

# TDA Mutación y TDA Enfermedad

## V1

Generado por Doxygen 1.8.11

## Índice

<b>1 Documentación Práctica</b>	<b>2</b>
1.1 Introducción . . . . .	2
1.1.1 Contexto . . . . .	3
1.1.2 Conjunto de Datos . . . . .	4
1.2 TDA enfermedad . . . . .	5
1.3 Mutación . . . . .	5
1.4 "Se Entrega / Se Pide" . . . . .	6
1.4.1 Se entrega . . . . .	6
1.4.2 Se Pide . . . . .	7
1.5 "Fecha Límite de Entrega" . . . . .	7
<b>2 Índice de clases</b>	<b>7</b>
2.1 Lista de clases . . . . .	7
<b>3 Índice de archivos</b>	<b>8</b>
3.1 Lista de archivos . . . . .	8
<b>4 Documentación de las clases</b>	<b>8</b>
4.1 Referencia de la Clase enfermedad . . . . .	8
4.1.1 Descripción detallada . . . . .	9
4.1.2 Documentación del constructor y destructor . . . . .	9
4.1.3 Documentación de las funciones miembro . . . . .	9
4.1.4 Documentación de los datos miembro . . . . .	12
4.2 Referencia de la Clase mutacion . . . . .	12
4.2.1 Descripción detallada . . . . .	14
4.2.2 Documentación del constructor y destructor . . . . .	14
4.2.3 Documentación de las funciones miembro . . . . .	15
4.2.4 Documentación de los datos miembro . . . . .	18

<b>5 Documentación de archivos</b>	<b>19</b>
5.1 Referencia del Archivo documentacion.dox . . . . .	19
5.2 Referencia del Archivo include/enfermedad.h . . . . .	19
5.2.1 Documentación de las funciones . . . . .	20
5.3 Referencia del Archivo include/enfermedad.hxx . . . . .	20
5.3.1 Documentación de las funciones . . . . .	20
5.4 Referencia del Archivo include/mutacion.h . . . . .	21
5.4.1 Documentación de las funciones . . . . .	21
5.5 Referencia del Archivo include/mutacion.hxx . . . . .	21
5.5.1 Documentación de las funciones . . . . .	21
5.6 Referencia del Archivo src/principal.cpp . . . . .	22
5.6.1 Documentación de las funciones . . . . .	22
<b>Índice</b>	<b>25</b>

## 1. Documentación Práctica

### Versión

v1

### Autor

Carlos Cano y Juan F. Huete

### 1.1. Introducción

En esta practica se pretende avanzar en el uso de las estructuras de datos mediante el diseño de distintos tipos de datos para manejar la información asociada a una base de datos de mutaciones del genoma humano con relevancia clínica (ClinVar-dbsnp).

### 1.1.1. Contexto

El ácido desoxirribonucleico, abreviado como ADN, es un ácido nucleico que contiene las instrucciones genéticas usadas en el desarrollo y funcionamiento de todos los organismos vivos conocidos y algunos virus, y es responsable de su transmisión hereditaria. En ocasiones, se compara al ADN con un programa de ordenador, ya que contiene las instrucciones necesarias para construir otros componentes de las células, como las proteínas y las moléculas de ARN, que son las responsables del funcionamiento celular. Los segmentos de ADN que llevan esta información genética son llamados genes.

Podemos representar el ADN como una secuencia de nucleótidos (Adenina A, Timina T, Citosina C, Guanina G). La disposición secuencial de estas cuatro bases a lo largo de la cadena es la que codifica la información genética. Por ejemplo, podemos representar una pequeña cadena de ADN como: "ACCCAGTCGGATTT".

En los organismos vivos, el ADN no suele existir como una molécula individual, sino como una pareja de moléculas que se enroscan sobre sí mismas formando una especie de escalera de caracol, denominada doble hélice. Esta estructura se sustenta en la complementariedad de sus bases (Citosina-Guanina y Adenina-Timina). Al ser las bases complementarias, podemos representar el ADN sin perder información especificando sólo una de sus cadenas.

El genoma humano es una secuencia de ADN contenida en 23 pares de cromosomas en el núcleo de cada célula humana (de cada pareja de cromosomas, uno es heredado del padre y otro de la madre). Los cromosomas 1 a 22 se numeran en orden creciente de tamaño. La pareja de cromosomas 23, también llamados cromosomas sexuales, se compone de un cromosoma X (de la madre) y uno X o Y (del padre).

El tamaño total del genoma humano haploide (es decir, considerando sólo uno de cada pareja de cromosomas) es de aproximadamente 3200 millones de pares de bases de ADN. Dado que una base se representa con un Byte ('A', 'C', 'G', 'T'), el tamaño aproximado de la secuencia completa de un genoma humano haploide es de 3 GBytes.

Dos seres humanos del mismo sexo comparten un porcentaje muy elevado (99,5 %) de su secuencia de ADN, pero estas secuencias no son idénticas. Estos millones de pequeñas variaciones en el genoma, junto con la influencia de factores del medio, son los responsables de que exhibamos distintos fenotipos, es decir, distintos rasgos físicos y conductales. Una variación en el genoma, por sustitución, inserción o delección de bases, se llama mutación o polimorfismo, y la principal fuente de variabilidad entre dos genomas humanos es el polimorfismo de una sola base (Single Nucleotide Polimorphism, SNP).

Un SNP es, por tanto, un cambio de una base en una misma posición entre dos genomas humanos. Un SNP suele representarse indicando el número de cromosoma en el que se localiza el cambio, la posición dentro del cromosoma, y el cambio de base respecto al genoma humano de referencia (el primer genoma humano para el que se conoce la secuencia, que se terminó de secuenciar por primera vez en 2001). Por ejemplo, el siguiente SNP indica un cambio en la posición 1014143 del cromosoma 1, que en el genoma humano de referencia presenta una 'C' y en otros genomas presenta una 'T':

1 1014143 C T

Los SNP constituyen hasta el 90 % de todas las variaciones genómicas humanas. Estas variaciones en la secuencia del ADN pueden afectar a la respuesta de los individuos a enfermedades, bacterias, virus, productos químicos, fármacos, etc.. De este modo, su estudio es de gran utilidad en la denominada Medicina Personalizada o Medicina de Precisión: el desarrollo de métodos de prevención, diagnóstico y tratamiento (fármacos) de forma individualizada para cada paciente.

Los estudios genéticos personalizados se basan en décadas de descubrimientos científicos publicados en la literatura especializada que muestran evidencia de que la presencia de un determinado SNP en el genoma de un individuo puede hacerle propenso a padecer una cierta enfermedad. La base de datos ClinVar-dbSNP recoge esta información.

Para leer más sobre el contexto del problema:

- [https://es.wikipedia.org/wiki/Ácido\\_desoxirribonucleico](https://es.wikipedia.org/wiki/Ácido_desoxirribonucleico)
- [https://es.wikipedia.org/wiki/Genoma\\_humano](https://es.wikipedia.org/wiki/Genoma_humano)
- [https://es.wikipedia.org/wiki/Polimorfismo\\_de\\_nucleótido\\_único](https://es.wikipedia.org/wiki/Polimorfismo_de_nucleótido_único)

### 1.1.2. Conjunto de Datos

El conjunto de datos con el que trabajaremos es la base de datos completa ClinVar-dbSNP descargada de la web del National Institute of Health (NIH) de los Estados Unidos: <https://www.ncbi.nlm.nih.gov/clinvar/>. Esta base de datos se puede obtener en formato VCF v4.0 (archivo: clinvar\_20160831.vcf), que representa de forma tabular más de 130.000 mutaciones (SNPs) conocidos hasta la fecha y su relación clínica con alguna enfermedad.

El fichero comienza con una cabecera (líneas que se inician con '#') que describe cada uno de los campos de la base de datos. A partir de la línea 67 se listan las entradas de la BD, con un SNP por línea, y los campos delimitados por tabulador ('\t'). Nota: algunos campos no relevantes se han omitido en este ejemplo para facilitar su lectura (los campos omitidos se han reemplazado por [...]).

```
#CHROM POS ID REF ALT QUAL FILTER INFO
1 1014143 rs786201005 C T . . RS=786201005; [...] GENEINFO=ISG15:9636; CLNSIG=5; CLNDSDB=MedGen:OMIM;
CLNDSDBID=CN221808:616126; CLNDBN=Immunodeficiency_38_with_basal_ganglia_calcification; [...]
1 1014316 rs672601345 C CG . . RS=672601345; [...] GENEINFO=ISG15:9636; CLNSIG=5; CLNDSDB=MedGen:OMIM;
CLNDSDBID=CN221808:616126; CLNDBN=Immunodeficiency_38_with_basal_ganglia_calcification; [...]
1 1053827 rs74685771 G A,C,T . . RS=74685771; [...] GENEINFO=AGRN:375790; [...] CLNSIG=3; CLNDSDB=MedGen;
CLNDSDBID=CN169374; CLNDBN=not_specified; [...]
1 11847114 rs202102042 C T . . RS=202102042; [...] GENEINFO=NPPA:4878|NPPA-AS1:100379251; [...] CLNSIG=5;
CLNDSDB=MedGen:OMIM; CLNDSDBID=C3810401:615745; CLNDBN=Atrial_standstill_2; [...] CAF=0.9998,0.0001997;COMMON
=0
1 11847311 rs755212754 G A . . RS=755212754; [...] GENEINFO=NPPA:4878|NPPA-AS1:100379251; [...] CLNSIG=3;
CLNDSDB=MedGen:OMIM; CLNDSDBID=C2677294:612201; CLNDBN=Atrial_fibrillation\2c_familial\2c_6; [...]
13 32316475 rs80359298 CAA C . . RS=80359298; [...] GENEINFO=BRCA2:675; [...] CLNSIG=1|5; CLNDSDB=MedGen:
OMIM;SNOMED_CT|MedGen:OMIM; CLNDSDBID=C0346153:114480:254843006|C2675520:612555; CLNDBN=
Familial_cancer_of_breast|Breast-ovarian_cancer\2c_familial_2; [...]
```

Los campos de interés en cada línea son los siguientes:

- CHROM: Número de cromosoma.
- POS: Posición del SNP dentro del cromosoma (comienza a numerarse en 1).
- ID: Identificador único del SNP ('rsXXXX').
- REF: Base(s) que aparecen en esa posición en el genoma humano de referencia. En caso de que aparezca una pequeña cadena de varias bases (ejemplo: "ATTGGAG"), el SNP que se indica reemplaza esta secuencia de bases por una sola.
- ALT: la(s) base(s) alternativa(s) que se han observado en la población. Si se han observado distintas mutaciones para la misma posición, éstas se indican delimitadas por coma (ejemplo: "A,C,T").
- INFO: Este campo representa información adicional sobre el SNP en forma de listado de atributos separados por ';'. Entre estos atributos, destacamos por su interés los siguientes:
  - GENEINFO: Nombre e identificador del gen que contiene este SNP. Ejemplo: GENEINFO=ISG15:9636 (Nombre del gen: ISG15, Identificador del gen: 9636). En caso de que se trate de varios genes, se separan con '|' o ';'. Ejemplo: GENEINFO=B3GALT6:126792|SDF4:51150
  - CAF: Frecuencia con que se observa cada base descrita en este SNP en la población. Ejemplo: CAF=0.9912,0.008786 indica que la base de la referencia se observa con frecuencia 0.9912 y la base alternativa con frecuencia 0.008786. El primer valor de CAF corresponde a frecuencia de la base REF, los siguientes a las bases indicadas en ALT, en el mismo orden.
  - COMMON: Indica si es un SNP común en la población (0 - no, 1 - si).
  - CLNSIG: relevancia clínica del SNP: 0/1 - Incierta, Desconocida, 2 - Benigno, 3 - Probablemente benigno, 4 - Probablemente patógeno, 5 - Patógeno, 6 - Relevante en respuesta a fármaco, 7 - Histocompatibilidad, 255 - Otro. En caso de que el SNP esté asociado con varias enfermedades se mostrará un código CLNSIG para cada enfermedad (delimitados por '|' o ';'), o un solo código CLNSIG, indicando que la relevancia clínica del SNP es la misma para todas las enfermedades.

- CLNDBN: Nombre de la enfermedad asociada al SNP. También se suministran el ID único de la enfermedad (CLNDSDBID) y la base de datos que provee este ID (CLNDSDB). En caso de que un SNP esté asociado a varias enfermedades, éstas se separan con '|' o ';'. El siguiente ejemplo hace referencia a tres enfermedades: CLNDSDB=MedGen|MedGen:OMIM|MedGen; CLNDSDBID=CN178850|C3809288:615373|CN169374; CLNDBN=Dilated\_cardiomyopathy\_1LL|Left\_ventricular\_noncompaction\_8|not\_specified;

## 1.2. TDA enfermedad

Para relacionar SNPs con enfermedades proponemos la creación de una clase enfermedad, que deberá tener entre otros los métodos abajo indicados. La especificación de la clase enfermedad se realizará en el fichero [enfermedad.h](#) y la implementación de la clase enfermedad en el fichero [enfermedad.hxx](#).

```
class enfermedad {
private:
    string  name;          // nombre de la enfermedad. Almacenar completo en minúscula.
    string  ID;            // ID único para la enfermedad
    string  database;      // Base de datos que provee el ID

public:
    enfermedad (); //Constructor de enfermedad por defecto
    enfermedad (const string & name, const string & ID, const string & database);

    void setName(const string & name);
    void setID(const string & ID);
    void setDatabase(const string & database);

    string getName();
    string getID();
    string getDatabase();

    enfermedad & operator=(const enfermedad & e);
    string toString() const;

    // Operadores relacionales
    bool operator==(const enfermedad & e) const;
    bool operator!=(const enfermedad & e) const;
    bool operator<(const enfermedad & e) const; //Orden alfabético por campo name.

    bool nameContains(const string & str) const; //Devuelve True si str está incluido en el
        nombre de la enfermedad, aunque no se trate del nombre completo. No debe ser sensible a mayúsculas/minúsculas.
}

ostream& operator<< (ostream& os, const enfermedad & e); //imprime enfermedad

#include "enfermedad.hxx" // Incluimos la implementacion.
```

Así, podremos trabajar con enfermedades como indica el siguiente código

```
...
enfermedad e1("Breast-ovarian_cancer\x2c_familial_2", "C2675520:612555", "MedGen:OMIM");
enfermedad e2("Prostate_cancer\x2c_susceptibility_to", "", "");
enfermedad e3 = e1;
...
if (e1.nameContains("cancer"))
    cout << e1 << " es un tipo de cancer. ";
...
```

## 1.3. Mutación

A igual que con la clase enfermedad, la especificación del tipo mutación y su implementación se realizará en los ficheros [mutacion.h](#) y [mutacion.hxx](#), respectivamente, y debe tener la información de los atributos (con su representación asociada)

- chr: identificador del cromosoma (string). Los cromosomas válidos son: "1", "2", "3", "4", "5", "6", "7", "8", "9", "10", "11", "12", "13", "14", "15", "16", "17", "18", "19", "20", "21", "22", "X", "Y", "MT".
- pos: identificador de la posición dentro del cromosoma (unsigned int).
- ID: identificador del SNP/mutación (string).
- ref\_alt: base(s) en el genoma de referencia y alternativa(s) posible(s) (vector de string). La primera posición la ocupará el string con la(s) base(s) del genoma de referencia, y, a continuación, aparecerán la(s) base(s) alternativas en el mismo orden que se indica en el fichero. Ejemplos:

```
X 154032548 rs61754422 C A,G,T
ref_alt: ["C", "A", "G", "T"]

1 1338032 rs797044840 GTAGGCAGG GC
ref_alt: ["GTAGGCAGG", "GC"]
```

- genes: gen(es) asociado(s) al SNP (vector de string). Ejemplo:

```
1 11847311 rs755212754 G A . . [...]GENEINFO=NPPA:4878|NPPA-AS1:100379251;[...]
genes: ["NPPA:4878", "NPPA-AS1:100379251"]
```

- common: indica si el SNP es común en la población (bool).
- caf: frecuencia de cada base del SNP en la población (vector de float). En primer lugar debe indicarse la frecuencia de la base 'ref' (posición 0 de ref-alt), seguida por las frecuencias de las bases alternativas indicadas en 'ref-alt', en el mismo orden. Ejemplo:

```
1 11847114 rs202102042 C T . . RS=202102042;[...];CAF=0.9998,0.0001997;COMMON=0
ref_alt: ["C", "T"]
caf: [0.9998, 0.0001997]
common: False
```

- enfermedades: enfermedades asociadas al SNP (vector de enfermedad).
- clnsig: relevancia clínica del SNP para cada enfermedad utilizando el código numérico del campo CLNSIG (vector de int). En caso de que existan varias enfermedades asociadas a la mutación, cada una de ellas puede presentar diferente código CLNSIG, por lo se deben almacenar en el vector clnsig en el mismo orden que las enfermedades asociadas. En caso de presentarse sólo un código CLNSIG y varias enfermedades, este código se aplica a todas ellas. Ejemplo:

```
13 32316475 rs80359298 CAA C . . RS=80359298;[...];CLNSIG=1|5;CLNDSDB=MedGen:OMIM:SNOMED_CT|MedGen:OMIM;
CLNDSDBID=C0346153:114480:254843006|C2675520:612555;CLNDBN=Familial_cancer_of_breast|Breast-
ovarian_cancer\x2c_familial_2;[...]

enfermedades: [ enfermedad("Familial_cancer_of_breast", "C0346153:114480:254843006", "
MedGen:OMIM:SNOMED_CT"),
                enfermedad("Breast-ovarian_cancer\x2c_familial_2", "C2675520:612555", "
MedGen:OMIM")]
clnsig: [1,5]
```

```
// Fichero mutacion.h
class mutacion {
    ....
}

#include "mutacion.hxx" // Incluimos la implementacion
```

## 1.4. "Se Entrega / Se Pide"

### 1.4.1. Se entrega

En esta práctica se entrega los fuentes necesarios para generar la documentación de este proyecto así como el código necesario para resolver este problema. En concreto los ficheros que se entregan son:

- `documentacion.pdf` Documentación de la práctica en pdf.
- `documentacion.dox` Este fichero contiene el fichero de configuración de doxygen necesario para generar la documentación del proyecto (html y pdf). Para ello, basta con ejecutar desde la línea de comando

```
doxygen doxPractica.txt
```

La documentación en html la podemos encontrar en el fichero `./html/index.html`. Para generar la documentación en latex es suficiente con hacer los siguientes pasos:

```
cd latex
make
```

obteniendo como resultado el fichero `refman.pdf` que incluye toda la documentación generada.

- `mutacion.h` Plantilla para la especificación del TDA mutación
- `mutacion.hxx` Plantilla para la implementación del TDA mutación
- `enfermedad.h` Plantilla para la especificación del TDA enfermedad
- `enfermedad.hxx` Plantilla para la implementación del TDA enfermedad
- `principal.cpp` Fichero donde se incluye el main del programa. Este programa toma como entrada el fichero de datos `"clinvar_20160831.vcf"`, carga las mutaciones en un vector, muestra el número total de mutaciones leídas del fichero y el número de mutaciones que están asociadas a una enfermedad que indica el usuario.

#### 1.4.2. Se Pide

- Diseñar la función de abstracción e invariante de la representación del tipo enfermedad.
- Diseñar la función de abstracción e invariante de la representación del tipo mutación.
- Implementar el código asociado a los ficheros `.hxx`.
- Implementar el código asociado a `principal.cpp`.

## 1.5. "Fecha Límite de Entrega"

La fecha límite de entrega será el 23 de Octubre a las 23:50 hrs.

## 2. Índice de clases

### 2.1. Lista de clases

Lista de las clases, estructuras, uniones e interfaces con una breve descripción:

<b>enfermedad</b>	
Clase enfermedad, asociada al TDA enfermedad	8
<b>mutacion</b>	
Clase mutacion, asociada a la definición de una mutación/SNP	12



### 3. Índice de archivos

#### 3.1. Lista de archivos

Lista de todos los archivos con descripciones breves:

<code>include/enfermedad.h</code>	19
<code>include/enfermedad.hxx</code>	20
<code>include/mutacion.h</code>	21
<code>include/mutacion.hxx</code>	21
<code>src/principal.cpp</code>	22

### 4. Documentación de las clases

#### 4.1. Referencia de la Clase enfermedad

Clase enfermedad, asociada al TDA enfermedad.

```
#include <enfermedad.h>
```

##### Métodos públicos

- `enfermedad ()`  
*Constructor de enfermedad por defecto.*
- `enfermedad (const string &name, const string &ID, const string &database)`  
*Constructor parametrizado de la clase.*
- `void setName (const string &name)`  
*Establece el nombre de la enfermedad.*
- `void setID (const string &ID)`  
*Establece el ID de la enfermedad.*
- `void setDatabase (const string &database)`  
*Establece la base de Datos de la enfermedad.*
- `string getName () const`  
*Obtiene el nombre de la enfermedad.*
- `string getID () const`  
*Obtiene el ID de la enfermedad.*
- `string getDatabase () const`  
*Obtiene la base de datos de la enfermedad.*
- `enfermedad & operator= (const enfermedad &e)`  
*operador de asignación*
- `string toString () const`  
*Muestra el valor de los atributos de la clase enfermedad.*
- `bool operator== (const enfermedad &e) const`  
*Compara si son iguales dos enfermedades.*
- `bool operator!= (const enfermedad &e) const`  
*Compara si son distintas dos enfermedades.*
- `bool operator< (const enfermedad &e) const`  
*Compara si una enfermedad es mayor que otra, comparando por orden alfabético por el campo name.*
- `bool nameContains (const string &str) const`  
*Comprueba si una cadena de caracteres está incluido en el nombre de la enfermedad.*

#### Atributos privados

- string `name`
- string `ID`
- string `database`

##### 4.1.1. Descripción detallada

Clase enfermedad, asociada al TDA enfermedad.

`enfermedad::enfermedad`, `enfermedad::setName`, `enfermedad::setID`, `enfermedad::setDatabase`, `enfermedad::getName`, `enfermedad::getID`, `enfermedad::getDatabase`, `enfermedad::operator=`, `enfermedad::toString`, `enfermedad::operator==`, `enfermedad::operator!=`, `enfermedad::operator<`, `enfermedad::nameContains`

Descripción: contiene toda la información asociada a una enfermedad almacenada en la BD ClinVar-dbSNP (nombre de la enfermedad, id, BD que provee el id)

#### Autor

Elena María Gómez Ríos

##### 4.1.2. Documentación del constructor y destructor

###### 4.1.2.1. `enfermedad::enfermedad ( )`

Constructor de enfermedad por defecto.

fichero de implementacion de la clase enfermedad

###### 4.1.2.2. `enfermedad::enfermedad ( const string & name, const string & ID, const string & database )`

Constructor parametrizado de la clase.

#### Parámetros

in	<i>name</i>	nombre de la enfermedad. Almacenar completo en minúscula.
in	<i>ID</i>	ID único para la enfermedad.
in	<i>database</i>	Base de datos que provee el ID.

##### 4.1.3. Documentación de las funciones miembro

###### 4.1.3.1. `string enfermedad::getDatabase ( ) const`

Obtiene la base de datos de la enfermedad.

#### Devuelve

base de datos que provee el ID de la enfermedad.

#### 4.1.3.2. `string enfermedad::getID ( ) const`

Obtiene el ID de la enfermedad.

Devuelve

ID de la enfermedad.

#### 4.1.3.3. `string enfermedad::getName ( ) const`

Obtiene el nombre de la enfermedad.

Devuelve

nombre de la enfermedad.

#### 4.1.3.4. `bool enfermedad::nameContains ( const string & str ) const`

Comprueba si una cadena de caracteres está incluido en el nombre de la enfermedad.

Parámetros

in	str	string a comprobar.
----	-----	---------------------

Devuelve

Devuelve true si el nombre de la enfermedad contiene a str.

#### 4.1.3.5. `bool enfermedad::operator!= ( const enfermedad & e ) const`

Compara si son distintas dos enfermedades.

Parámetros

in	e	enfermedad a comparar.
----	---	------------------------

Devuelve

Devuelve true si son distintas y false si no lo son.

#### 4.1.3.6. `bool enfermedad::operator< ( const enfermedad & e ) const`

Compara si una enfermedad es mayor que otra, comparando por orden alfabético por el campo name.

Parámetros

in	e	enfermedad a comparar.
----	---	------------------------

**Devuelve**

Devuelve true si e es mayor que el que lo llama. False en otro caso.

**4.1.3.7. enfermedad & enfermedad::operator= ( const enfermedad & e )**

operador de asignación

**Parámetros**

in	e	enfermedad a copiar.
----	---	----------------------

**Devuelve**

copia de la enfermedad.

**4.1.3.8. bool enfermedad::operator== ( const enfermedad & e ) const**

Compara si son iguales dos enfermedades.

**Parámetros**

in	e	enfermedad a comparar.
----	---	------------------------

**Devuelve**

Devuelve true si son iguales y false si no lo son.

**4.1.3.9. void enfermedad::setDatabase ( const string & database )**

Establece la base de Datos de la enfermedad.

**Parámetros**

in	database	Base de datos que provee el ID.
----	----------	---------------------------------

**4.1.3.10. void enfermedad::setID ( const string & ID )**

Establece el ID de la enfermedad.

**Parámetros**

in	ID	ID único para la enfermedad.
----	----	------------------------------

**4.1.3.11. void enfermedad::setName ( const string & name )**

Establece el nombre de la enfermedad.

**Parámetros**

<code>in</code>	<code>name</code>	nombre de la enfermedad. Almacenar completo en minúscula.
-----------------	-------------------	---

**4.1.3.12. `string enfermedad::toString ( ) const`**

Muestra el valor de los atributos de la clase enfermedad.

**Devuelve**

Devuelve un string con los atributos de la clase enfermedad.

**4.1.4. Documentación de los datos miembro****4.1.4.1. `string enfermedad::database [private]`**

Base de datos que provee el ID

**4.1.4.2. `string enfermedad::ID [private]`**

ID único para la enfermedad

**4.1.4.3. `string enfermedad::name [private]`**

Nombre de la enfermedad.

**Precondición**

Almacenar completo en minúscula.

La documentación para esta clase fue generada a partir de los siguientes ficheros:

- `include/enfermedad.h`
- `include/enfermedad.hxx`

**4.2. Referencia de la Clase mutacion**

Clase mutacion, asociada a la definición de una mutación/SNP.

```
#include <mutacion.h>
```

## Métodos públicos

- **mutacion** ()  
*Constructor de enfermedad por defecto.*
- **mutacion** (const **mutacion** &m)  
*Constructor de copia de la clase.*
- **mutacion** (const string &str)  
*Constructor parametrizado de la clase.*
- void **setID** (const string &id)  
*Establece el identificador del SNP/mutación.*
- void **setChr** (const string &chr)  
*Establece el identificador del cromosoma.*
- void **setPos** (const unsigned int &pos)  
*Establece el identificador de la posición dentro del cromosoma.*
- void **setRef\_alt** (const std::vector< string > &ref\_alt)  
*Establece la(s) base(s) en el genoma de referencia y alternativa(s) posible(s).*
- void **setGenes** (const std::vector< string > &genes)  
*Establece los gen(es) asociado(s) al SNP.*
- void **setCommon** (const bool &common)  
*Establece el indicador de si el SNP es común en la población.*
- void **setCaf** (const std::vector< float > &caf)  
*Establece la frecuencia de cada base del SNP en la población.*
- void **setEnfermedades** (const std::vector< enfermedad > &enfermedades)  
*Establece las enfermedades asociadas al SNP.*
- void **setClnsig** (const std::vector< int > &clnsig)  
*Establece la relevancia clínica del SNP para cada enfermedad.*
- string **getID** () const  
*Obtiene el identificador del SNP/mutación.*
- string **getChr** () const  
*Obtiene identificador del cromosoma.*
- unsigned int **getPos** () const  
*Obtiene el identificador de la posición dentro del cromosoma.*
- const vector< string > & **getRef\_alt** () const  
*Obtiene la(s) base(s) en el genoma de referencia y alternativa(s) posible(s).*
- const vector< string > & **getGenes** () const  
*Obtiene los gen(es) asociado(s) al SNP.*
- bool **getCommon** () const  
*Obtiene el indicador de si el SNP es común en la población.*
- const vector< float > & **getCaf** () const  
*Obtiene la frecuencia de cada base del SNP en la población.*
- const vector< enfermedad > & **getEnfermedades** () const  
*Obtiene las enfermedades asociadas al SNP.*
- const vector< int > & **getClnsig** () const  
*Obtiene la relevancia clínica del SNP para cada enfermedad.*
- **mutacion** & **operator=** (const **mutacion** &m)  
*operador de asignación*
- bool **operator==** (const **mutacion** &m) const  
*Compara si son iguales dos mutaciones.*
- bool **operator<** (const **mutacion** &m) const  
*Compara si una mutacion es mayor que otra, comparando por Chr y pos.*

### Atributos privados

- string `ID`
- string `chr`
- unsigned int `pos`
- vector< string > `ref_alt`
- vector< string > `genes`
- bool `common`
- vector< float > `caf`
- vector< enfermedad > `enfermedades`
- vector< int > `clnsig`

#### 4.2.1. Descripción detallada

Clase `mutacion`, asociada a la definición de una mutación/SNP.

`mutacion::mutacion`, `mutacion::setID`, `mutacion::setChr`, `mutacion::setPos`, `mutacion::setRef_alt`, `mutacion::setGenes`, `mutacion::setCommon`, `mutacion::setCaf`, `mutacion::setEnfermedades`, `mutacion::setClnsig`, `mutacion::getID`, `mutacion::getChr`, `mutacion::getPos`, `mutacion::getRef_alt`, `mutacion::getGenes`, `mutacion::getCommon`, `mutacion::getCaf`, `mutacion::getEnfermedades`, `mutacion::getClnsig`, `mutacion::operator=`, `mutacion::operator==`, `mutacion::operator<`

Descripción: contiene toda la información asociada a una mutación almacenada en la BD ClinVar-dbSNP

### Autor

Elena María Gómez Ríos

#### 4.2.2. Documentación del constructor y destructor

##### 4.2.2.1. `mutacion::mutacion ( )`

Constructor de enfermedad por defecto.

fichero de implementacion de la clase `mutacion`

##### 4.2.2.2. `mutacion::mutacion ( const mutacion & m )`

Constructor de copia de la clase.

### Parámetros

<code>in</code>	<code>m</code>	mutacion a copiar
-----------------	----------------	-------------------

##### 4.2.2.3. `mutacion::mutacion ( const string & str )`

Constructor parametrizado de la clase.

### Parámetros

<code>in</code>	<code>str</code>	cadena que contiene todos los atributos de la clase
-----------------	------------------	---

#### 4.2.3. Documentación de las funciones miembro

##### 4.2.3.1. `const vector< float > & mutacion::getCaf ( ) const`

Obtiene la frecuencia de cada base del SNP en la población.

Devuelve

la frecuencia de cada base del SNP en la población.

##### 4.2.3.2. `string mutacion::getChr ( ) const`

Obtiene identificador del cromosoma.

Devuelve

identificador del cromosoma.

##### 4.2.3.3. `const vector< int > & mutacion::getCInsig ( ) const`

Obtiene la relevancia clínica del SNP para cada enfermedad.

Devuelve

la relevancia clínica del SNP para cada enfermedad.

##### 4.2.3.4. `bool mutacion::getCommon ( ) const`

Obtiene el indicador de si el SNP es común en la población.

Devuelve

el indicador de si el SNP es común en la población.

##### 4.2.3.5. `const vector< enfermedad > & mutacion::getEnfermedades ( ) const`

Obtiene las enfermedades asociadas al SNP.

Devuelve

las enfermedades asociadas al SNP.

##### 4.2.3.6. `const vector< string > & mutacion::getGenes ( ) const`

Obtiene los gen(es) asociado(s) al SNP.

Devuelve

los gen(es) asociado(s) al SNP.



**4.2.3.7. string mutacion::getID ( ) const**

Obtiene el identificador del SNP/mutación.

Devuelve

identificador del SNP/mutación.

**4.2.3.8. unsigned int mutacion::getPos ( ) const**

Obtiene el identificador de la posición dentro del cromosoma.

Devuelve

identificador de la posición dentro del cromosoma.

**4.2.3.9. const vector< string > & mutacion::getRef\_alt ( ) const**

Obtiene la(s) base(s) en el genoma de referencia y alternativa(s) posible(s).

Devuelve

la(s) base(s) en el genoma de referencia y alternativa(s) posible(s).

**4.2.3.10. bool mutacion::operator< ( const mutacion & m ) const**

Compara si una mutacion es mayor que otra, comparando por Chr y pos.

Parámetros

in	e	enfermedad a comparar.
----	---	------------------------

Devuelve

Devuelve true si e es mayor que el que lo llama. False en otro caso.

**4.2.3.11. mutacion & mutacion::operator= ( const mutacion & m )**

operador de asignación

Parámetros

in	m	mutación a copiar.
----	---	--------------------

Devuelve

copia de la mutación.

4.2.3.12. `bool mutacion::operator==( const mutacion & m ) const`

Compara si son iguales dos mutaciones.

**Parámetros**

in	<i>m</i>	mutación a comparar.
----	----------	----------------------

**Devuelve**

Devuelve true si son iguales y false si no lo son.

4.2.3.13. `void mutacion::setCaf ( const std::vector< float > & caf )`

Establece la frecuencia de cada base del SNP en la población.

**Parámetros**

in	<i>caf</i>	frecuencia de cada base del SNP en la población.
----	------------	--

4.2.3.14. `void mutacion::setChr ( const string & chr )`

Establece el identificador del cromosoma.

**Parámetros**

in	<i>chr</i>	nombre de el identificador del cromosoma.
----	------------	---

4.2.3.15. `void mutacion::setClnsig ( const std::vector< int > & clnsig )`

Establece la relevancia clínica del SNP para cada enfermedad.

**Parámetros**

in	<i>clnsig</i>	relevancia clínica del SNP para cada enfermedad.
----	---------------	--

4.2.3.16. `void mutacion::setCommon ( const bool & common )`

Establece el indicador de si el SNP es común en la población.

**Parámetros**

in	<i>common</i>	indicador de si el SNP es común en la población.
----	---------------	--

4.2.3.17. `void mutacion::setEnfermedades ( const std::vector< enfermedad > & enfermedades )`

Establece las enfermedades asociadas al SNP.

## Parámetros

in	<i>enfermedades</i>	enfermedades asociadas al SNP.
----	---------------------	--------------------------------

## 4.2.3.18. void mutacion::setGenes ( const std::vector&lt; string &gt; &amp; genes )

Establece los gen(es) asociado(s) al SNP.

## Parámetros

in	<i>genes</i>	gen(es) asociado(s) al SNP.
----	--------------	-----------------------------

## 4.2.3.19. void mutacion::setID ( const string &amp; id )

Establece el identificador del SNP/mutación.

## Parámetros

in	<i>id</i>	nombre de el identificador del SNP/mutación.
----	-----------	--

## 4.2.3.20. void mutacion::setPos ( const unsigned int &amp; pos )

Establece el identificador de la posición dentro del cromosoma.

## Parámetros

in	<i>pos</i>	nombre de el identificador de la posición dentro del cromosoma.
----	------------	---

## 4.2.3.21. void mutacion::setRef\_alt ( const std::vector&lt; string &gt; &amp; ref\_alt )

Establece la(s) base(s) en el genoma de referencia y alternativa(s) posible(s).

## Parámetros

in	<i>ref_alt</i>	base(s) en el genoma de referencia y alternativa(s) posible(s).
----	----------------	---

## 4.2.4. Documentación de los datos miembro

## 4.2.4.1. vector&lt;float&gt; mutacion::caf [private]

Frecuencia de cada base del SNP en la población. En primer lugar debe indicarse la frecuencia de la base 'ref' (posición 0 de ref-alt), seguida por las frecuencias de las bases alternativas indicadas en 'ref-alt', en el mismo orden.

## 4.2.4.2. string mutacion::chr [private]

Identificador del cromosoma

#### 4.2.4.3. `vector<int> mutacion::clnsig` [private]

Relevancia clínica del SNP para cada enfermedad utilizando el código numérico del campo CLNSIG. En caso de que existan varias enfermedades asociadas a la mutación, cada una de ellas puede presentar diferente código CLNSIG, por lo que se deben almacenar en el vector clnsig en el mismo orden que las enfermedades asociadas. En caso de presentarse sólo un código CLNSIG y varias enfermedades, este código se aplica a todas ellas.

#### 4.2.4.4. `bool mutacion::common` [private]

Indica si el SNP es común en la población

#### 4.2.4.5. `vector<enfermedad> mutacion::enfermedades` [private]

Enfermedades asociadas al SNP

#### 4.2.4.6. `vector<string> mutacion::genes` [private]

Gen(es) asociado(s) al SNP

#### 4.2.4.7. `string mutacion::ID` [private]

Identificador del SNP/mutación

#### 4.2.4.8. `unsigned int mutacion::pos` [private]

Identificador de la posición dentro del cromosoma

#### 4.2.4.9. `vector<string> mutacion::ref_alt` [private]

Base(s) en el genoma de referencia y alternativa(s) posible(s). La primera posición la ocupará el string con la(s) base(s) del genoma de referencia, y, a continuación, aparecerán la(s) base(s) alternativas en el mismo orden que se indica en el fichero

La documentación para esta clase fue generada a partir de los siguientes ficheros:

- `include/mutacion.h`
- `include/mutacion.hxx`

## 5. Documentación de archivos

### 5.1. Referencia del Archivo `documentacion.dox`

### 5.2. Referencia del Archivo `include/enfermedad.h`

```
#include <string>
#include <iostream>
#include <algorithm>
#include "enfermedad.hxx"
```

## Clases

- class `enfermedad`

*Clase enfermedad, asociada al TDA enfermedad.*

## Funciones

- `ostream & operator<< (ostream &os, const enfermedad &e)`

*Imprime enfermedad con todos sus campos.*

### 5.2.1. Documentación de las funciones

#### 5.2.1.1. `ostream& operator<< ( ostream & os, const enfermedad & e )`

Imprime enfermedad con todos sus campos.

#### Parámetros

in	<i>os</i>	flujo.
in	<i>e</i>	enfermedad a mostrar.

#### Devuelve

La enfermedad por salida estandar.

## 5.3. Referencia del Archivo include/enfermedad.hxx

## Funciones

- `ostream & operator<< (ostream &os, const enfermedad &e)`

*Imprime enfermedad con todos sus campos.*

### 5.3.1. Documentación de las funciones

#### 5.3.1.1. `ostream& operator<< ( ostream & os, const enfermedad & e )`

Imprime enfermedad con todos sus campos.

#### Parámetros

in	<i>os</i>	flujo.
in	<i>e</i>	enfermedad a mostrar.

#### Devuelve

La enfermedad por salida estandar.

## 5.4. Referencia del Archivo include/mutacion.h

```
#include <string>
#include <iostream>
#include <vector>
#include <sstream>
#include "enfermedad.h"
#include "mutacion.hxx"
```

### Clases

- class `mutacion`

*Clase mutacion, asociada a la definición de una mutación/SNP.*

### Funciones

- ostream & `operator<<` (ostream &os, const `mutacion` &m)

*Imprime mutacion con todos sus campos.*

#### 5.4.1. Documentación de las funciones

##### 5.4.1.1. ostream& operator<< ( ostream & os, const mutacion & m )

Imprime mutacion con todos sus campos.

#### Parámetros

in	os	flujo.
in	m	mutacion a mostrar.

### Devuelve

La enfermedad por salida estandar.

## 5.5. Referencia del Archivo include/mutacion.hxx

### Funciones

- ostream & `operator<<` (ostream &os, const `mutacion` &m)

*Imprime mutacion con todos sus campos.*

#### 5.5.1. Documentación de las funciones

##### 5.5.1.1. ostream& operator<< ( ostream & os, const mutacion & m )

Imprime mutacion con todos sus campos.

**Parámetros**

in	os	flujo.
in	m	mutacion a mostrar.

**Devuelve**

La enfermedad por salida estandar.

**5.6. Referencia del Archivo src/principal.cpp**

```
#include "mutacion.h"
#include "enfermedad.h"
#include <iostream>
#include <fstream>
#include <vector>
```

**Funciones**

- **bool load** (vector< **mutacion** > &vm, const string &s)  
*lee un fichero de mutaciones, linea a linea*
- **int cuentaMutacionesEnfermedad** (vector< **mutacion** > &vm, const string &s)  
*Recorre un vector de mutaciones y devuelve cuántas de estas mutaciones están asociadas a un nombre de enfermedad s.*
- **int main** (int argc, char \*argv[ ])

**5.6.1. Documentación de las funciones****5.6.1.1. int cuentaMutacionesEnfermedad ( vector< mutacion > & vm, const string & s )**

Recorre un vector de mutaciones y devuelve cuántas de estas mutaciones están asociadas a un nombre de enfermedad s.

**Parámetros**

in	vm	vector de mutaciones
in	s	texto asociado al nombre de la enfermedad.

**Devuelve**

int número de mutaciones asociadas a enfermedades cuyo nombre contiene s

**5.6.1.2. bool load ( vector< mutacion > & vm, const string & s )**

lee un fichero de mutaciones, linea a linea

**Parámetros**

<code>in</code>	<code>s</code>	nombre del fichero
<code>in, out</code>	<code>vm</code>	vector sobre el que se lee

**Devuelve**

true si la lectura ha sido correcta, false en caso contrario

5.6.1.3. `int main ( int argc, char * argv[] )`





## Índice alfabético

- caf
  - mutacion, 18
- chr
  - mutacion, 18
- clnsig
  - mutacion, 18
- common
  - mutacion, 19
- cuentaMutacionesEnfermedad
  - principal.cpp, 22
- database
  - enfermedad, 12
- documentacion.dox, 19
- enfermedad, 8
  - database, 12
  - enfermedad, 9
  - getDatabase, 9
  - getID, 9
  - getName, 10
  - ID, 12
  - name, 12
  - nameContains, 10
  - operator!=, 10
  - operator<, 10
  - operator=, 11
  - operator==, 11
  - setDatabase, 11
  - setID, 11
  - setName, 11
  - toString, 12
- enfermedad.h
  - operator<<, 20
- enfermedad.hxx
  - operator<<, 20
- enfermedades
  - mutacion, 19
- genes
  - mutacion, 19
- getCaf
  - mutacion, 15
- getChr
  - mutacion, 15
- getClnsig
  - mutacion, 15
- getCommon
  - mutacion, 15
- getDatabase
  - enfermedad, 9
- getEnfermedades
  - mutacion, 15
- getGenes
  - mutacion, 15
- getID
  - enfermedad, 9
  - mutacion, 15
- getName
  - enfermedad, 10
- getPos
  - mutacion, 16
- getRef\_alt
  - mutacion, 16
- ID
  - enfermedad, 12
  - mutacion, 19
- include/enfermedad.h, 19
- include/enfermedad.hxx, 20
- include/mutacion.h, 21
- include/mutacion.hxx, 21
- load
  - principal.cpp, 22
- main
  - principal.cpp, 23
- mutacion, 12
  - caf, 18
  - chr, 18
  - clnsig, 18
  - common, 19
  - enfermedades, 19
  - genes, 19
  - getCaf, 15
  - getChr, 15
  - getClnsig, 15
  - getCommon, 15
  - getEnfermedades, 15
  - getGenes, 15
  - getID, 15
  - getPos, 16
  - getRef\_alt, 16
  - ID, 19
  - mutacion, 14
  - operator<, 16
  - operator=, 16
  - operator==, 16
  - pos, 19
  - ref\_alt, 19
  - setCaf, 17
  - setChr, 17
  - setClnsig, 17
  - setCommon, 17
  - setEnfermedades, 17
  - setGenes, 18
  - setID, 18
  - setPos, 18
  - setRef\_alt, 18
- mutacion.h
  - operator<<, 21

mutacion.hxx  
    operator<<, 21

name  
    enfermedad, 12

nameContains  
    enfermedad, 10

operator!=  
    enfermedad, 10

operator<  
    enfermedad, 10  
    mutacion, 16

operator<<  
    enfermedad.h, 20  
    enfermedad.hxx, 20  
    mutacion.h, 21  
    mutacion.hxx, 21

operator=  
    enfermedad, 11  
    mutacion, 16

operator==  
    enfermedad, 11  
    mutacion, 16

pos  
    mutacion, 19

principal.cpp  
    cuentaMutacionesEnfermedad, 22  
    load, 22  
    main, 23

ref\_alt  
    mutacion, 19

setCaf  
    mutacion, 17

setChr  
    mutacion, 17

setClnsig  
    mutacion, 17

setCommon  
    mutacion, 17

setDatabase  
    enfermedad, 11

setEnfermedades  
    mutacion, 17

setGenes  
    mutacion, 18

setID  
    enfermedad, 11  
    mutacion, 18

setName  
    enfermedad, 11

setPos  
    mutacion, 18

setRef\_alt  
    mutacion, 18

src/principal.cpp, 22

toString  
    enfermedad, 12