Choose a disease of your interest and investigate about it using BOTH Open Targets and HumanMine platforms.

Write what you have learned about it and upload a report here (max 4-5 pages, you can include screenshots).

Akira Hirata & Marc Trujillo

Introduction

For this practical we were given the option to choose a disease of interest and to perform an investigation about it. This investigation must be by using the Open Targets and the HumanMine platforms. First of all, we have chosen the Osteogenesis Imperfecta (OI) which is a disease that makes the one that has it break their bones easily, it's also known as the crystal bones disease.

Open Target
If we go into the main page for Open Targets and search our disease we obtain this:
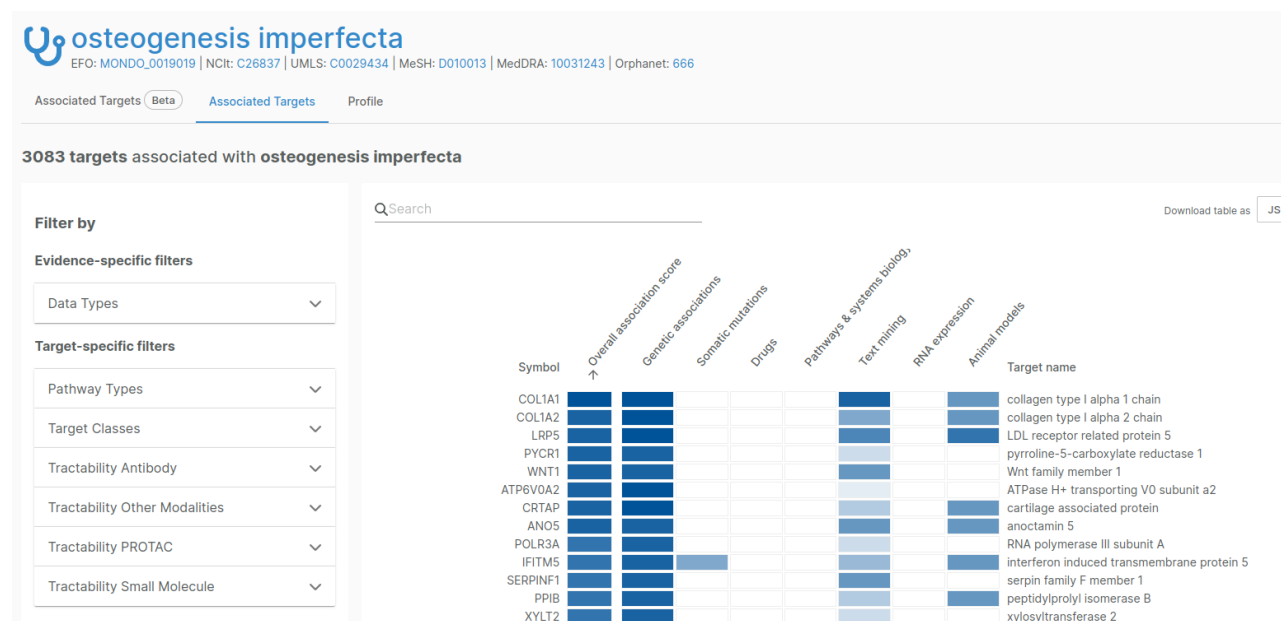
As we know from the practical, there are 3083 targets associated with our disease, and the data types that support this association are Animal models, Somatic mutations, Drugs, Text mining and Genetic associations. Now that we identified the basic elements, we go into the content. The first thing to notice are the 2 targets that show the highest association with the Osteogenesis Imperfecta, which are the COL1A1 and the COL1A2, that show the highest score and their colors are similar, and we go into their information their description is the same: Type I collagen is a member of group I collagen (fibrillar forming collagen), but the difference is the alpha chain that they belong to. Also the third target shows similar colors to

the first ones. Before going into the page of the highest scoring target we will take a look at the Profile section and observe the Known Drugs that we have for our disease:

Disease information	Drug information				Target information		Clinical trials information		
Disease	Drug	Type	Mechanism Of Action	Action Type	Symbol	Name	Phase	Status	Source
postmenopausal osteoporosis	BAZEDOXIFENE	Small molecule	Estrogen receptor modulator	Modulator	ESR2	estrogen receptor 2	Phase IV	Completed	ClinicalTrials.
osteogenesis imperfecta	CHOLECALCIFEROL	Small molecule	Vitamin D receptor agonist	Agonist	VDR	vitamin D receptor	Phase IV	Completed	ClinicalTrials.
osteoporosis	ERGOCALCIFEROL	Small molecule	Vitamin D receptor agonist	Agonist	VDR	vitamin D receptor	Phase IV	Unknown status	ClinicalTrials.
postmenopausal osteoporosis	CALCITONIN SALMON RECOMBINANT	Protein	Calcitonin receptor agonist	Agonist	CALCR	calcitonin receptor	Phase IV	N/A	FDA
osteoporosis	ROMOSOZUMAB	Antibody	Sclerostin inhibitor	Inhibitor	SOST	sclerostin	Phase IV	Active, not recruiting	ClinicalTrials.
osteoporosis	IBANDRONIC ACID	Small molecule	Farnesyl diphosphate synthase inhibitor	Inhibitor	FDPS	farnesyl diphosphate synthase	Phase IV	Terminated	ClinicalTrials.
postmenopausal osteoporosis	ERGOCALCIFEROL	Small molecule	Vitamin D receptor agonist	Agonist	VDR	vitamin D receptor	Phase IV	Unknown status	ClinicalTrials.
osteoporosis	TERIPARATIDE	Protein	Parathyroid hormone receptor agonist	Agonist	PTH1R	parathyroid hormone 1 receptor	Phase IV	Completed	10 references
osteogenesis imperfecta	ZOLEDRONIC ACID	Small molecule	Farnesyl diphosphate synthase inhibitor	Inhibitor	FDPS	farnesyl diphosphate synthase	Phase IV	Recruiting	ClinicalTrials.
osteoporosis	TERIPARATIDE	Protein	Parathyroid hormone receptor agonist	Agonist	PTH1R	parathyroid hormone 1 receptor	Phase IV	Terminated	ClinicalTrials.

Rows per page: 10 ▾   41–50 of 368   ‹   ›

We had to go to page 41 because the description tells us that our disease comprises a heterogeneous group of genetic disorders, and that's why we find drugs for diseases like Osteoporosis or Postmenopausal Osteoporosis. The first drug that we find is Cholecalciferol and it's a Vitamin D receptor agonist and has been applied to a Phase IV of Clinical trial with completed status. A main reason why they could have chosen this small molecule is because it is necessary for the body to properly absorb calcium and phosphate. In order to form strong and healthy bones are developed and maintained in large part by these minerals, Cholecalciferol or also known as Vitamin D3 .

Now we go into the best hit, which is the COL1A1 and the first thing we observe is that the only sections that it has are about the Genetic associations, Text mining and Animal models

COL1A1

If we center our investigation in this COL1A1 we can find some valuable information. Also known as collagen type I alpha 1 chain and located in the endoplasmic reticulum. COL1A1 is involved in the synthesis of the alpha 1 chain of type I collagen. Type I collagen is a fibrillar collagen that forms a triple helical structure with two alpha 1 chains and one alpha 2 chain, creating a robust and flexible framework. It's also worth mentioning the fact that it is the most abundant type of collagen in the body and plays an important role in providing support and strength to various tissues, being a major component of the extracellular matrix of connective tissue.

Pathways

It participates in a lot of important cellular processes such as:

Pathway	Top-level parent pathway	View target and pathway
ECM proteoglycans	Extracellular matrix organization	📍 Reactome pathway browser
GPVI-mediated activation cascade	Hemostasis	📍 Reactome pathway browser
Platelet Adhesion to exposed collagen	Hemostasis	📍 Reactome pathway browser
MET activates PTK2 signaling	Signal Transduction	📍 Reactome pathway browser
Cell surface interactions at the vascular wall	Hemostasis	📍 Reactome pathway browser
Non-integrin membrane-ECM interactions	Extracellular matrix organization	📍 Reactome pathway browser
Collagen biosynthesis and modifying enzymes	Extracellular matrix organization	📍 Reactome pathway browser
Collagen degradation	Extracellular matrix organization	📍 Reactome pathway browser
Assembly of collagen fibrils and other multimeric structures	Extracellular matrix organization	📍 Reactome pathway browser
Crosslinking of collagen fibrils	Extracellular matrix organization	📍 Reactome pathway browser

Rows per page: 10 ▾   1–10 of 20 |◁ ◁ ▷ ▷|

So we prove that the COL1A1 plays a key role in the extracellular matrix, a structural + helping in the organization. On the other hand we've also found that it participates in the process of homeostasis of the cell even though it's rather indirect, as the COL1A1 only supports the tissues and structures that participate in the homeostasis.

Molecular interactions

We can also search a lot more information about the COL1A1. For example the it's most important and frequent molecular interactions such as the FN1(fibronectin)CRTAP(Cartilage associated protein) or the COL1A2.

FN1 Alt ID: P02751	0.71	A B	3	▶
CRTAP Alt ID: O75718	0.58	A B	3	
COL1A2 Alt ID: P08123	0.58	A B	5	
P4HB Alt ID: P07237	0.58	A B	3	
HYOU1 Alt ID: Q9Y4L1	0.56	A B	3	
DNAJB11 Alt ID: Q9UBS4	0.56	A B	3	
KEAP1 Alt ID: Q14145	0.56	A B	3	

The interaction with the better score(0.71) is with the FN1 also known as fibronectin. Fibronectin is a critically important ECM protein that mediates cell-ECM interaction during fundamental events such as development, wound healing, fibrosis, and tumor progression. And that's why it has such a highscore with the COL1A1, because it's located in the ECM as well and regulates the interaction of that part with the rest of the cell(which makes FN1 interact with the most abundant component of the connective tissues of the ECM

Another important remark to make is something that took our attention which is the Known Drugs targeting the COL1A1,

We have found one that is extremely repeated and is the COLLAGENASE CLOSTRIDIUM. This disease is an enzyme drug used for therapeutic purposes. It affects the COL1A1 as it has the ability to break down collagen. As appears on the screenshot taken directly from Open Target, this drug is really useful when treating multiple diseases related with tissues, obviously thanks to the ability of targeting collagen.

HumanMine

Now if we go into the HumanMine the first thing we have to do is to look for our disease. Now we will go and look at the associated diseases of this gene and the tissues where is most highly expressed:

Diseases (7)

Showing 1 to 5 of 7 rows

Diseases Primary Identifier	Diseases Name
OMIM:166210	OSTEOGENESIS IMPERFECTA, TYPE II
OMIM:166220	OSTEOGENESIS IMPERFECTA, TYPE IV
OMIM:166710	OSTEOPOROSIS
OMIM:225320	EHLERS-DANLOS SYNDROME, CARDIAC VALVULAR TYPE
OMIM:259420	OSTEOGENESIS IMPERFECTA, TYPE III

Atlas Expression (156)

Showing 1 to 5 of 156 rows

Atlas Expression T Statistic	Atlas Expression Type	Atlas Expression Condition	Atlas Expression Expression	Atlas Expression P Value
2.5	disease_state	astrocytoma	UP	0.040972765535116196
2.5	disease_state	breast epithelial carcinosarcoma	UP	0.04251117631793022
2.5	organism_part	ovary	UP	0.04111003130674362
2.5	organism_part	placental basal plate	UP	0.04749387502670288
2.5	organism_part	right knee	UP	0.04426585137844086

We can also find information about the most the most important genes, as their length, identifier or the organism were they belong

Genes (2)

Showing 1 to 2 of 2 rows

Genes Symbol	Genes Name	Genes Primary Identifier	Genes Secondary Identifier	Genes Length	Organism Name
COL1A1	collagen type I alpha 1 chain	1277	ENSG00000108821	17554	Homo sapiens
COL1A2	collagen type I alpha 2 chain	1278	ENSG00000164692	36333	Homo sapiens

Following the example of stepik we've also found all different types of osteogenesis imperfecta and following the steps in stepik we have generated a column to see them(all 26). Stepik is also useful to see the upregulated and downregulated genes finding the template Tissue --> Gene Expression (Array Express) and putting the condition of "connective tissue" Will let you see a huge table with the format

Atlas Expression Condition	Gene Primary Identifier	Gene Symbol	Gene Name	Atlas Expression Expression	Atlas Expression P Value
connective tissue	10	NAT2	N-acetyltransferase 2	DOWN	0.0000044065477

That will allow you to see, if it's up or down, the gene identifier and the P value

Showing 1 to 10 of 26 rows
Rows per page: 10 — Page 1

Organism Name	Diseases Name	Gene Symbol
Homo sapiens	COMBINED OSTEOGENESIS IMPERFECTA AND EHLERS-DANLOS SYNDROME 1	COL1A1
Homo sapiens	COMBINED OSTEOGENESIS IMPERFECTA AND EHLERS-DANLOS SYNDROME 2	COL1A2
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE I	COL1A1
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE II	COL1A1
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE II	COL1A2
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE III	COL1A1
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE III	COL1A2
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE IV	COL1A1
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE IV	COL1A2
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE IX	PPIB

Rows per page: 10 — Page 2

Organism Name	Diseases Name	Gene Symbol
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE V	IFITM5
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE VI	SERPINF1
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE VII	CRTAP
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE VIII	P3H1
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE X	SERPINH1
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE XI	FKBP10
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE XII	SP7
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE XIII	BMP1
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE XIV	TMEM38B
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE XIX	MBTPS2

Rows per page: 10 — Page 3

Organism Name	Diseases Name	Gene Symbol
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE XV	WNT1
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE XVI	CREB3L1
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE XVII	SPARC
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE XVIII	TENT5A
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE XX	MESD
Homo sapiens	OSTEOGENESIS IMPERFECTA, TYPE XXI	KDELR2

(photos of all osteogenesis)