

RELAZIONE DI BIOLOGIA MOLECOLARE 2

BRUNO FRANCESCA mat.576193

Analisi di proteine. Descrivi un database di interazioni proteiche e mostra con quali proteine interagisce la proteina KEN di Drosophila.

Uno dei database utilizzati durante una ricerca di interazioni proteiche è String. String ospita la banca dati di interazioni proteina-proteina stimate e quelle sperimentali. Le Reti di interazione proteina-proteina sono un ingrediente importante per la comprensione di molti processi cellulari. Tali reti possono essere utilizzate per il filtraggio e la valutazione dei dati della genomica funzionale e per fornire una piattaforma intuitiva per l'annotazione di proprietà strutturali, funzionali ed evolutive delle proteine. Esplorando le reti di interazione previste si possono suggerire nuove direzioni per il futuro della ricerca sperimentale e fornire previsioni cross-specie per la mappatura.

Come molti altri database che memorizzano l'associazione tra proteine STRING importa dati derivati da interazioni proteina-proteina attraverso la letteratura. Inoltre, STRING memorizza anche interazioni computazionalmente predette da: text mining di testi scientifici, le interazioni calcolate dalle caratteristiche genomiche, e le interazioni trasferite da organismi modello basate sull'ortologia. Tutte le interazioni previste o importate, sono confrontate con un riferimento comune di partnership funzionale commentato fatto da KEGG (Kyoto Enciclopedia dei geni e dei genomi). String viene aggiornato continuamente, e contiene attualmente 261 033 ortologi in 89 genomi completamente sequenziati. Il database prevede interazioni funzionali ad un livello atteso di precisione di almeno l'80% per più della metà dei geni. Collegamenti funzionali tra proteine spesso possono essere dedotti dalle associazioni genomiche tra i geni che li codificano: gruppi di geni che sono necessari per la stessa funzione tendono a mostrare la copertura specie simile, sono spesso situate in prossimità del genoma (nei procarioti), e tendono ad essere coinvolti in eventi gene-fusion. STRING è una risorsa pre-calcolata globale per l'esplorazione e l'analisi di queste associazioni. Dal momento che i tre tipi di prove differiscono concettualmente, e il numero di interazioni previste è molto grande, è essenziale essere in grado di valutare e confrontare il significato delle singole previsioni. Così, string contiene un unico score-frameshift basato su parametri di riferimento dei diversi tipi di associazioni nei confronti di un set comune di riferimento, integrati in un unico punteggio di confidenza per la previsione. La rappresentazione grafica della rete di interazioni fornisce una visione di alto livello del collegamento funzionale, facilitando l'analisi della modularità nei processi biologici.

String presenta una pagina iniziale nella quale è presente un box dove è possibile inserire la nostra sequenza query o il nome della proteina.(Fig.1)

L' output che si presenta è una pagina con una mappa colorata, una rete di interazione nella quale è rappresentata la nostra proteina query come una sfera collegata, tramite dei fili di colori diversi, ad altre sfere, ovvero le altre proteine con le quali questa interagisce.(Fig.2) Nella parte sottostante la rappresentazione grafica è presente un riquadro grazie al quale si possono scegliere le opzioni di visualizzazione. Ad esempio tramite i tasti "+ more/- less" si possono inserire o nascondere ulteriori interazioni cioè avere la nostra proteina rappresentata solo con le interazioni più importanti o anche quelle più lontane. Si può scegliere di salvare su disco il nostro lavoro tramite "Save" o scegliendo l' opzione "interactive" si possono spostare a nostro piacimento i nodi della rete in modo da facilitare la visualizzazione dei rapporti. Si può inoltre vedere sotto la rappresentazione grafica un riquadro che ci indica i pattern funzionali predetti elencati in ordine di maggiore probabilità.(Fig.3)

Per iniziare la nostra ricerca sui pattern funzionali della proteina Ken di Drosophila ricerchiamo innanzitutto la nostra proteina sul database "Uniprot" dal quale possiamo ricavare diverse informazioni riguardanti la proteina stessa come la lunghezza della sequenza, la sua funzione, la localizzazione a livello cellulare.(Fig.4) Nella parte finale della pagina di Uniprot sono presenti le referenze di KEN nei database, da qui scegliamo la referenza del database String e ci clicchiamo su in modo da avere il collegamento su String.(Fig.5)

Una volta entrati tramite il collegamento di Uniprot sul database di interazioni proteiche si visualizza direttamente la nostra proteina query con le proteine con le quali interagisce.(Fig.6).

In particolare KEN interagisce con 10 pattern funzionali: fzy, Cdc27, BubR1, mad2, rap, mr, aur, CG4882, CG4585, wibg con diverse probabilità. La proteina con la quale KEN interagisce con maggiore probabilità è fzy con uno score dello 0,817.

Il grafo che rappresenta KEN insieme alle altre proteine interagenti si può modificare visualizzando solo le proteine più probabili tramite il tasto "- less"(Fig.7).

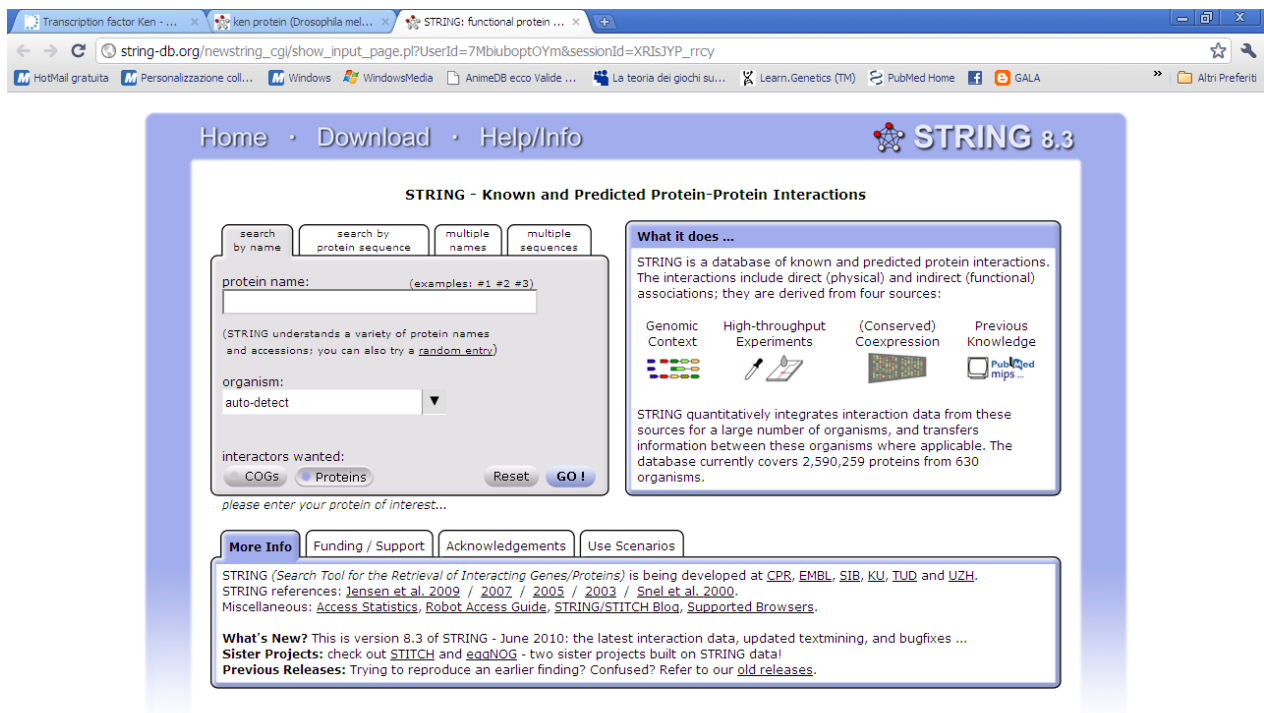


Figura 1

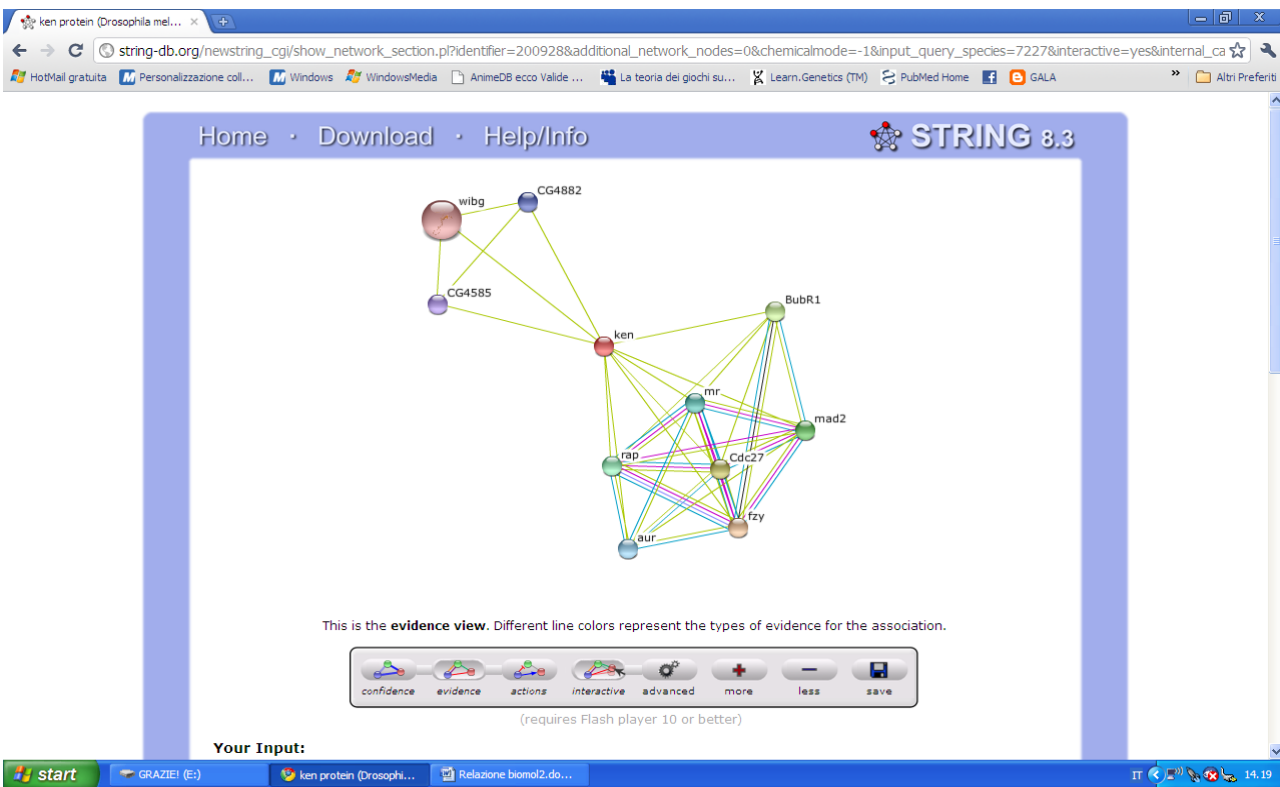


Figura 2

ken protein (Drosophila mel... x)

string-db.org/newstring.cgi/show_network_section.pl?identifier=200928&additional_network_nodes=0&chemicalmode=-1&input_query_species=7227&interactive=yes&internal_ca

HotMail gratuita Personalizzazione coll... Windows WindowsMedia AnimeDB ecco Valde ... La teoria dei giochi su... Learn.Genetics (TM) PubMed Home GALA Altri Preferiti

Home · Download · Help/Info **STRING 8.3**

This is the **evidence view**. Different line colors represent the types of evidence for the association.

confidence evidence actions interactive advanced more less save

(requires Flash player 10 or better)

Your Input:

- ken Probable transcription factor Ken (Protein Ken and Barbie); Transcription factor required for terminalia development. Negative regulator of the JAK/STAT pathway- represses JAK/STAT-dependent expression of ventral veins lacking (vvl) in the posterior spiracles (601 aa) (*Drosophila melanogaster*)

Predicted Functional Partners:

	Neighborhood	Gen: Fusion	Coincidence	Co-expression	Co-localization	Path: bases	Textmining	[Homology]	Score
fzy	fizzy CG4274-PA (526 aa)	•	•	•	•	•	•	•	0.817
Cdc27	Cdc27 CG8610-PA (900 aa)	•	•	•	•	•	•	•	0.754
BubR1	Bub1 CG7838-PA (1460 aa)	•	•	•	•	•	•	•	0.744
mad2	mad2 CG17498-PA (207 aa)	•	•	•	•	•	•	•	0.736
rap	retina aberrant in pattern CG3000-PB, isoform B (478 aa)	•	•	•	•	•	•	•	0.710

start GRAZIE! (E:) ken protein (Drosophi... Relazione biomi2.do... IT 14.15

Figura 3

Transcription factor Ken - ... ken protein (Drosophila mel... STRING: functional protein ...

www.uniprot.org/uniprot/O77459

HotMail gratuita Personalizzazione coll... Windows WindowsMedia AnimeDB ecco Valde ... La teoria dei giochi su... Learn.Genetics (TM) PubMed Home GALA Altri Preferiti

UniProtKB Downloads Contact Documentation/Help

Search Blast * Align Retrieve ID Mapping *

Search in Query Search Clear Advanced Search »

O77459 (KEN_DROME) ★ Reviewed, UniProtKB/Swiss-Prot
 Last modified November 30, 2010. Version 86. History...

Contribute
 Send feedback
 Read comments (0) or add your own

Clusters with 100%, 90%, 50% identity | Documents (2) | Third-party data
 text xml rdf/xml gif fasta

Names - Attributes - General annotation - Ontologies - Interactions - Sequence annotation - Sequences - References - Cross-refs - Entry info - Documents - Customize order

Names and origin

Protein names	Recommended name: Transcription factor Ken Alternative name(s): Protein Ken and Barbie
Gene names	Name: ken ORF Names:CG5575
Organism	Drosophila melanogaster (Fruit fly) [Complete proteome]
Taxonomic identifier	7227 [NCBI]
Taxonomic lineage	Eukaryota > Metazoa > Arthropoda > Hexapoda > Insecta > Pterygota > Neoptera > Endopterygota > Diptera > Brachycera > Muscomorpha > Ephydroidea > Drosophilidae > Drosophila > Sophophora

Protein attributes

Sequence length	601 AA.
Sequence status	Complete.
Protein existence	Evidence at protein level.

General annotation (Comments)

start Transcription factor K... Analisi.doc [Modalità ...] 16.05

Figura 4

Transcription factor Ken - ... ken protein (Drosophila mel... STRING: functional protein ...

www.uniprot.org/uniprot/O77459

HotMail gratuita Personalizzazione coll... Windows WindowsMedia AnimeDB ecco Valde ... La teoria dei giochi su... Learn.Genetics (TM) PubMed Home GALA Altri Preferiti

Names - Attributes - General annotation - Ontologies - Interactions - Sequence annotation - Sequences - References - Cross-refs - Entry info - Documents - Customize order

Protein-protein interaction databases

DIP	DIP-22018N.
IntAct	O77459. 2 interactions.
MINT	MINT-981522.
STRING	O77459.

Proteomic databases

PRIDE	O77459.
-------	---------

Genome annotation databases

EnsemblMetazoa	FBtr0072129; FBpp0072038; FBgn0011236.
GeneID	37785.
KEGG	dme:Dmel_CG5575.
NMPDR	figI7227.3 pag. 7156.

Organism-specific databases

CTD	37785.
FlyBase	FBgn0011236. ken.

Phylogenomic databases

eggNOG	inNOG08862.
InParanoid	O77459.
OMA	REYRCEY.
OrthoDB	EOG9RBQBG.
PhylomeDB	O77459.

Gene expression databases

ArrayExpress	O77459.
--------------	---------

start Transcription factor K... Analisi.doc [Modalità ...] 16.07

Figura 5

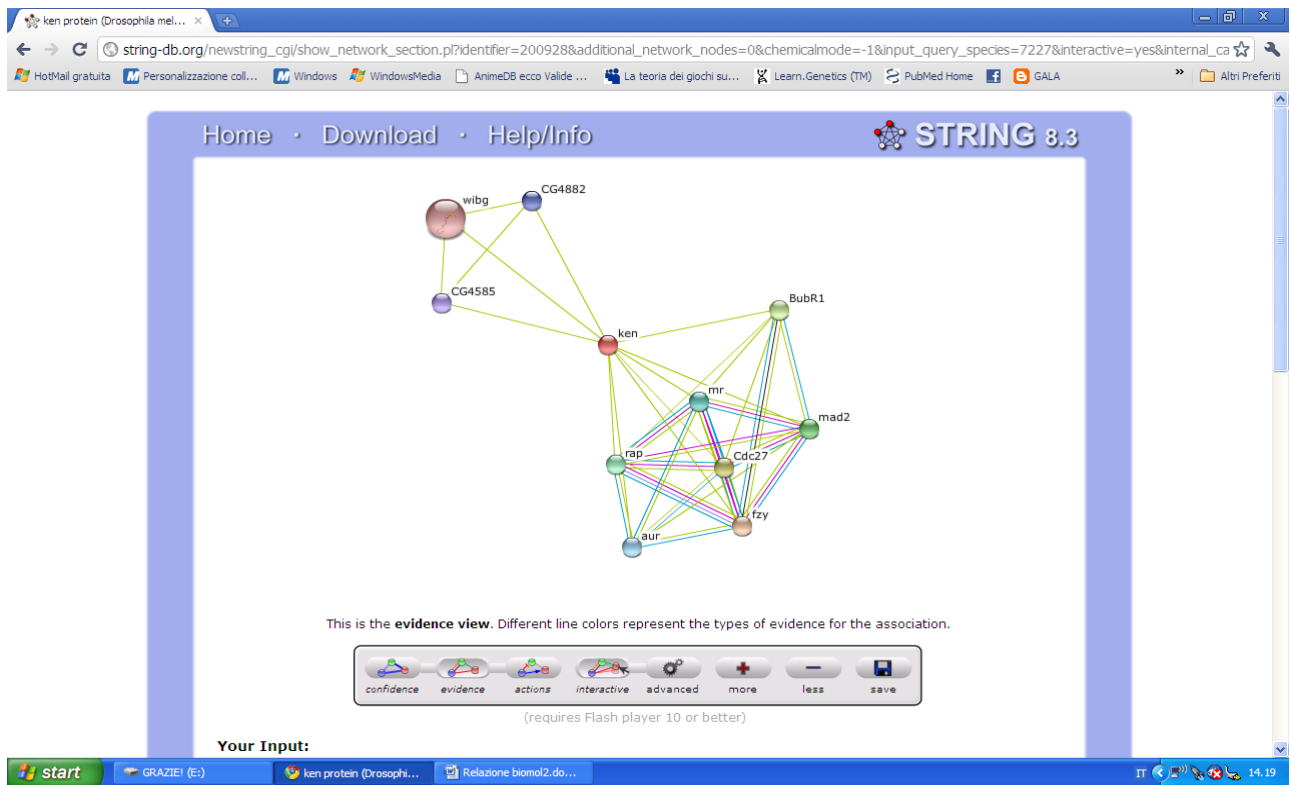


Figura 6

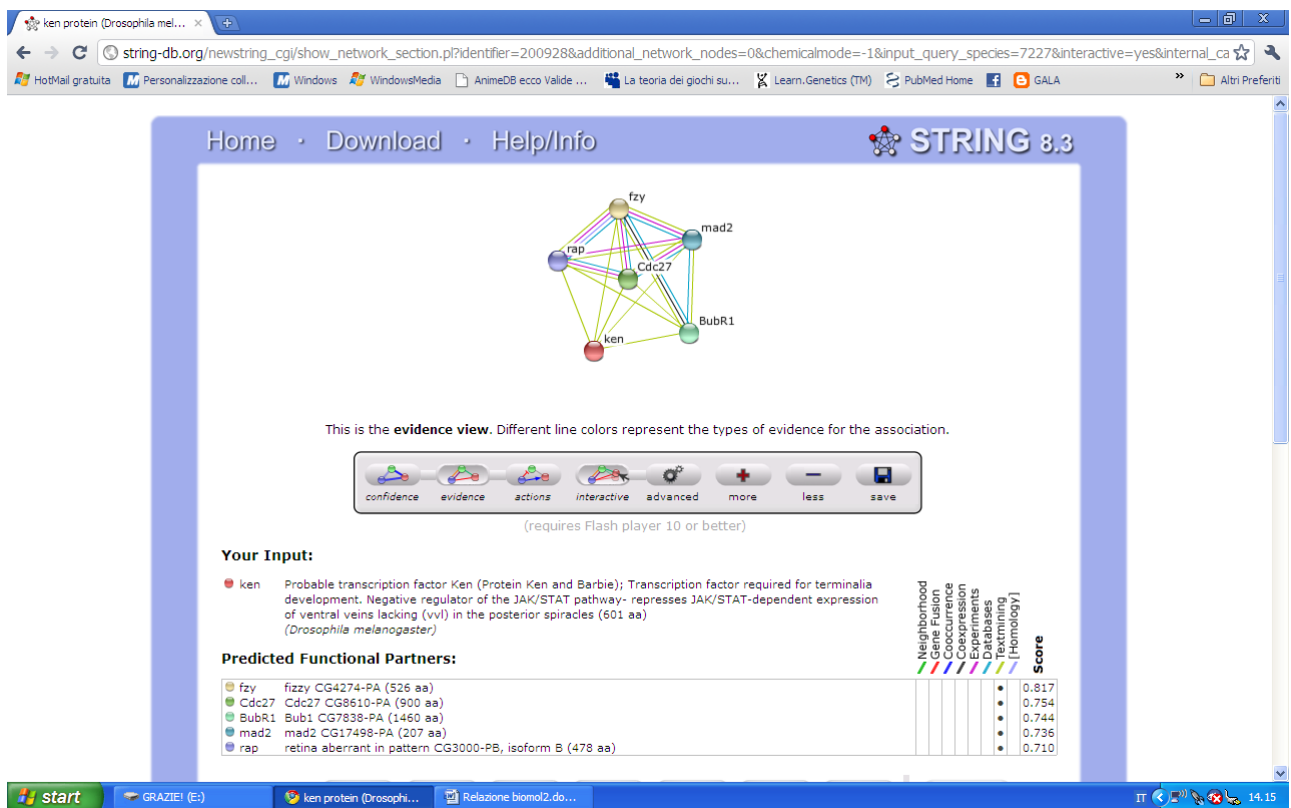


Figura 7