per line):       Swiss-Prot accession number, position on the Swiss-Prot accession number, position on the Swiss-Prot accession acid, potional identifier, separated by a ons or tabs.         B       c       W WAR 023004       Or upload a text file:	o acid variations  t the species (Heno septens ) and the genome version (current version (insembl 50 )  t (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot eence, wild type amino acid, mutated amino acid, optional identifier, separated by o colons or tabs.  77 58 c W VAR 022004 Or upload a text file:	acid variations         the species       Homo suppose         (one variation per line):       Swiss-Prot accession number, position on the Swiss-Prot accession number, position on the Swiss-Prot accession or tabs.         7       56       V       VAR. 923064       Or upload a text file:	to acid variations  t the species Homo sapiens and the genome version current version (insembl SB) and the genome version (insemble SB) are constrained and the genome version (insemble SB) are constrained at the genome version (insemble SB) are	t the species (Hemo sapems ) and the genome version (current version (insembli 58) ) ( (one variation per line); Swiss-Prot accession number, position on the Swiss-Prot ence, wild type amino acid, mutated amino acid, optional identifier, separated by colons or tabs.	o acid variations t the species (Homo sapiens ) and the genome version (current version (Ensembl 56) ) (one variation per line): Swiss-Prot accession number, position on the Swiss-Pro ence, wild type amino acid, mutated amino acid, optional identifier, separated by polons or tabs.
tot the species (Homo suplens ) and the genome version (current version (Ensembl 58) ) it (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot uence, wild type amino acid, mutated amino acid, optional identifier, separated by co icolons or tabs.	t the species (Homo sapens ) and the genome version (current version (insembl 58) ) ( (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot enco, wild type amino acid, mutated amino acid, optional identifier, separated by colors or tabs.	t the species Homo superse c) and the genome version (current version (threemblisis) c) ( (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot ance, wild type amino acid, mutated amino acid, optional identifier, separated by o olons or tabs.	he species (Homo sapiens ) and the genome version (current version (thremb158) ) one variation per line): Swiss-Prot accession number, position on the Swiss-Prot uco, wild type amino acid, mutated amino acid, optional identifier, separated by o ons or tabs.	tt the species (Hemo sapiens ) and the genome version (current version (fusemblish) ) t (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot ence, wild type amino acid, mutated amino acid, optional identifier, separated by o colons or tabs.	the species Homo supiens ; and the genome version current version (insembl 58); ( (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot noc, wild type amino acid, mutated amino acid, optional identifier, separated by c plons or tabs.	ct the species items supiens and the genome version (current version (charmed 5.5) )	t the species (Homo supiens ) and the genome version (current version (threembl 58) ;) ( (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot ence, wild type amino acid, mutated amino acid, optional identifier, separated by coordinate to the set of the set	t the species (Homo saplens ) and the genome version (current version (Insembl 56) ) (one variation per line): Swiss-Prot accession number, position on the Swiss-Pro ence, wild type amino acid, mutated amino acid, optional identifier, separated by jolons or tabs.
tc the species (Home sapiens ) and the genome version (current version (thread to )) It (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot uence, wild type amino acid, mutated amino acid, optional identifier, separated by co colons or tabs.	t the species [Homo sapiens ] and the genome version [current version itement[58] ] (one variation per line): Swiss-Prot accession number,position on the Swiss-Prot accession number, position number, p	t the species <u>Homo septens</u> and the genome version <u>current version (insemble 58)</u> ( (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot ance, wild type amino acid, mutated amino acid, optional identifier, separated by c olors or tabs.	he species Homo sapins   and the genome version (current version (insemble SB 2)  ne variation per line): Swiss-Prot accession number, position on the Swiss-Prot accevild type amino acid, mutated amino acid, optional identifier, separated by o ons or tabs.  B C N VAR 023004 Or upload a text file:	t the species Hemo supers and the genome version (current version (insembl 58) ) t (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot sence, wild type amino acid, mutated amino acid, optional identifier, separated by o coins or tabs. 97 58 c W VAR 022004 Or upload a text file:	the species terms appens 2 and the genome version current version (current version terms to be compared to be c	tt (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot accession number, position on the Swiss-Prot uence, wild type amino acid, mutated amino acid, optional identifier, separated by coolons or tabs.	t the species [Homo sapiens ] and the genome version [current version (insemble 58) ]	t the species Homo supiers and the genome version (current version (Curren
ut (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot uence, wild type amino acid, mutated amino acid, optional identifier, separated by co icolons or tabs.	t (one variation per line): Swiss-Prot accession number,position on the Swiss-Prot ence,wild type amino acid,mutated amino acid,optional identifier, separated by c colons or tabs.	(one variation per line): Swiss-Prot accession number, position on the Swiss-Prot ence, wild type amino acid, mutated amino acid, optional identifier, separated by o olons or tabs.	one variation per line): Swiss-Prot accession number,position on the Swiss-Prot ice,wild type amino acid,mutated amino acid,optional identifier, separated by o ons or tabs.	t (one variation per line): Swiss-Prot accession number,position on the Swiss-Prot sence,wild type amino acid,mutated amino acid,optional identifier, separated by o colons or tabs. 197 S8 C W VAR 022004 Or upload a text file:	(one variation per line): Swiss-Prot accession number,position on the Swiss-Prot ince,wild type amino acid,mutated amino acid,optional identifier, separated by c plons or tabs.	It (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot uence, wild type amino acid, mutated amino acid, optional identifier, separated by c icolons or tabs.	(one variation per line): Swiss-Prot accession number, position on the Swiss-Prot ence, wild type amino acid, mutated amino acid, optional identifier, separated by c colons or tabs.	: (one variation per line): Swiss-Prot accession number, position on the Swiss-Pro ence, wild type amino acid, mutated amino acid, optional identifier, separated by colons or tabs.
It (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot uence, wild type amino acid, mutated amino acid, optional identifier, separated by co icolons or tabs.	t (one variation per line): Swiss-Prot accession number,position on the Swiss-Prot ence,wild type amino acid,mutated amino acid,optional identifier, separated by c colors or tabs.	(one variation per line): Swiss-Prot accession number,position on the Swiss-Prot ence,wild type amino acid,mutated amino acid,optional identifier, separated by o olons or tabs.	wind variation per line): Swiss-Prot accession number, position on the Swiss-Prot accession number, position number, positi	t (one variation per line): Swiss-Prot accession number,position on the Swiss-Prot ence.wild type amino acid,mutated amino acid,optional identifier, separated by o colons or tabs.	(one variation per line): Swiss-Prot accession number,position on the Swiss-Prot nec,wild type amino acid,mutated amino acid,optional identifier, separated by c plons or tabs.	it (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot uence, wild type amino acid, mutated amino acid, optional identifier, separated by c icolons or tabs.	(one variation per line): Swiss-Prot accession number, position on the Swiss-Prot ence, wild type amino acid, mutated amino acid, optional identifier, separated by c zolons or tabs.	(one variation per line): Swiss-Prot accession number, position on the Swiss-Prot ence, wild type amino acid, mutated amino acid, optional identifier, separated by solons or tabs.
It (one variation per line): Swiss-Prot accession number,position on the Swiss-Prot uence,wild type amino acid,mutated amino acid,optional identifier, separated by co icolons or tabs.	t (one variation per line): Swiss-Prot accession number,position on the Swiss-Prot ence,wild type amino acid,mutated amino acid,optional identifier, separated by o colons or tabs.	(one variation per line): Swiss-Prot accession number, position on the Swiss-Prot ence, wild type amino acid, mutated amino acid, optional identifier, separated by o olons or tabs.	one variation per line): Swiss-Prot accession number,position on the Swiss-Prol ce,wild type amino acid,mutated amino acid,optional identifier, separated by ons or tabs.	t (one variation per line): Swiss-Prot accession number,position on the Swiss-Prot sence,wild type amino acid,mutated amino acid,optional identifier, separated by o colons or tabs. 197 S8 C W VAR 023004 Or upload a text file:	(one variation per line): Swiss-Prot accession number,position on the Swiss-Prot accession number,position on the Swiss-Prot accession number,position on the Swiss-Prot accession number, position number, positi numer, position number, position numer, position number,	It (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot uence, wild type amino acid, mutated amino acid, optional identifier, separated by o icolons or tabs.	: (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot ence, wild type amino acid, mutated amino acid, optional identifier, separated by c colons or tabs.	: (one variation per line): Swiss-Prot accession number, position on the Swiss-Pro ence, wild type amino acid, mutated amino acid, optional identifier, separated by colons or tabs.
at (one variation per line): Swiss-Prot accession number,position on the Swiss-Prot uence,wild type amino acid,mutated amino acid,optional identifier, separated by co icolons or tabs.	t (one variation per line): Swiss-Prot accession number,position on the Swiss-Prot lence,wild type amino acid,mutated amino acid,optional identifier, separated by c colors or tabs.	(one variation per line): Swiss-Prot accession number,position on the Swiss-Prot ence,wild type amino acid,mutated amino acid,optional identifier, separated by o olons or tabs.	one variation per line): Swiss-Prot accession number,position on the Swiss-Prot ice,wild type amino acid,mutated amino acid,optional identifier, separated by ons or tabs.	t (one variation per line): Swiss-Prot accession number,position on the Swiss-Prot sence,wild type amino acid,mutated amino acid,optional identifier, separated by o colons or tabs. 197 SB C W VAR 022004 Or upload a text file:	(one variation per line): Swiss-Prot accession number,position on the Swiss-Prot ance,wild type amino acid,mutated amino acid,optional identifier, separated by c plons or tabs. 7 88 C W VAR, 623004 Or upload a text file:	It (one variation per line): Swiss-Prot accession number, position on the Swiss-Prot uence, wild type amino acid, mutated amino acid, optional identifier, separated by c colons or tabs.	: (one variation per line): Swiss-Prot accession number,position on the Swiss-Prot ence,wild type amino acid,mutated amino acid,optional identifier, separated by c colons or tabs.	: (one variation per line): Swiss-Prot accession number,position on the Swiss-Prot ence,wild type amino acid,mutated amino acid,optional identifier, separated by colons or tabs.
a (one vanadon per inte): Swiss-Prot accession number, position on the Swiss-Prot uence, wild type amino acid, mutated amino acid, optional identifier, separated by co icolons or tabs.	( one vehacion per inte): swiss-prot accession number,position on the swiss-prot ence,wild type amino acid,mutated amino acid,optional identifier, separated by c colons or tabs.	(one venacion per line): SWISP-Prot accession number,position on the SWISP-Prot ence,wild type amino acid, mutated amino acid, optional identifier, separated by t olons or tabs.	c wids type amino acid, mutated amino acid, optional identifier, separated by one of tabs.      c w VAR 023004 Or upload a text file:	t (one vanadom per lane): swiss-rrot accession number,position on the swiss-rrot ence,wild type amino acid,mutated amino acid,optional identifier, separated by i colons or tabs.	Cone vanadon per intej: swiss-prot accession number, position on the swiss-prot ence, wild type amino acid, mutated amino acid, optional identifier, separated by c plons or tabs.	It (one vanadon per inte): Swiss-Prot accession number, position on the Swiss-Prot uence, wild type amino acid, mutated amino acid, optional identifier, separated by c icolons or tabs.	(one venation per line): Swiss-Prot accession number, position on the swiss-Prot ence, wild type amino acid, mutated amino acid, optional identifier, separated by c colons or tabs.	(one venation per line): Swiss-Prot accession number, position on the Swiss-Prot ence, wild type amino acid, mutated amino acid, optional identifier, separated by colons or tabs.
uence, wild type amino acid, mutated amino acid, optional identifier, separated by co icolons or tabs.	ence, wild type amino acid, mutated amino acid, optional identifier, separated by c colons or tabs.	ence,wild type amino acid,mutated amino acid,optional identifier, separated by i plons or tabs.	cc, wild type amino acid, mutated amino acid, optional identifier, separated by ons or tabs. 88 C N VAR 023004 Or upload a text file:	ence, wild type amino acid, mutated amino acid, optional identifier, separated by i colons or tabs.	ence, wild type amino acid, mutated amino acid, optional identifier, separated by c olons or tabs. 7 88 c w VAR 523064 Or upload a text file:	uence, wild type amino acid, mutated amino acid, optional identifier, separated by c icolons or tabs.	ence, wild type amino acid, mutated amino acid, optional identifier, separated by coolons or tabs.	ence, wild type amino acid, mutated amino acid, optional identifier, separated by polons or tabs.
icolons or tabs.	colons or tabs. 97 88 C W VAR 923004 Or upload a text file:	olons or tabs.	ons or tabs.         S8         C         W         VAR 023004         Or upload a text file:	colons or tabs.	olons or tabs. 7 86 C W VAR_023004 Or upload a text file:	icolons or tabs.	colons or tabs.	colons or tabs.
897 88 C W VAR 023004 Or upload a text file:	97 88 C W VAR 023004 Or upload a text file:	7 56 C W VAR 023004 Or upload a text file:	88 C W VAR 023004 Or upload a text file:	197 88 C N VAR 023004 Or upload a text file:	7 56 C W VAR 023004 Or upload a text file:	197 58 C N VAR 023004 Or upload a text file:	on so a w use annot film	
897 88 C W VAR 023004 Or upload a text file:	97 88 C W VAR 023004 Or upload a text file:	17 88 C W VAR 023004 Or upload a text file:	58 C W VAR 023004 Or upload a text file:	197 58 C W VAR 023004 Or upload a text file:	7 88 C W VAR_023004 Or upload a text file:	197 88 C N VAR 023004 Or upload a text file:	67 60 P M Han 632664 On unload a text fil-	
oy/ oo C w vax v23004 Or upidad a text me.	197 00 C W VAR 023004 I OF UDIDAU & LEAL FILE.	VAR 023001 OF UDIDAU & LEAL THE.	oo c w vax v23004 or upload a text me.	197 be C w YAR 023004 OF upidad a LEAL THE.	vak_023004 Or upidad a text me.		The literation of the second sec	at the stand of the stand a taxt file:
200 140 P P Van 012005	409 160 H B VAR 012805	49 160 H R VAR 012805	160 H B VAR 012805	NG9 160 H R VAR 012805	160 H B VAR 012805	200 120 W T TTTT 012005	or upload a text file:	97 56 C W VAR_023004 OF upload a text file:
						THE REAL PROPERTY AND A RE	G9 160 H R VAR 012805	C9 160 H B VAR 012805
235 61 I V VAR 038312 (Browse)	35 61 I V VAR 038312 (Browse)	15 61 I V VAR 038312 (Browse)	61 I V VAR 038312 (Browse)	135 61 I V VAR 038312 (Browse)	5 61 I V VAR 038312 Browse.	235 61 I V VAR 038312 (Browse)	G9 160 H R VAR 012805 35 61 I V VAR 038312 (Browse)	G9 160 H R VAR 012805 35 61 I V VAR 038312 (Browse)
235 61 I V VAR 038312 Browse	35 61 I V VAR 038312 Browse	35 61 I V VAR_038312 Browse 9 182 C W VAR_029604	61 I V VAR 038312 Browse	135 61 I V VAR 038312 129 182 C W VAR 03864	5 61 I V VAR 038312 Browse	235 61 I V VAR 018905 799 182 C W VAR 018912 799 182 C W VAR 029604	US9 160 H R VAR 012805 35 61 I V VAR 038312 Y9 182 C W VAR 029604	ND9 160 H R VAR_012805 35 61 I V VAR_038312 79 182 C W VAR_02604
235 61 I V VAR_04802 D19 182 C W VAR_029604 ex : File.txt	35 61 I V VAR 038312 Y9 182 C W VAR 028604 ex.: File.bxt	55 61 I V VAR 038312 19 182 C W VAR 029604 ex.: File.txt	61 1 V VAR_018012 182 C W VAR_029604 ex.: File.txt	135 61 1 V VAR_038312 199 182 C W VAR_028604 ex.: File.txt	5 61 I V VAR 038312 9 182 C W VAR 029604 ex.: File.txt	Nov         n         x         vAx         0.12803           25         61         1         v         vAx         0.8812           Dr9         182         C         w         vAx         0.29604           ex.:         File.txt	UT 160 H R VAR 012805 35 61 I V VAR 013812 39 182 C W VAR 028604 ex: File.txt	udg)         160         H         R         VAR_012805           15         61         1         V         VAR_03812           19         182         C         N         VAR_029604           ex.: File.txt
215 51 I V VAR_01111 219 182 C V VAR_01111 Q4L235,865,T,A,VAR_038312 Q4L235,865,T,A,VAR_038312 ex.: File.txt	335         61         1         0         VMA_018312         IBTORNE           193         182         C         M         VMA_0294604         IBTORNE         IBTORNE           Q4L235,865,T,A,VAR_038312         ex.: File.txt         File.txt	24L235,865,T,A,VAR_038312 24L235,865,T,A,VAR_038312 24L235,865,T,A,VAR_038312	61         I         V         VAR_038312         Browse           182         C         W         VAR_028604         Browse           1L235,865,T,A,VAR_038312         ex.: File.txt	235         61         I         V         VAR_051812         Works           193         182         C         H         VAR_028604         ex.: File.txt           Q4L235,865,T,A,VAR_038312         ex.: File.txt         Ex.: File.txt	13 61 T V VAR.038312 9 182 C N VAR.028604 24L235,865,T,A,VAR_038312 ex.: File.txt	315         617         2         VAL_016502         Browse           079         682         C         x         VAL_016502         Browse           Q4L235,865,T,A,VAR_038312         ex.: File.txt	G9         160         H         R         VAR.012005           35         61         1         V         VAR.038112           182         C         W         VAR.028604           Q4L235,865,T,A,VAR_038312         ex.: File.txt	09 160 H R VAR_012805 35 61 I V VAR_039312 91 182 C W VAR_039312 Q4L235,865,T,A,VAR_038312 ex.: File.txt
225 51 I V VAR_025512 079 182 C W VAR_025604 Q4L235,865,7,A,VAR_038312 ex.: File.txt	335         61         2         VMA_018312         Browse           195         182         C         W         VAL_029601         Browse           Q4L235,865,T,A,VAR_038312         ex.: File.txt         File.txt	55         1         1         V         VAR_0589112         Browse           9         182         C         M         VAR_029604         Browse           24L235,865,T,A,VAR_038312         ex.: File.txt         Ex.         Ex.         Ex.	61 I V VAR_038312 182 C W VAR_029604 1L235,865,T,A,VAR_038312 L235,865,T,A,VAR_038312 L235,865,T,A,VAR_038312 Etowse	235         61         I         V         VAR_058512         Works           293         182         C         W         VAR_028604         ex.: File.txt           Q4L235,865,T,A,VAR_038312         ex.: File.txt         Ex.: File.txt	13 61 T V VAR.038312 9 182 C W VAR.028604 24L235,865,T,A,VAR_038312 ex.: File.txt	235         61         7         VAL         2555         100	GP         IsO         H         R         VAR. (0212005)           35         61         1         V         VAR. (0212005)         (Browse)           y3         182         C         M         VAR. (0212005)         (Browse)           Q4L235,865,T,A,VAR_038312         ex.: File.txt         Ex.: File.txt	039         160         R         R         VAA         VAA         12805         S         <
235 61 I V VAR_018012 219 182 C W VAR_022604 Q4L235,865,T,A,VAR_038312 ex.: File.txt	35         61         I         V         VRR_058012         Image: Constraint of the state o	55         61         1         V         VAR_038312         Browse           19         162         C         W         VAR_022604         ex.: File.txt           24L235,865,T,A,VAR_038312         ex.: File.txt         Erowse         ex.: File.txt	61         1         V         VAR_038312         Browse           182         C         W         VAR_0228604         ex.: File.txt	133         61         1         V         VAR_058512         Browse           273         182         C         W         VAR_028604         ex.: File.txt           Q4L235,865,T,A,VAR_038312         ex.: File.txt         Erowse         ex.: File.txt	15         61         1         V         VAR_018312         (Browse)           9         182         C         W         VAR_029604         (Browse)           24L235,865,T,A,VAR_038312         ex.: File.txt	AUS 100         II         V         VAR 012602         Browse           215         61         I         V         VAR 01812         Browse           219         182         C         W         VAR 023664         Browse         Browse           Q4L235,865,T,A,VAR_038312         ex.: File.txt         File.txt         Browse         Browse	UP 160 H R VAR.012005 35 61 I V VAR.02812 39 182 C W VAR.02812 Q4L235,865,T,A,VAR_038312 ex.: File.txt	03         160         R         R         VAR_012805           35         61         1         V         VAR_02812           23         182         C         N         VAR_02804           Q4L235,865,T,A,VAR_038312         ex.: File.txt
215 51 7 V VAR_025512 079 182 C K VAR_0256312 Q41235,865,T,A,VAR_038312 ex.: File.txt	313         61         I         V         VAR_01121         Browse           113         182         C         H         VAR_012904         Errowse           V4L235,865,T,A,VAR_038312         ex.: File.txt         File.txt	2442355665,T,A,VAR_038312 ex.: File.txt	61 I V VAR 018312 182 C K VAR 0239664 IL235,865,T,A,VAR_038312 IL235,865,T,A,VAR_038312 Ex.: File.txt	135         61         I         V         VAR_018312         Browse           139         182         C         W         VAR_028604         ex.: File.txt	15         61         T         V         VAX_018312         Browse           9         162         C         N         VAX_029604         Errowse           V4L235,865,T,A,VAR_038312         ex.: File.txt         Errowse         Errowse	015 610 1 VAR 016012 019 182 C X VAR 016012 0241235,865,T,A,VAR_038312 ex.: File.txt	UP 160 H R VAR 012805 35 61 I V VAR 038312 Y9 182 C W VAR 038312 Q4L235,865,T,A,VAR_038312 ex.: File.txt	G9         160         H         R         VAA_012805           35         61         I         V         VAA_038312           91         182         C         W         VAA_025604           Q4L235,865,T,A,VAR_038312         ex.: File.txt
215 51 I V VAR_01111 219 182 C V VAR_01111 Q4L235,865,T,A,VAR_038312 Q4L235,865,T,A,VAR_038312 ex.: File.txt	335         61         1         0         VMA_018312         IBTORNE           193         182         C         M         VMA_0294604         IBTORNE         IBTORNE           Q4L235,865,T,A,VAR_038312         ex.: File.txt         File.txt	24L235,865,T,A,VAR_038312 24L235,865,T,A,VAR_038312 24L235,865,T,A,VAR_038312	61         I         V         VAR_038312         Browse           182         C         W         VAR_028604         Browse         Browse           1L235,865,T,A,VAR_038312         ex.: File.txt         File.txt         Browse         Browse <td>235         61         I         V         VAR_058912         Workson           293         182         C         H         VAR_029604         ex.: File.txt           Q4L235,865,T,A,VAR_038312         ex.: File.txt         Ex.: File.txt</td> <td>13 61 T V VAR.038312 9 182 C N VAR.028604 24L235,865,T,A,VAR_038312 ex.: File.txt</td> <td>315         617         2         VAL_016502         Browse           079         682         C         x         VAL_016502         Browse           Q4L235,865,T,A,VAR_038312         ex.: File.txt</td> <td>G9         160         H         R         VAR.012005           35         61         1         V         VAR.038112           182         C         W         VAR.028604           Q4L235,865,T,A,VAR_038312         ex.: File.txt</td> <td>09 160 H R VAR_012805 35 61 I V VAR_039312 91 182 C W VAR_029604 Q4L235,865,T,A,VAR_038312 ex.: File.txt</td>	235         61         I         V         VAR_058912         Workson           293         182         C         H         VAR_029604         ex.: File.txt           Q4L235,865,T,A,VAR_038312         ex.: File.txt         Ex.: File.txt	13 61 T V VAR.038312 9 182 C N VAR.028604 24L235,865,T,A,VAR_038312 ex.: File.txt	315         617         2         VAL_016502         Browse           079         682         C         x         VAL_016502         Browse           Q4L235,865,T,A,VAR_038312         ex.: File.txt	G9         160         H         R         VAR.012005           35         61         1         V         VAR.038112           182         C         W         VAR.028604           Q4L235,865,T,A,VAR_038312         ex.: File.txt	09 160 H R VAR_012805 35 61 I V VAR_039312 91 182 C W VAR_029604 Q4L235,865,T,A,VAR_038312 ex.: File.txt
235 61 I V VAR_014802 D19 182 C W VAR_029604 Q4L235,865,T,A,VAR_038312 ex.: File.txt	35         61         1         V         VRE_018012         Browse           19         182         C         W         VRE_029604         ex.:         File.txt           Q4L235,865,T,A,VAR_038312         ex.:         File.txt         File.txt	55         61         I         V         VAR_038312         Browse           19         182         C         W         VAR_029604         ex.: File.bxt           Q4L235,865,T,A,VAR_038312         ex.: File.bxt         Erowse         ex.: File.bxt	61         1         V         VAR 038312         Browse           182         C         W         VAR 028604         ex.: File.txt	333         61         1         V         VAR_038312         Browse           123         182         C         W         VAR_029604         ex.: File.txt           Q4L235,865,T,A,VAR_038312         ex.: File.txt         Erowse         Erowse         Erowse	561         I         V         VAR_038312         Browse           9         182         C         W         VAR_023604         ex.: File.txt           24L235,865,T,A,VAR_038312         ex.: File.txt         Ex.: File.txt         Ex.: File.txt	Aug         Low         n         x         VAR.012602           255         61         1         V         VAR.012602         Browse           279         182         C         W         VAR.029604         Erowse         Erowse           Q4L235,865,7,A,VAR_038312         ex.: File.txt         File.txt         Erowse         Erowse         Erowse	UP 160 H R VAR.012005 35 61 I V VAR.028012 39 182 C W VAR.02812 Q4L235,865,T,A,VAR_038312 ex.: File.txt	03         160         R         R         VAX_0212805           35         61         1         V         VAX_021805           29         182         C         N         VAX_029604           Q4L235,865,T,A,VAR_038312         Ex.: File.txt
235 61 I V VAR 048012 DY9 182 C W VAR 029604 Q4L235,865,T,A,VAR_038312 ex.: File.txt	35         61         1         V         VAR_018012         Browse           19         182         C         M         VAR_029604         ex.: File.txt           Q4L235,865,T,A,VAR_038312         ex.: File.txt         Ex.: File.txt	55         61         1         V         VAR_038312         Browse           19         182         C         M         VAR_029604         Erowse         Erowse           24L235,865,T,A,VAR_038312         ex.: File.txt         Erowse         Erowse         Erowse         Erowse	61 I V VAR_038312 182 C W VAR_029604 LL235,865,T,A,VAR_038312 LL235,865,T,A,VAR_038312 LL235,865,T,A,VAR_038312	133         61         1         V         VAR_038012         Browse           131         2         K         VAR_028604         Erowse         Erowse           Q4L235,865,T,A,VAR_038312         ex.: File.txt         Erowse         Erowse         Erowse	5         61         I         V         VAR_038312         Browse           9         182         C         W         VAR_029604         Erosse         Erosse           24L235,865,T,A,VAR_038312         ex.: File.txt         File.txt	Auf         100         II         V         VAR.012603         Browse           235         61         I         V         VAR.012603         Browse         Browse           291         182         C         W         VAR.029604         Browse         Browse           Q4L235,865,T,A,VAR_038312         ex.: File.txt         Ex.: File.txt         Ex.: File.txt         Ex.: File.txt	GP         IsO         H         R         VAR. (022005)           35         61         I         V         VAR. (028112)         (Browse)           29         182         C         M         VAR. (029604)         (Browse)           Q4L235,865,T,A,VAR_038312         ex.: File.txt         Ex.: File.txt	039         160         R         R         VAA         0212005           35         61         I         V         VAA         021305           29         182         C         W         VAA         029104           Q4L235,865,T,A,VAR_038312         ex.: File.txt
235 61 1 2 VAR 036912 079 162 C V VAR 036912 079 162 C V VAR 036912 041235,665,T,A,VAR_036312 ex.: File.txt	135 61 1 V VAR 018312 179 182 C N VAR 018312 Q4L235,865,T,A,VAR 038312 ex.: File.txt	25 61 I V VAR 038312 09 182 C W VAR 038312 44235,865,T,A,VAR 038312 ex.: File.txt	61:         1         V         VAR_0513312         Browse           182         C         W         VAR_0529664         Erosse           18235,865,T,A,VAR_038312         ex.: File.txt	125         51         Y         YAR_038312         Browse           179         182         C         YAR_029604         ex.: File.txt	15 61 I V VAA 038312 9 182 C W VAA 029604 442235,665,T,A,VAR_038312 ex.: File.txt	Xiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii	UP 160 H R VAR.012805 35 61 U V VAR.01812 19 182 C V VAR.02812 Q42235665,T,A,VAR_03812 ex.: File.txt	G9 160 H R VAA 012805 35 61 I V VAA 038312 241235,865,T,A,VAR 038312 Erowsz Erowsz Erowsz Erowsz Erowsz
235 61 I V VAR_04802 D19 182 C W VAR_029604 ex : File.txt	35 61 I V VAR 038312 Y9 182 C W VAR 028604 ex.: File.bxt	55 61 I V VAR 038312 19 182 C W VAR 029604 ex.: File.txt	61 1 V VAR_018012 182 C W VAR_029604 ex.: File.txt	135 61 1 V VAR_038312 199 182 C W VAR_028604 ex.: File.txt	5 61 I V VAR 038312 9 182 C W VAR 029604 ex.: File.txt	Nov         n         x         VAR_012803           25         61         1         y         VAR_028012           Dr9         182         C         w         VAR_029604           ex:         File.txt	UT 160 H R VAR 012805 35 61 I V VAR 013812 39 182 C W VAR 028604 ex: File.txt	udg)         160         H         R         VAR_012805           15         61         1         V         VAR_03812           19         182         C         N         VAR_029604           ex.: File.txt
235 61 I V VAR 038312 DY9 182 C W VAR 029604	35 61 I V VAR 038312 Y9 182 C W VAR 029604	15 61 I V VAR 038312 19 182 C W VAR 029604	61 I V VAR 038312 182 C W VAR_029604	135 61 1 V VAR 038312 199 182 C W VAR 029604	5 61 1 V VAR_038312 9 182 C W VAR_029604	AUS 100 R VAR 012803 235 61 I V VAR 038312 DY9 182 C W VAR 029604	G9 160 H R VAR_012805 35 61 I V VAR_038312 Y9 182 C W VAR_029604	KI9         160         H         R         VAR_012805           35         61         I         V         VAR_038312         Browse           79         182         C         W         VAR_028604         Browse
235 61 I V VAR 038312 DY9 182 C W VAR 038612 Browse	35 61 I V VAR 038312 9 182 C W VAR 038604	35 61 I V VAR 038312 9 182 C W VAR 038364	61 I V VAR 038312 Browse	135 61 I V VAR 038312 199 182 C W VAR 038312 Browse	5 61 I V VAR_038312 Browse	235 61 I V VAR 018905 299 182 C W VAR 038312 299 182 C W VAR 02864	G9 160 H R VAR 012805 35 61 1 V VAR 038312 99 182 C W VAR 029604	NG9 160 H R VAR 012805 35 61 I V VAR 038912 79 182 C W VAR 029604
235 61 I V VAR 038312 (Browse)	35 61 I V VAR 038312 (Browse)	15 61 I V VAR 038312 (Browse)	61 I V VAR 038312 (Browse)	135 61 I V VAR 038312 (Browse)	5 61 I V VAR 038312 (Browse)	235 61 I V VAR 038312 (Browse)	G9 160 H R VAR 012805 35 61 I V VAR 038312 (Browse)	G9 160 H R VAR 012805 35 61 I V VAR 038312 (Browse)
AND AND A R VAR UILOUD				A TAR VIEWS		NUS 100 B R VAR 012805	G9 160 H R VAR_012805	G9 160 H R VAR_012805
					ANN IS AN TANKA		TH 160 H D USD 012905	TO 160 P P 1120 012005
200 140 B B V12 012005	419 160 H B VAR 012805	49 160 H B VAR 012805	160 H B VAR 012805	AG9 160 H B VAR 012805	9 160 H R VAR 012805	3/0 1/0 H B HBB 013005		
						and a second a second a second s	V se c w VAR_023004 Or upload a text file:	97 58 C W VAR_023004 OF upload a text file:
							VAR 023004 OF UDIOAD A TEXT THE	97 se C W VAR 023004 Or upload a text file:
							VAR 023004 OF upload a text file:	97 56 C W VAR 023004 Or upload a text file:
							VAR 023004 OF upload a text file:	97 56 C W VAR 023004 Or upload a text file:
							VAR 023004 OF upload a text file:	97 56 C W VAR 023004 Or upload a text file:
						The second	VAR 023004 OF upload a text file:	97 58 C W VAR 023004 Or upload a text file:
						The second	VAR 023004 OF upload a text file:	97 58 C W VAR 023004 Or upload a text file:
the second						The second	VAR_023004 OF upload a text file:	97 88 C W VAR_023004 Or upload a text file:
						The second	VAR 023004 OF upload a text file:	97 58 C W VAR 023004 Or upload a text file:
the second						The second	VAR_023004 OF upload a text file:	97 58 C W VAR_023004 Or upload a text file:
						and an and a state that	V vo v vax_v23004 Or upload a text file:	97 58 C W VAR_023004 OF upload a text file:
						and an and a state that	V vo v vax_v23004 Or upload a text file:	97 58 C W VAR_023004 OF upload a text file:
						and a second a second a second s	VAR_023004 OF upload a text file:	97 56 C W VAR_023004 OF upload a text file:
						and a second a second a second s	V se c w VAR_023004 Or upload a text file:	97 58 C W VAR_023004 OF upload a text file:
						and an and a state that	V vo v vax_v23004 Or upload a text file:	97 58 C W VAR_023004 OF upload a text file:
						and an and a state that	V vo v vax_v23004 Or upload a text file:	97 88 C W VAR_023004 Or upload a text file:
the second						The second	VAR_023004 OF upload a text file:	97 58 C W VAR_023004 Or upload a text file:
						The second	VAR 023004 OF upload a text file:	97 58 C W VAR 023004 Or upload a text file:
						The second	VAR 023004 OF upload a text file:	97 58 C W VAR 023004 Or upload a text file:
the second						The second	VAR_023004 OF upload a text file:	97 58 C W VAR_023004 Or upload a text file:
the second						The second	VAR_023004 OF upload a text file:	97 58 C W VAR_023004 Or upload a text file:
the second						The second	VAR_023004 OF upload a text file:	97 58 C W VAR_023004 Or upload a text file:
						The second	VAR 023004 OF upload a text file:	97 58 C W VAR 023004 Or upload a text file:
						The second	VAR 023004 OF upload a text file:	97 58 C W VAR 023004 Or upload a text file:
the second						The second	VAR_023004 OF upload a text file:	97 58 C W VAR_023004 Or upload a text file:
						The second	VAR 023004 OF upload a text file:	97 58 C W VAR 023004 Or upload a text file:
						The second	VAR 023004 OF upload a text file:	97 58 C W VAR 023004 Or upload a text file:
the second						The second	VAR_023004 OF upload a text file:	97 58 C W VAR_023004 Or upload a text file:
the second						The second	VAR_023004 OF upload a text file:	97 56 C W VAR_023004 OF upload a text file:
						and an and a state that	VAR_023004 Or upload a text file:	97 56 C W VAR_023004 OF upload a text file:
						The second	or upload a text me.	The very of upload a text file.
2000 160 W B USB 015006	20 160 W B USB 013005	20 140 H B HBB 013005	140 11 11 1111 012005	160 H H H H H H H H H H H H H H H H H H H	140 11 11 1100.0100.00	300 160 H H H H H H H A 19005		
200 160 P P Vap 012905	160 B B VIAD 012905	29 140 P P USP 012905	140 B B USD 012005	NT0 160 P P VID 012805	140 H P Vap 012905	200 160 H B THE A13005		
200 160 P P Vap 012905	160 B B VIAD 012905	29 140 P P USP 012905	140 B B USD 012005	NT0 160 P P VID 012805	140 H P Vap 012905	200 160 H B THE A13005		
000 160 8 9 930 012005	419 160 H H VAR 012805	49 160 H B VAR 012805	160 H B VAR 012805	AC9 160 H B VAR 012805	9 160 H R VAR 012805	200 140 H H H HAD 012005		
000 160 8 9 930 012005	419 160 H H VAR 012805	49 160 H B VAR 012805	160 H B VAR 012805	AC9 160 H B VAR 012805	9 160 H R VAR 012805	200 140 H H H HAD 012005		
000 160 8 9 930 012005	419 160 H H VAR 012805	49 160 H B VAR 012805	160 H B VAR 012805	AC9 160 H B VAR 012805	9 160 H R VAR 012805	200 140 H H H HAD 012005		
000 160 8 9 930 012005	419 160 H H VAR 012805	49 160 H B VAR 012805	160 H B VAR 012805	AC9 160 H B VAR 012805	9 160 H R VAR 012805	200 140 H H H HAD 012005		
200 140 B B V12 012005	419 160 H B VAR 012805	49 160 H B VAR 012805	160 H B VAR 012805	AG9 160 H B VAR 012805	160 H B VAR 012805	3/0 1/0 H B HBB 013005		
2000 160 W B USB 015006	20 160 W B USB 013005	20 140 H B HBB 013005	140 11 11 1111 012005	160 H H H H H H H H H H H H H H H H H H H	140 11 11 1100.0100.00	300 160 H H H H H H H A 19005		
174 174 H H H H H H H H H H H H H H H H H H H	100 100 m m m 010000	140 H H H H H H H H H H H H H H H H H H H	140	100 100 m m m 10000	10.100 00 00 00000	140 H H H H H H H H H H H H H H H H H H H	or uproud a case the	or uprove a case the
174 174 H H H H H H H H H H H H H H H H H H H	100 100 m m m 010000	140 H H H H H H H H H H H H H H H H H H H	140	100 100 m m m 10000	10.100 00 00 00000	200 120 H H H H H H H H H H H H H H H H H H H		
200 160 P P Vap 012905	160 B B VIAD 012905	29 140 P P USP 012905	140 B B USD 012005	NO 160 B B VID 012805	140 H P Vap 012905	200 160 H B THE A13005		
000 160 8 9 930 012005	419 160 H H VAR 012805	49 160 H B VAR 012805	160 H B VAR 012805	AC9 160 H B VAR 012805	9 160 H H VAR 012805	200 140 H H H HAD 012005		
0/0 160 12 12 1/2005	409 160 H B VAR 012805	49 160 H R VAR 012805	160 H B VAR 012805	409 160 H B VAR 012805	9 160 H R VAR 012805	APR 120 11 11 1110 A190AE		
3/19 160 B B UND 012905	419 160 H R VAR 012805	49 160 H R VAR 012805	160 H B VAR 012805	AG9 160 H R VAR 012805	9 160 H R VAR 012805	ACR 160 H B HAD 013005		
3/19 160 B B UND 012905	419 160 H R VAR 012805	49 160 H R VAR 012805	160 H B VAR 012805	AG9 160 H R VAR 012805	9 160 H R VAR 012805	ACR 160 H B HAD 013005		
3/19 160 B B UND 012905	419 160 H R VAR 012805	49 160 H R VAR 012805	160 H B VAR 012805	AG9 160 H R VAR 012805	9 160 H R VAR 012805	ACR 160 H B HAD 013005		
3/19 160 B B UND 012905	419 160 H R VAR 012805	49 160 H R VAR 012805	160 H B VAR 012805	AG9 160 H R VAR 012805	9 160 H R VAR 012805	ACR 160 H B HAD 013005		
3/19 160 B B UND 012905	419 160 H R VAR 012805	49 160 H R VAR 012805	160 H B VAR 012805	AG9 160 H R VAR 012805	9 160 H R VAR 012805	ACR 160 H B HAD 013005		
201 160 2 2 122 012205	A19 160 H R VAR 012805	49 160 H R VAR 012805	160 H H VAR 012805	AC9 160 H R VAR 012805	9 160 H R VAR 012805			
AND THE ME AND	AN IND M H VAN DIZED	AN END M MAR DIZED	IND M H VAN DIZKUN	ANY THU M N VAN ULZBUS	9 160 H H VAR 012803			10 100 m m m mm 010000
					1 1 N XAR V140V3		THE 160 THE DISCON	TA 140 11 11 1110 015005
					1 1 N XAR V140V3		THE 160 THE DISCON	TA 140 11 11 1110 015005
201 160 2 2 122 012205	A19 160 H R VAR 012805	49 160 H R VAR 012805	160 H H VAR 012805	AC9 160 H R VAR 012805	9 160 H R VAR 012805			
0/0 160 8 9 1100 012005	419 160 H R VAR 012805	19 160 H B VAR 012805	160 H B VAR 012805	409 160 H B VAR 012805	9 160 H R VAR 012805	ACO 140 11 11 1120 012005		
2010 140 P P Van 012005	409 160 H B VAR 012805	49 160 H R VAR 012805	160 H B VAR 012805	NG9 160 H R VAR 012805	9 160 H B VAR 012805	200 120 H H H H H H H H A 19005	an aprove a construction	
2010 140 B B USD 012005	409 160 H B VAR 012805	49 160 H R VAR 012805	160 H B VAR 012805	NG9 160 H R VAR 012805	160 H B VAR 012805	300 140 H B H B H B H B B B B B B B B B B B B		
ACG 160 B B V3D 012005	419 160 H B VAR 012805	49 160 H B VAR 012805	160 H B VAR 012805	AG9 160 H B VAR 012805	9 160 H R VAR 012805	APR 160 H T THE 019005		
900 160 B B Vap 012005	419 160 H B VAR 012805	49 160 H B VAR 012805	160 H B VAR 012805	AG9 160 H B VAR 012805	9 160 H R VAR 012805	APR 160 H T TAT 019005		
200 140 8 9 1122 012005	419 160 H H VAR 012805	49 160 H B VAR 012805	160 H B VAR 012805	AC9 160 H B VAR 012805	9 160 H R VAR 012805	200 140 H H H HAD 012005		
200 140 B B V10 012005	419 160 H B VAR 012805	49 160 H B VAR 012805	160 H B VAR 012805	AG9 160 H B VAR 012805	160 H B VAR 012805	3/0 1/0 H B HBB 013005		

Fig. 1. Input page of PICMI for amino acid variations.

input data	chr	chr position	strand	wt base	mut base	wt	mut codon	ENSG	ENST	cDNA pos	ENSP	var id
P33897,88,C,W	х	152990985	1	C	G	TGC	TGG	101986	218104	264	218104	VAR_02300
Q9NRG9,160,H,R	12	53708601	-1	A	G	CAC	CGC	94914	209873	479	209873	VAR_01280
Q4L235,61,1,V	4	57250285	-1	A	G	ATT	GTT	157426	205214	181	205214	VAR_03831
Q4L235,61,1,V	4	57250285	-1	A	G	ATT	GTT	157426	514745	181	427298	VAR_03831
Q4L235,61,1,V	4	57250285	-1	A	G	ATT	GTT	157426	502617	181	421171	VAR_03831
Q4L235,61,I,V	4	57250285	-1	A	G	ATT	GTT	157426	451613	181	409656	VAR_03831
Q68DY9,182,C,W	19	57985566	-1	C	G	TGC	TGG	197128	319969	384	321015	VAR_02960
Q68DY9,182,C,W	19	57985566	-1	C	G	TGC	TGG	197128	343280	546	341165	VAR_02960
Q68DY9,182,C,W	19	57985566	+1	C	G	TGC	TGG	197128	427512	423	395967	VAR_02960
Q68DY9,182,1,M	19	\$7985566	-1	C	G	ATC	ATG	197128	450712	321	400754	VAR_02960
Q68DY9,182,1,M	19	57985566	-1	C	G	ATC	ATG	197128	356584	264	348992	VAR_02960
Q68DY9,182,1,M	19	57985566	-1	C	G	ATC	ATG	197128	415705	486	413487	VAR 02950

**Fig. 2.** Example of the output of PICMI for amino acid variations falling in a coding region.

The user can also input one or more amino acid mutations in a given protein when the information on the corresponding nucleotide mutation is not available, as is the case for those reported in the SwissProt 'Natural variant' field, in the SwissVar portal and in a number of entries in OMIM. Given the UniProt identifier of the protein, the position of the mutation and the wild-type and mutated amino acid in the protein sequence, the system will retrieve the coordinates of the corresponding gene in the genome, identify the wild-type codon and verify whether the mutated amino acid can be unambiguously obtained by a single-nucleotide mutation. If this is the case, the identified nucleotide variation is treated as in the case of an input nucleotide variation (Fig. 2).

The system relies on the Perl APIs provided by Ensembl. For nucleotide variations, it first verifies whether the input data are consistent with the genome sequence and, next, it maps the identified position on all the genes/isoforms spanning it. For amino acid variations, after a consistency check, it aligns the UniProt sequence to the corresponding Ensembl gene products and proceeds as in the case of nucleotide variations. As an example of the usefulness of the amino acid variation option of the tool, entry 600509.0011 of the OMIM resource reports two mutations of the ABCC8 protein associated to hyperinsulinemic hypoglycemia, E1506K (Huopio *et al.*, 2000) and E1507K (Pinney *et al.*, 2008); however, the two mutations correspond to the same nucleotide variation, and the discrepancy in the numbering is due to the fact that they were originally mapped by the authors on different splicing isoforms of the protein.

The question obviously arises about how often an amino acid variation can be unambiguously assigned to a single nucleotide polymorphism. We tested the PICMI server on the whole collection of polymorphisms in the SwissVar knowledgebase that provides information on about 53 000 amino acid variations (release 56.8). (Results are available at www.biocomputing.it/picmi/SwissVar). Interestingly, >85% of the amino acid variations could be unambiguously associated to single nucleotide mutations and therefore mapped on all alternative isoforms of the corresponding analyzed genes.

#### **3 CONCLUSIONS**

We believe that this easy-to-use tool can reveal to be very useful both to simplify the mapping of nucleotide variations and, especially, to analyze a number of pathological and physiological variations at the nucleotide level when they are only reported at the protein level. In this way, the server can add value to existing amino acid variation data. We will continuously update it by adding more genomes, as soon as sufficient mutation data will accumulate. We also plan to allow mapping of insertions and deletions in the next release and to make the tool available as a web service.

### ACKNOWLEDGEMENTS

We are grateful to the Biocomputing group for useful discussions.

*Funding*: KAUST (Award N. KUK-I1-012-43) and FIRB (Italbionet and Proteomica).

Conflict of Interest: none declared.

### REFERENCES

- Amberger, J. et al. (2009) McKusick's online mendelian inheritance in man (OMIM) Nucleic Acids Res., 37, D793–D796.
- Hubbard, T.J. et al. (2009) Ensembl 2009. Nucleic Acids Res., 37, D690-D697.
- Huopio, H. et al. (2000) Dominantly inherited hyperinsulinism caused by a mutation in the sulfonylurea receptor type 1. J. Clin. Invest., 106, 897–906.
- Kumar, P. et al. (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat. Protocols, 4, 1073–1081.
- McLaren, W. et al. (2010) Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor, *Bioinformatics*, 26, 2069–2070.
- Mottaz, A. et al. (2010) Easy retrieval of single amino-acid polymorphisms and phenotype information using SwissVar. Bioinformatics, 26, 851–852.
- Pinney,S.E. et al. (2008) Clinical characteristics and biochemical mechanisms of congenital hyperinsulinism associated with dominant KATP channel mutations. J. Clin. Invest., 118, 2877–2886.
- Sherry,S.T. et al. (2001) dbSNP: the NCBI database of genetic variation. Nucleic Acids Res., 29, 308–311.
- The UniProt Consortium (2010) The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Res.*, **38**, D142–D148.