



DEPARTAMENT DE BIOQUÍMICA
I BIOLOGIA MOLECULAR

UNIVERSITAT DE BARCELONA

**ANÀLISI BIOINFORMÀTICA
DE LES MUTACIONS PUNTUALS
PATOLÒGIQUES**

CARLES FERRER COSTA
2005

DEPARTAMENT DE BIOQUÍMICA I BIOLOGIA
MOLECULAR

UNIVERSITAT DE BARCELONA

**ANÀLISI BIOINFORMÀTICA DE LES
MUTACIONS PUNTUALS PATOLÒGIQUES**

CARLES FERRER COSTA

2005

Programa de Doctorat de Biomedicina
Departament de Bioquímica i Biologia Molecular
Universitat de Barcelona
Bienni 2000-2001

Aquesta tesi ha estat realitzada pel llicenciat Carles Ferrer Costa sota la direcció del Dr. Modesto Orozco Lòpez i el Dr. Xavier de La Cruz Montserrat per optar al títol de doctor per la Universitat de Barcelona en el Programa de Doctorat de Biomedicina.

Modesto Orozco Lòpez

Xavier de la Cruz Montserrat

També vull agrair a la Rusbi, per proximitat, tots els anys de vivència, primer a BCN i després a Calella, amb tot el que ha estat tan bo i també les discussions. I a la resta de Ferrers, perquè posar-los un a un ocuparien massa, per fer-me gran i per fer-me sentir que som la família Ulisses ;))

Gràcies també a l'Alberto, per la teva amistat i tot el teu saber inacabable, i el teu cor tan gran.

Al Josep Lluís Gelpí, només li puc dir-li gràcies pel seu coneixement immens, sobre tot, i fer tan fàcil allò que semblava impossible.

Ara si que vull espavilar, no vull deixar-me a ningú, i si ho faig, disculpes, però les neurones ja no responen com voldria. Gràcies a tots els MMBs, la gent del grup, als que hi són i als que han marxat, per la seva companyia, el seu silenci i la seva ajuda i també les nostres diferències.

Gràcies als de Calella, als de Barcelona, als de Nova York i als de Cambridge, a tots els que s'han creuat en el meu camí i han deixat sens dubte un record inoblidable, i un regust agradable. Espero després d'aquest no viure d'acabar amb aquesta tesi poder recuperar, la vostra amistat i alegria.

Aquí s'acaba el punt i seguit, ara comença el camí cap al desconegut...

Carles Ferrer Costa

Calella, juny de 2005

You'll Never Walk Alone

When you walk through a storm
Hold your head up high,
And don't be afraid of the dark.
At the end of a storm,
There's a golden sky,
And the sweet silver song of a lark.

Walk on through the wind,
Walk on through the rain,
Though your dreams be tossed and
blown...
Walk on, walk on, with hope in your
heart,
And you'll never walk alone,
You'll never walk alone.

Walk on, walk on, with hope in your
heart,
And you'll never walk alone...
You'll never walk alone.

ÍNDEX

Índex	i
Capítol 1. Introducció	1
I. El context evolutiu de les mutacions	6
II. Mutacions puntuals/SNPs	8
III. Mutacions i patologia	10
IV. Bases moleculars de l'impacte funcional/ estructural de les mutacions	12
L'efecte estructural de les mutacions puntuals	13
L'efecte de l'entorn cel·lular sobre l'efecte de les mutacions puntuals	15
V. Mutacions patològiques	17
Anèmia falciforme	17
Agammaglobulinèmia lligada al X	23
Insulina	29
VI. Mutacions patològiques puntuals: descripció	32
VII. Mutacions patològiques puntuals: predicció	39
Bibliografia del capítol.....	49
Capítol 2. Objectius	57
Capítol 3. Caracterització estructural i evolutiva de mutacions puntuals patològiques.	61
I. Introducció.....	63
II. Materials i mètodes	64
Polimorfismes d'un sol aminoàcid associats a malaltia	64
Variants neutres	65
Les estructures de les proteïnes	66

Càlculs d'estructura secundària i accessibilitat	66
Alineaments Múltiples de Seqüències	67
Mesures de la variabilitat en les posicions els alineaments múltiples	67
Mesures relacionades amb les propietats de seqüència i de residu	68
Tests Estadístics	69
Procediment de Predicció usant matrius de mutació	69
III. Article “Characterization of disease-associated single amino acid polymorphisms in terms of sequence and structure properties”	71
IV. Bibliografia del capítol	89
Capítol 4. Predicció de mutacions patològiques usant xarxes neurals	93
I. Introducció	95
II. Materials i mètodes	96
El protocol de predicció	96
Mutacions associades a malaltia	97
Mutacions neutres	98
Selecció d'estructures per les proteïnes	99
Propietats derivades de la seqüència i de l'estructura	99
Propietats derivades de l'estructura	100
Propietats de seqüència i de residu	102
Paràmetres evolutius	103
Informació de la base de dades	105
La xarxa neural	106
La validació creuada	106
Mesures del rendiment	107
Índex de fiabilitat	108

III. Article “Sequence-based prediction of pathological mutations.”	109
IV. Bibliografia del capítol.....	120
Capítol 5. Pmut: Un servidor d’internet per la predicció de mutacions puntuals	123
I.I ntroducció	125
II. Materials i mètodes	126
Estructuració del servidor	126
Predictor Pmut	126
La base de dades pmut	128
Implementació del programari i el seu ús	129
III. Article “PMUT: a web-based tool for the annotation of pathological mutations on proteins.”	133
IV. Article “PUPASVIEW: A visual tool for selecting suitable SNPs, with putative pathological effect in genes, for genotyping purpose.”	139
V. Bibliografia	146
Capítol 6. Ús d’eines bioinformàtiques per l’anotació de mutacions puntuals en models animals	149
I. Introducció	151
II. Materials i mètodes	152
Descriptors de les mutacions	153
Els grups de dades de mutacions	153
Les xarxes neurals	155
Mesures de rendiment	155
III. Article “Use of bioinformatics tools for the annotation of disease-associated mutations in animal models.”	157
IV. Bibliografia del capítol	202

Capítol 7. Variació en el caràcter patològic de les mutacions puntuals dependents de l'entorn estructural	207
I. Introducció	209
II. Materials i mètodes	212
Mutacions patològiques neutres i CPDs	212
Descriptors de les mutacions	214
Accessibilitat al solvent	214
Propietats de residu/seqüència	214
Entorn espacial del residu	214
Distribució de probabilitat d'identitat de seqüència local	215
III. Article “How can non-human proteins accommodate as wild-type human disease-associated residues.”	217
IV. Bibliografia del capítol	247
Capítol 8. Resum	249
Bibliografia del capítol	264
Capítol 9. Conclusions	267
Capítol 10. Altres treballs	271
Article “Functional and structural conservation of CBS domains from CLC chloride channels”	275
Apèndix	293
Dna, gens, proteïnes i mutacions	295
Gens	295
Proteïnes	298