Genetic and BioInformatics Analysis of **Celecoxib** Metabolism: A **CYP2C9**-Focused Study

- I certify that I have not used AI tools, like ChatGPT, or other AI tools to construct this document or responses to questions in this form. I certify the item is created by myself without any AI.
- All graphics created by student researcher unless otherwise cited therein.

Introduction

Objective	• To investigate how the gene, CYP2C9, and its genetic variants impact Celecoxib metabolism, a key non-steroidal anti-inflammatory drug (NSAID) for arthritis. • To analyze how similar genes and gene variants affect drug efficacy and their side effects, using Bioinformatics tools (NCBI, PharmGKB, PubChem, STRING, EMBOSS Water, SIFT, PolyPhen-2, MutationTaster)
Motivation (Why)	• Inspired by my family's history of arthritis and varying responses to drugs like Celecoxib. • To start my inquisitive journey in the field of Bioinformatics.
Problem	• CYP2C9 variations in the human body can slow Celecoxib metabolism, leading to higher drug levels, which increases the risk of side effects (GI bleeding, cardiovascular issues) if doses aren't adjusted.
Research Findings	• SNP analysis revealed *CYP2C9**2 and *CYP2C9**3 variants slow down Celecoxib metabolism, increasing risks of side effects. This reinforces the need to reduce Celecoxib dosage. • Highlights the need for genetic testing to personalize Celecoxib dosing, especially for those with a family history of arthritis. • Aims to uncover genetic factors influencing drug metabolism for better treatment strategies.

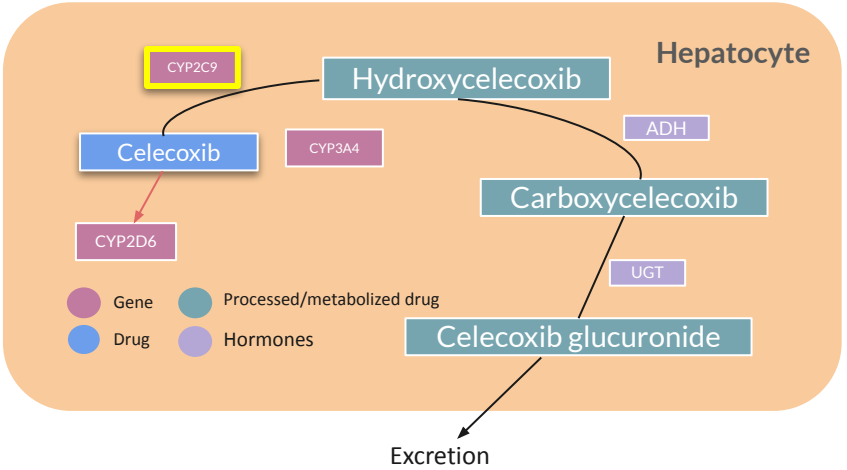
Celecoxib Overview

Celecoxib is a selective Cyclooxygenase-2 (COX-2) inhibitor for arthritis, pain, inflammation, dysmenorrhea, and polyp reduction.

Pharmacokinetics:

- Absorption: Peaks ~3 hours post-dose.
- Metabolism: CYP2C9 → hydroxycelecoxib → ADH → carboxycelecoxib (inactive) → Celecoxib glucuronide.
- Excretion: Urine & bile.

Mechanism of Action: Blocks COX-2, reducing prostaglandins to relieve pain & inflammation.

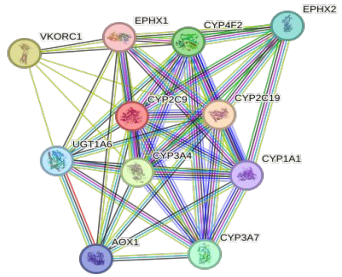


Drug Advantages	Potential Side Effects	Precautions	Undiscovered Genetic Factors
Arthritis Relief: Improves joint function, reduces stiffness, lowers GI risks vs. NSAIDs. **Benefits:** Fewer stomach issues, safer for ulcers, minimal blood clotting impact.	**Risks:** High doses/long-term use may increase heart attack, stroke, GI bleeding, or kidney issues.	**Precautions:** Avoid in NSAID allergies; use cautiously with heart, liver, or kidney conditions; not for late pregnancy.	**Genetics & Metabolism:** CYP2C9 variants slow metabolism, increasing drug buildup and risk of side effects.

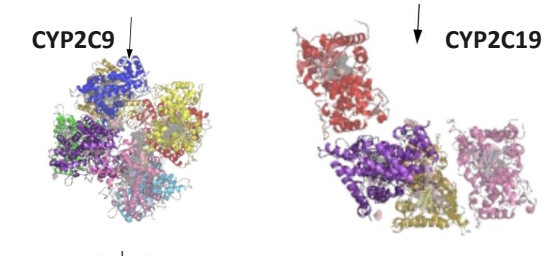
Methodology

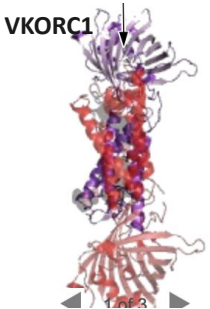
Steps	Outcome	Slide
1. Collect Celecoxib pharmacokinetic data from PubChem & PharmGKB. Find CYP2C9 information.	To understand Celecoxib's function, metabolism, and key genes involved (helped identify CYP2C9).	3
2. Analyze the gene, CYP2C9's protein interactions using STRING.	To identify other genes interacting with CYP2C9 to hypothesize their role in Celecoxib metabolism.	5
3. Retrieve CYP2C9 and related protein sequences from NCBI. Align sequences using EMBOSS Water.	To compare genetic differences of the identified genes in step 2, and to predict their impact on metabolism.	6
4. Find Level 1A variants in PharmGKB, searching for CYP2C9 and Celecoxib.	To find gene variants CYP2C9*2 (rs1799853) and CYP2C9*3 (rs1057910).	7
5. Input the variants from step 4 in dbSNP.	To determine nucleotide and amino acid changes.	8
6. Compare Celecoxib metabolism in CYP2C9 vs. the variants.	To assess the impact of polymorphisms on drug metabolism and side effects, to draw final conclusions.	8
7. Predict amino acid substitution impact using SIFT, PolyPhen-2 & MutationTaster.	To evaluate whether genetic variations are harmful.	9, 10
8. Analyze variants' population frequencies.	To determine significance.	9, 10

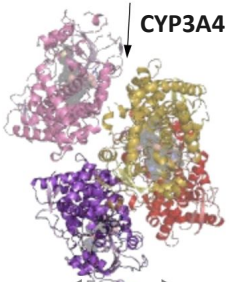
CYP2C9 Protein Interactions in Humans (STRING Database Results)



- **CYP2C19, VKORC1, and CYP3A4** had the highest interaction scores across multiple evidence categories, including **neighborhood, gene fusion, co-occurrence, coexpression, experiments, databases, text mining, and homology**.
- **VKORC1** showed particularly strong associations in text mining.
- The slide presents **protein interaction networks and coexpression data**, highlighting CYP2C9's role in broader metabolic and pharmacokinetic pathways.







CYP2C9 *including fatty acids and steroids. Mechanistically, uses molecular oxygen inserting one oxygen atom into a substrate, and reducing the second into a water molecule, with two electrons provided by NADPH via cytochrome P450 reductase (NADPH–hemoprotein reductase). Catalyzes the epoxidation of double bonds of polyunsaturated fatty acids (PUFA). Catalyzes the hydroxylation of carbon-hydrogen bonds. Metabolizes cholesterol toward 25-hydroxycholesterol, a physiological regulato [...] (490 aa)*

Predicted Functional Partners:

		Neighborhood	Gene Fusion	Cooccurrence	Coexpression	Experiments	Databases	Textmining	[Homology]	Score
CYP2C19	*Cytochrome P450 2C19; A cytochrome P450 monooxygenase involved in the metabolism of polyunsaturated fatty acids (PUF...*				●	●	●	●	●	0.988
VKORC1	*Vitamin K epoxide reductase complex subunit 1; Involved in vitamin K metabolism. Catalytic subunit of the vitamin K epoxide ...*							●		0.980
CYP3A4	*Cytochrome P450 3A4; A cytochrome P450 monooxygenase involved in the metabolism of sterols, steroid hormones, retinoid...*			●	●	●	●	●	●	0.980
CYP4F2	*Cytochrome P450 4F2; A cytochrome P450 monooxygenase involved in the metabolism of various endogenous substrates, in...*			●	●		●	●	●	0.974
CYP3A7	*Cytochrome P450 3A7; A cytochrome P450 monooxygenase involved in the metabolism of steroid hormones and vitamins d...*			●	●		●	●	●	0.971
EPHX2	*Bifunctional epoxide hydrolase 2; Bifunctional enzyme. The C-terminal domain has epoxide hydrolase activity and acts on epo...*						●	●		0.967
UGT1A6	*UDP-glucuronosyltransferase 1-6; UDPGT is of major importance in the conjugation and subsequent elimination of potentially...*				●		●	●		0.964
AOX1	*Aldehyde oxidase; Oxidase with broad substrate specificity, oxidizing aromatic azaheterocycles, such as N1-methylnicotinami...*				●		●	●		0.964
CYP1A1	*Cytochrome P450 1A1; A cytochrome P450 monooxygenase involved in the metabolism of various endogenous substrates, i...*			●	●		●	●	●	0.964
EPHX1	*Epoxide hydrolase 1; Biotransformation enzyme that catalyzes the hydrolysis of arene and aliphatic epoxides to less reactive ...*				●	●	●	●		0.963

Results (EMBOSS Water Pairwise Sequence Alignment for CYP2C9 and similar genes)

VKORC1

```
# Aligned_sequences: 2
# 1: NP_000762.2
# 2: VKOR1_HUMAN
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 117
# Identity:      27/117 (23.1%)
# Similarity:    41/117 (35.0%)
# Gaps:          45/117 (38.5%)
# Score: 48.5
#
#
#=======================================

NP_000762.2       62 GPVFTLYFGLKPIVVLHGYEAVK------EALIDLGEEFSGRGIFPLAER    105
                     |.|.:||      .|| .:|.:      .||.|:|...|...:|  :.|
VKOR1_HUMAN       19 GLVLSLY-------ALH-VKAARARDRDYRALCDVGTAISCSRVF--SSR     58

NP_000762.2      106 ANRGFGIVFSNGKKWKEIRRFSLMTLRNFGMGKRSIEDRVQEEEARCLVEE    155
                     ..||||:|                ...:|:.||.::.....|:...
VKOR1_HUMAN       59 WGRGFGLV------------------EHVLGQDSILNQSNSIFGCIFYT     89

NP_000762.2      156 LRKTKASPCDPTFILGC    172
                     |:          .:|||
VKOR1_HUMAN       90 LQ----------LLLGC     96


#---------------------------------------
#---------------------------------------
```

CYP2C19

```
# Aligned_sequences: 2
# 1: NP_000762.2
# 2: NP_000760.1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 490
# Identity:     448/490 (91.4%)
# Similarity:   467/490 (95.3%)
# Gaps:           0/490 ( 0.0%)
# Score: 2374.0
#
#
#=======================================

NP_000762.2        1 MDSLVVLVLCLSCLLLLSLWRQSSGRGKLPPGPTPLPVIGNILQIGIKDI     50
                     ||..|||||||||||||:|||||||||||||||||||||||.|||:
NP_000760.1        1 MDPFVVLVLCLSCLLLLSIWRQSSGRGKLPPGPTPLPVIGNILQIDIKDV     50

NP_000762.2       51 SKSLTNLSKVYGPVFTLYFGLKPIVVLHGYEAVKEALIDLGEEFSGRGIF    100
                     |||||||||:|||||||||:.:|||||||.|||||||||||||||.|
NP_000760.1       51 SKSLTNLSKIYGPVFTLYFGLERMVVLHGYEVVKEALIDLGEEFSGRGHF    100

NP_000762.2      101 PLAERANRGFGIVFSNGKKWKEIRRFSLMTLRNFGMGKRSIEDRVQEEAR    150
                     ||||||||||||||||:|||||||||||||||||||||||||||||||||
NP_000760.1      101 PLAERANRGFGIVFSNGKRWKEIRRFSLMTLRNFGMGKRSIEDRVQEEAR    150

NP_000762.2      151 CLVEELRKTKASPCDPTFILGCAPCNVICSIIFHKRFDYKDQQFLNLMEK    200
                     ||||||||||||||||||||||||||||||.|||||||||||||||||||
NP_000760.1      151 CLVEELRKTKASPCDPTFILGCAPCNVICSIIFQKRFDYKDQQFLNLMEK    200

NP_000762.2      201 LNENIKILSSPWIQICNNFSPIIDYFPGTHNKLLKNVAFMKSYILEKVKE    250
                     |||||:|:|:|||||||||..|||||||||||||||:|||:|.|||||||
NP_000760.1      201 LNENIRIVSTPWIQICNNFPTIIDYFPGTHNKLLKNLAFMESDILEKVKE    250

NP_000762.2      251 HQESMDMNNPQDFIDCFLMKMEKEKHNQPSEFTIESLENTAVDLFGAGTE    300
                     |||||||:|||:||||||:|||||.||.|||||||:..||.||.||||||
NP_000760.1      251 HQESMDINNPRDFIDCFLIKMEKEKQNQQSEFTIENLVITAADLLGAGTE    300

NP_000762.2      301 TTSTTLRYALLLLLKHPEVTAKVQEEIERVIGRNRSPCMQDRSHMPYTDA    350
                     ||||||||||||||||||||||||||||||||||||||||||.|||||||
NP_000760.1      301 TTSTTLRYALLLLLKHPEVTAKVQEEIERVIGRNRSPCMQDRGHMPYTDA    350

NP_000762.2      351 VVHEVQRYIDLLPTSLPHAVTCDIKFRNYLIPKGTTILISLTSVLHDNKE    400
                     ||||||||||:||||||||||:|||||||||||||.|||||||||||||
NP_000760.1      351 VVHEVQRYIDLIPTSLPHAVTCDVKFRNYLIPKGTTILTSLTSVLHDNKE    400

NP_000762.2      401 FPNPEMFDPHHFLDEGGNFKKSKYFMPFSAGKRICVGEALAGMELFLFLT    450
                     |||||||||.|||||||||||.|||||||||||||||.||.|||||||||
NP_000760.1      401 FPNPEMFDPRHFLDEGGNFKKSNYFMPFSAGKRICVGEGLARMELFLFLT    450

NP_000762.2      451 SILQNFNLKSLVDPKNLDTTPVVNGFASVPPFYQLCFIPV    490
                     .||||||||||:|||:|||||||||||||||||||||||
NP_000760.1      451 FILQNFNLKSLIDPKDLDTTPVVNGFASVPPFYQLCFIPV    490
```

CYP3A4

```
# Aligned_sequences: 2
# 1: NP_000762.2
# 2: AAF21034.1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 495
# Identity:     141/495 (28.5%)
# Similarity:   212/495 (42.8%)
# Gaps:          83/495 (16.8%)
# Score: 385.5
#
#
#=======================================

NP_000762.2        7 LVLCLSCLLLLSLWRQSSGRGKL--PPGPTPLPVIGNILQIGIKDISKSL     54
                     |:|.:|.:||......|.|..|.  .|||||||.:||||... |......
AAF21034.1        13 LLLAVSLVLLYLYGTHSHGLFKKLGIPGPTPLPFLGNILSYH-KGFCMFD     61

NP_000762.2       55 TNLSKVYGPVFTLYFGLKPIVVLHGYEAVKEALI-DLGEEFSGRGIF-PL    102
                     ....|.||.|:..|.|.:|::.:...:.:|..|: :....|:.|..| |:
AAF21034.1        62 MECHKKYGKVWGFYDGQQPVLAITDPDMIKTVLVKECYSVFTNRRPFGPV    111

NP_000762.2      103 AERANRGF---GIVFSNGKKWKEIRRFSLMTLRNFGMGK-RSIEDRVQEE    148
                        ||   .|..:..::||.:|  ||:: ..|..|| :.:.:..:.:.
AAF21034.1       112 ------GFMKSAISIAEDEEWKRLR--SLLS-PTFTSGKLKEMVPIIAQY    152

NP_000762.2      149 ARCLVEELRK--TKASPCDPTFILGCAPCNVICSIIFH------------    184
                     ...||..||:  ....|.....:.|....:||.|..|.
AAF21034.1       153 GDVLVRNLRREAETGKPVTLKDVFGAYSMDVITSTSFGVNIDSLNNPQDP    202

NP_000762.2      185 --------KRFDYKDQQFLNLMEKLNENIKILSSPWIQICNNFSPIIDYF    226
                             .|||:.|..||::      .:.....|.:::.|    |..|
AAF21034.1       203 FVENTKKLLRFDFLDPFFLSI------TVFPFLIPILEVLN-----ICVF    241
```

- **CYP2C9** has the highest similarity (**91.4% identity, 95.3% similarity**).
- **VKORC1 (35%)** and **CYP3A4 (42.8%)** show moderate similarity.
- High similarity suggests **CYP2C19 may collaborate with CYP2C9** in Celecoxib metabolism.
- Supports PharmGKB data on **CYP2C9 and CYP3A4 interaction** in drug metabolism.

Results (EMBOSS Water Pairwise Sequence Alignment for CYP2C9 and similar genes)

VKORC1

```
# Aligned_sequences: 2
# 1: NP_000762.2
# 2: VKOR1_HUMAN
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 117
# Identity:      27/117 (23.1%)
# Similarity:    41/117 (35.0%)
# Gaps:          45/117 (38.5%)
# Score: 48.5
#
#
#=======================================

NP_000762.2       62 GPVFTLYFGLKPIVVLHGYEAVK------EALIDLGEEFSGRGIFPLAER    105
                        |.|.:||      .|| .:|.:      .||.|:|...|...:|  :.|
VKOR1_HUMAN       19 GLVLSLY-------ALH-VKAARARDRDYRALCDVGTAISCSRVF--SSR     58

NP_000762.2      106 ANRGFGIVFSNGKKWKEIRRFSLMTLRNFGMGKRSIEDRVQEEARCLVEE    155
                        ..||||:|                 ...:|:.||.::.....|:...
VKOR1_HUMAN       59 WGRGFGLV------------------EHVLGQDSILNQSNSIFGCIFYT     89

NP_000762.2      156 LRKTKASPCDPTFILGC    172
                        |:          .:|||
VKOR1_HUMAN       90 LQ----------LLLGC     96


#---------------------------------------
#---------------------------------------
```

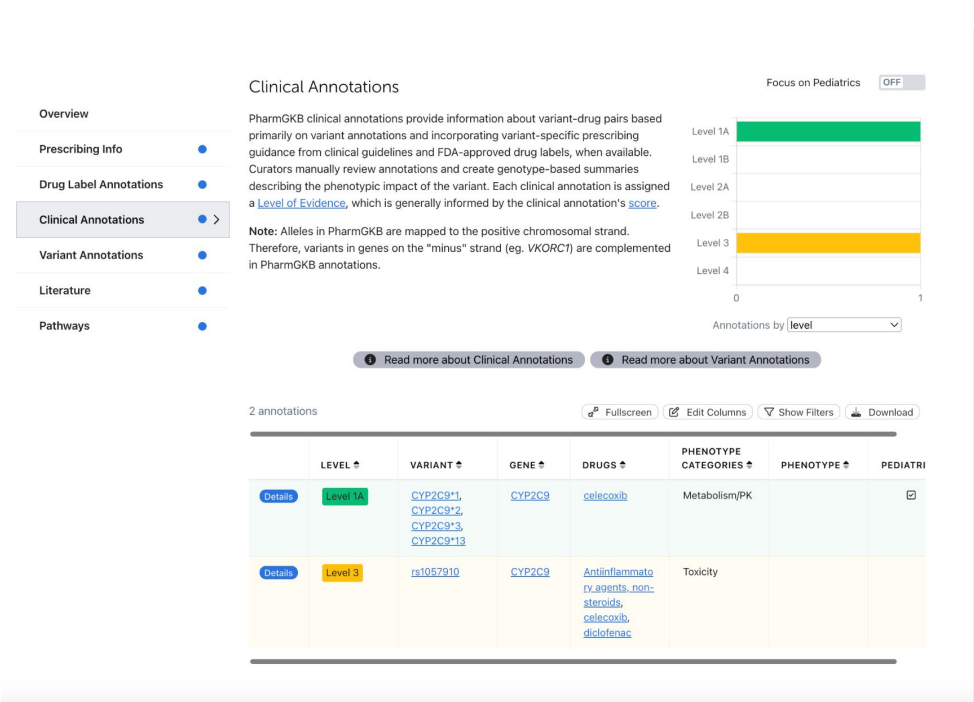
CYP2C19

```
# Aligned_sequences: 2
# 1: NP_000762.2
# 2: NP_000760.1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 490
# Identity:     448/490 (91.4%)
# Similarity:   467/490 (95.3%)
# Gaps:           0/490 ( 0.0%)
# Score: 2374.0
#
#
#=======================================

NP_000762.2        1 MDSLVVLVLCLSCLLLLSLWRQSSGRGKLPPGPTPLPVIGNILQIGIKDI     50
                     ||..||||||||||||||:|||||||||||||||||||||||||.|||:
NP_000760.1        1 MDPFVVLVLCLSCLLLLSIWRQSSGRGKLPPGPTPLPVIGNILQIDIKDV     50

NP_000762.2       51 SKSLTNLSKVYGPVFTLYFGLKPIVVLHGYEAVKEALIDLGEEFSGRGIF    100
                     |||||||||:||||||||||:.:|||||||.|||||||||||||||.|
NP_000760.1       51 SKSLTNLSKIYGPVFTLYFGLERMVVLHGYEVVKEALIDLGEEFSGRGHF    100

NP_000762.2      101 PLAERANRGFGIVFSNGKKWKEIRRFSLMTLRNFGMGKRSIEDRVQEEAR    150
                     ||||||||||||||||:|||||||||||||||||||||||||||||||||
NP_000760.1      101 PLAERANRGFGIVFSNGKRWKEIRRFSLMTLRNFGMGKRSIEDRVQEEAR    150

NP_000762.2      151 CLVEELRKTKASPCDPTFILGCAPCNVICSIIFHKRFDYKDQQFLNLMEK    200
                     ||||||||||||||||||||||||||||||||.|||||||||||||||||
NP_000760.1      151 CLVEELRKTKASPCDPTFILGCAPCNVICSIIFQKRFDYKDQQFLNLMEK    200

NP_000762.2      201 LNENIKILSSPWIQICNNFSPIIDYFPGTHNKLLKNVAFMKSYILEKVKE    250
                     |||||:|:|:|||||||||..|||||||||||||||:|||:|.|||||||
NP_000760.1      201 LNENIRIVSTPWIQICNNFPTIIDYFPGTHNKLLKNLAFMESDILEKVKE    250

NP_000762.2      251 HQESMDMNNPQDFIDCFLMKMEKEKHNQPSEFTIESLENTAVDLFGAGTE    300
                     |||||||:|||:|||||||:||||||.||.|||||:|..||.||.|||||
NP_000760.1      251 HQESMDINNPRDFIDCFLIKMEKEKQNQQSEFTIENLVITAADLLGAGTE    300

NP_000762.2      301 TTSTTLRYALLLLLKHPEVTAKVQEEIERVIGRNRSPCMQDRSHMPYTDA    350
                     ||||||||||||||||||||||||||||||||||||||||.|||||||||
NP_000760.1      301 TTSTTLRYALLLLLKHPEVTAKVQEEIERVIGRNRSPCMQDRGHMPYTDA    350

NP_000762.2      351 VVHEVQRYIDLLPTSLPHAVTCDIKFRNYLIPKGTTILISLTSVLHDNKE    400
                     ||||||||||:||||||||||:|||||||||||||||.|||||||||||
NP_000760.1      351 VVHEVQRYIDLIPTSLPHAVTCDVKFRNYLIPKGTTILTSLTSVLHDNKE    400

NP_000762.2      401 FPNPEMFDPHHFLDEGGNFKKSKYFMPFSAGKRICVGEALAGMELFLFLT    450
                     |||||||||.|||||||||||.|||||||||||||||.||.||||||||
NP_000760.1      401 FPNPEMFDPRHFLDEGGNFKKSNYFMPFSAGKRICVGEGLARMELFLFLT    450

NP_000762.2      451 SILQNFNLKSLVDPKNLDTTPVVNGFASVPPFYQLCFIPV    490
                     .||||||||||:|||:|||||||||||||||||||||||
NP_000760.1      451 FILQNFNLKSLIDPKDLDTTPVVNGFASVPPFYQLCFIPV    490
```

CYP3A4

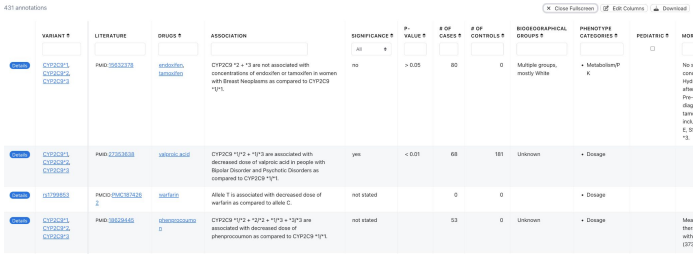
```
# Aligned_sequences: 2
# 1: NP_000762.2
# 2: AAF21034.1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 495
# Identity:     141/495 (28.5%)
# Similarity:   212/495 (42.8%)
# Gaps:          83/495 (16.8%)
# Score: 385.5
#
#
#=======================================

NP_000762.2        7 LVLCLSCLLLLSLWRQSSGRGKL--PPGPTPLPVIGNILQIGIKDISKSL     54
                     |:|.:|.:||......|.|..|.  .|||||||.:||||... |......
AAF21034.1        13 LLLAVSLVLLYLYGTHSHGLFKKLGIPGPTPLPFLGNILSYH-KGFCMFD     61

NP_000762.2       55 TNLSKVYGPVFTLYFGLKPIVVLHGYEAVKEALI-DLGEEFSGRGIF-PL    102
                     ....|.||.|:..|.|.:|::.:...:.:|..|: :....|:.|..| |:
AAF21034.1        62 MECHKKYGKVWGFYDGQQPVLAITDPDMIKTVLVKECYSVFTNRRPFGPV    111

NP_000762.2      103 AERANRGF---GIVFSNGKKWKEIRRFSLMTLRNFGMGK-RSIEDRVQEE    148
                              ||   .|..:..::||.:|  ||:: ..|..|| :.:....:.:.
AAF21034.1       112 ------GFMKSAISIAEDEEWKRLR--SLLS-PTFTSGKLKEMVPIIAQY    152

NP_000762.2      149 ARCLVEELRK--TKASPCDPTFILGCAPCNVICSIIFH------------    184
                     ...||..||:  ....|.....:.|....:||..|.|.
AAF21034.1       153 GDVLVRNLRREAETGKPVTLKDVFGAYSMDVITSTSFGVNIDSLNNPQDP    202

NP_000762.2      185 --------KRFDYKDQQFLNLMEKLNENIKILSSPWIQICNNFSPIIDYF    226
                              .|||:.|..||::       .:.....|.:::.|     |..|
AAF21034.1       203 FVENTKKLLRFDFLDPFFLSI------TVFPFLIPILEVLN-----ICVF    241

NP_000762.2      227 P-GTHNKLLKNVAFMKSYILEKVKEHQESMDMNNPQDFIDCFL---MKME    272
                     | ...|.|.|:|..||...||..::|:        .||:...:   ...|
AAF21034.1       242 PREVTNFLRKSVKRMKESRLEDTQKHR--------VDFLQLMIDSQNSKE    283

NP_000762.2      273 KEKHNQPSEFTIESLENTAVDLFGAGTETTSTTLRYALLLLLKHPEVTAK    322
                     .|.|...|:  :|.:...:.:.:| ||.||||:.|.:.:..|..||:|..|
AAF21034.1       284 TESHKALSD--LELVAQSIIFIF-AGYETTSSVLSFIMYELATHPDVQQK    330

NP_000762.2      323 VQEEIERVIGRNRSPCMQDRSHMPYTDAVVHEVQRYIDLLPTSLPHAVTC    372
                     :||||:.|:.....|.......|.|.|.||:|..|   |.|.::.....|
AAF21034.1       331 LQEEIDAVLPNKAPPTYDTVLQMEYLDMVVNETLR---LFPIAMRLERVC    377

NP_000762.2      373 --DIKFRNYLIPKGTTILISLTSVLHDNKEFPNPEMFDPHHFLDEGGNFK    420
                       |:.....||||..:.:|..:.:.|.|.:..||.|.|..|      .|
AAF21034.1       378 KKDVEINGMFIPKGVVVMIPSYALHRDPKYWTEPEKFLPERF------SK    421

NP_000762.2      421 KSK------YFMPFSAGKRICVGEALAGMELFLFLTSILQNFNLK    459
                     |:|      .:.||.:|.|.|:|...|.|.:.|.|..:|||||:|
AAF21034.1       422 KNKDNIDPYIYTPFGSGPRNCIGMRFALMNMKLALIRVLQNFSFK    466
```

Results (Finding PharmGKB Variants When Searching Celecoxib + CYP2C9)

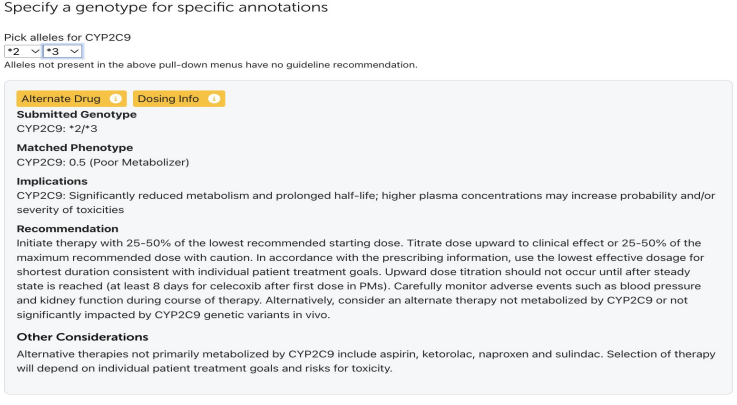


These clinical annotations were in Level 1A, and the most prevalent in the research was CYP2C9*2 and CYP2C9*3, which is why those variants were further analyzed, especially in the category of metabolism.

This table showcases Variant Annotations, showing them in more detail.



Prescribing Info Section, showcasing carriers of CYP2C9*2 and CYP2C9*3 as poor metabolizers.



Results (SNP Database Nucleotide and Amino Acid Changes) and collated Information from PharmGKB

Organism: Homo sapiens (humans)
Variant type: Single Nucleotide Variation

CYP2C9*2 (rs1799853)

Alleles: C>A / C>T
Frequency: T=0.085923 (22743/264690, TOPMED), T=0.092016 (23117/251228, GnomAD_exome), T=0.089751 (12575/140110, GnomAD) (+ 22 more)

Genomic Placements

Sequence name	Change
CYP2C9 RefSeqGene (LRG_1195)	NG_008385.2:g.9133C>A
CYP2C9 RefSeqGene (LRG_1195)	NG_008385.2:g.9133C>T
GRCh37.p13 chr 10	NC_000010.10:g.96702047C>A
GRCh37.p13 chr 10	NC_000010.10:g.96702047C>T
GRCh38.p14 chr 10	NC_000010.11:g.94942290C>A
GRCh38.p14 chr 10	NC_000010.11:g.94942290C>T

Gene: CYP2C9, cytochrome P450 family 2 subfamily C member 9 (plus strand)

Molecule type	Change	Amino acid[Codon]	SO Term
CYP2C9 transcript	NM_000771.4:c.430C>A	R [CGT] > S [AGT]	Coding Sequence Variant
CYP2C9 transcript	NM_000771.4:c.430C>T	R [CGT] > C [TGT]	Coding Sequence Variant
cytochrome P450 2C9	NP_000762.2:p.Arg144Ser	R (Arg) > S (Ser)	Missense Variant
cytochrome P450 2C9	NP_000762.2:p.Arg144Cys	R (Arg) > C (Cys)	Missense Variant

CYP2C9*3 (rs1057910)

Alleles: A>C / A>G
Frequency: C=0.064700 (19529/301838, ALFA), C=0.063706 (7725/121260, ExAC), C=0.02428 (686/28258, 14KJPN) (+ 22 more)

Genomic Placements

Sequence name	Change
CYP2C9 RefSeqGene (LRG_1195)	NG_008385.2:g.48139A>C
CYP2C9 RefSeqGene (LRG_1195)	NG_008385.2:g.48139A>G
GRCh37.p13 chr 10	NC_000010.10:g.96741053A>C
GRCh37.p13 chr 10	NC_000010.10:g.96741053A>G
GRCh38.p14 chr 10	NC_000010.11:g.94981296A>C
GRCh38.p14 chr 10	NC_000010.11:g.94981296A>G

Gene: CYP2C9, cytochrome P450 family 2 subfamily C member 9 (plus strand)

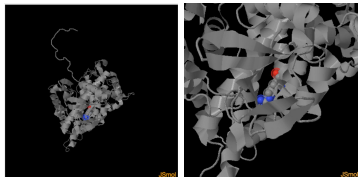
Molecule type	Change	Amino acid[Codon]	SO Term
CYP2C9 transcript	NM_000771.4:c.1075A>C	I [ATT] > L [CTT]	Coding Sequence Variant
CYP2C9 transcript	NM_000771.4:c.1075A>G	I [ATT] > V [GTT]	Coding Sequence Variant
cytochrome P450 2C9	NP_000762.2:p.Ile359Leu	I (Ile) > L (Leu)	Missense Variant
cytochrome P450 2C9	NP_000762.2:p.Ile359Val	I (Ile) > V (Val)	Missense Variant

Feature	CYP2C9	CYP2C9*2 (rs1799853)	CYP2C9*3 (rs1057910)
Enzyme Activity	Normal (100%)	Reduced (60-70%)	Severely reduced (10-30%)
Metabolism of Celecoxib	Efficient metabolism	Slower metabolism → Higher celecoxib plasma levels	Very slow metabolism → Significantly increased celecoxib plasma levels
Celecoxib Clearance Rate	Normal clearance	~50% reduced clearance	~75-90% reduced clearance
Impact on Arthritis Treatment	Normal anti-inflammatory effect with standard celecoxib dose	Increased risk of side effects (e.g., GI toxicity, cardiovascular risk) due to slower clearance	Severely increased risk of celecoxib-induced side effects, requires lower dose or alternative drug
Recommended Celecoxib Dose Adjustments	Standard dose (e.g., 200 mg/day)	Consider reducing dose (e.g., 25-50% of standard dose)	Consider reducing dose (e.g., 25-50% of standard dose)
Adverse Effects Risk	Normal risk	Higher risk of gastrointestinal (GI) bleeding and cardiovascular events	Very high risk of celecoxib-induced side effects, especially CV and GI toxicity
Therapeutic Monitoring Needed?	No special monitoring	Yes, monitor drug levels, side effects	Yes, requires frequent monitoring

Results (CYP2C9*2 (rs1799853) negative effect likelihood) and Population Samples

CYP2C9*2 was analyzed using SIFT, PolyPhen-2, and MutationTaster to predict mutation impact, with allele frequencies assessed using gnomAD (4,000 genomes, 185,242 samples, 11,971 homozygotes).

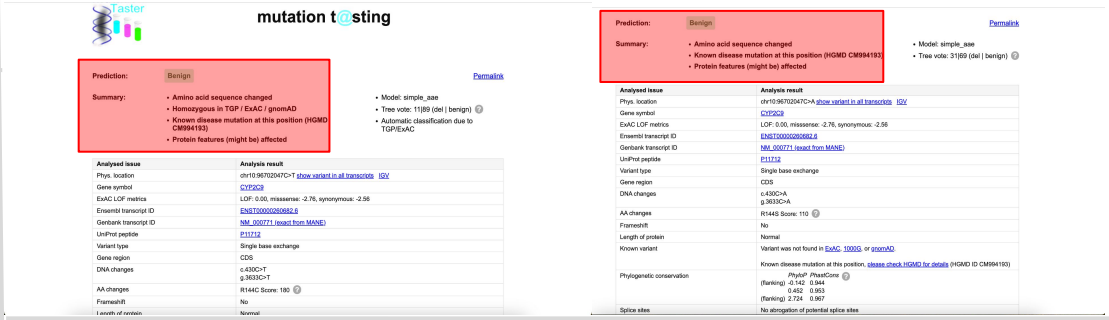
Structural changes



SIFT

SNP	ORGANISM/BUILD	CHR	COORDINATE	REF ALLELE	ALT ALLELE	AMINO ACID CHANGE	GENE NAME	GENE ID	TRANSCRIPT ID	PROTEIN ID	REGION	SIFT SCORE	SIFT MEDIAN	NO OF SEQS AT POSITION	SIFT PREDICTION	AVG ALLELE FREQ	EAS ALLELE FREQ	AMR ALLELE FREQ	AFR ALLELE FREQ	EUR ALLELE FREQ	SAS ALLELE FREQ
rs1799853	Homo_sapiens/GRCh37.74	10	96702047	C	T	R144C	CYP2C9	ENSG00000138109	ENST00000260682	ENSP00000260682	CDS	0.019	2.48	309	DELETERIOUS	0.048	0.001	0.099	0.008	0.124	0.035

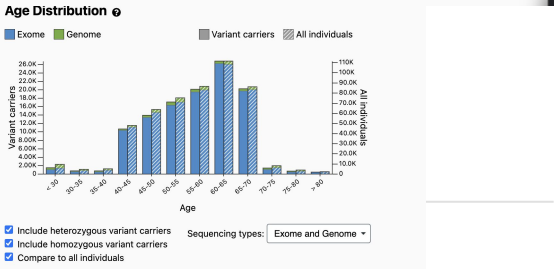
MutationTaster



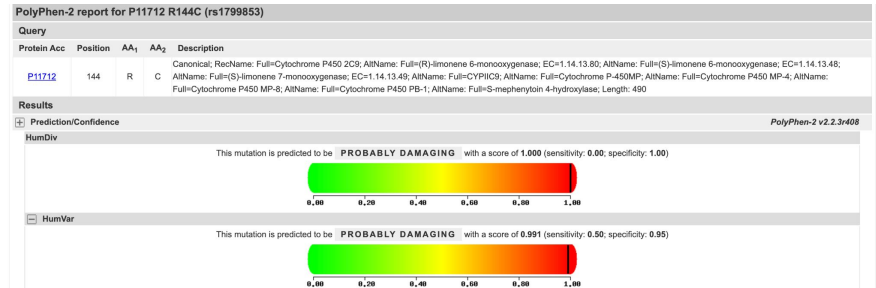
gnomAD

gnomAD | HGDP | 1KG | Local Ancestry

Genetic Ancestry Group	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
Amish	139	912	12	0.1524
Ashkenazi Jewish	4069	29596	280	0.1375
European (non-Finnish)	155585	1179834	10435	0.1319
Middle Eastern	734	6058	41	0.1212
European (Finnish)	7312	64006	416	0.1142
Remaining	6626	62476	410	0.1061
Admixed American	4720	59962	205	0.07872
South Asian	4417	91074	153	0.04850
African/African American	1629	74996	19	0.02172
East Asian	11	44858	0	0.0002452
XX	94366	812258	6145	0.1162
XY	90876	801514	5826	0.1134
Total	**185242**	**1613772**	**11971**	**0.1148**



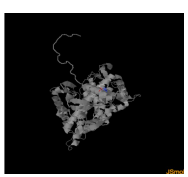
PolyPhen-2

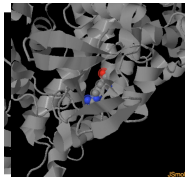


Results (CYP2C9*3 (rs1057910) negative effect likelihood) and Population Samples

CYP2C9*3 was analyzed using SIFT, PolyPhen-2, and MutationTaster to predict mutation impact, with allele frequencies assessed using gnomAD (4,000 genomes, 102,334 samples, 3,705 homozygotes).

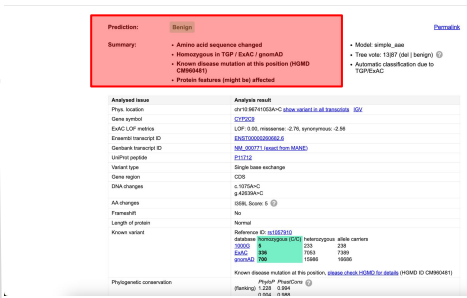
Structural changes

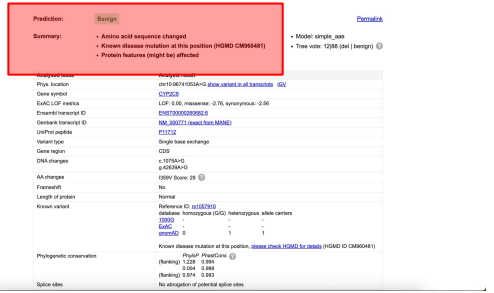




SIFT

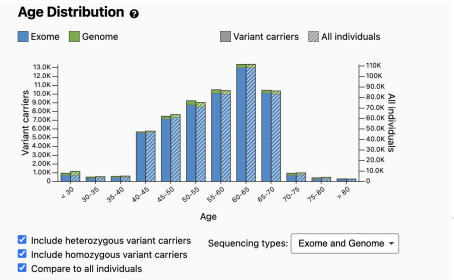
SNP	ORGANISM/BUILD	CHR	COORDINATE	REF ALLELE	ALT ALLELE	AMINO ACID CHANGE	GENE NAME	GENE ID	TRANSCRIPT ID	PROTEIN ID	REGION	SIFT SCORE	SIFT MEDIAN	NO OF SEQS AT POSITION	SIFT PREDICTION	AVG ALLELE FREQ	EAS ALLELE FREQ	AMR ALLELE FREQ	AFR ALLELE FREQ	EUR ALLELE FREQ	SAS ALLELE FREQ
rs1057910	Homo_sapiens/GRCh37.74	10	96741053	A	C	I359L	CYP2C9	ENSG00000138109	ENST00000260682	ENSP00000260682	CDS	0.362	2.48	312	TOLERATED	0.049	0.034	0.037	0.002	0.073	0.109



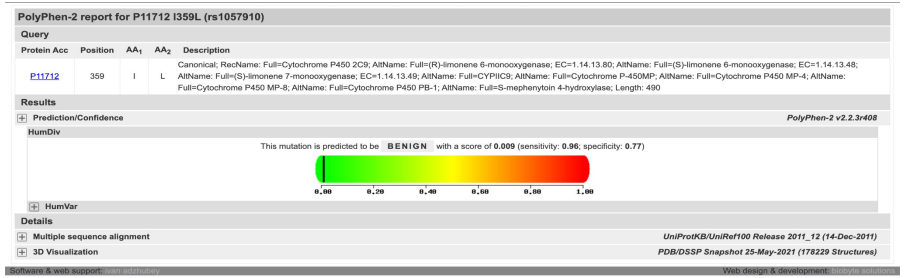


gnomAD

Genetic Ancestry Group	Allele Count	Allele Number	Number of Homozygotes	Allele Frequency
South Asian	10029	91086	605	0.1101
Ashkenazi Jewish	2463	29596	112	0.08322
Middle Eastern	441	6058	26	0.07280
European (non-Finnish)	76626	1179884	2596	0.06494
Remaining	3990	62492	136	0.06385
European (Finnish)	3941	64020	143	0.06156
Admixed American	2516	59948	62	0.04197
East Asian	1355	44870	19	0.03020
Amish	27	912	0	0.02961
African/African American	946	75024	6	0.01261
XX	50492	812328	1790	0.06216
XY	51842	801562	1915	0.06468
Total	**102334**	**1613890**	**3705**	**0.06341**



PolyPhen-2



Discussions (Analysis of Results)

Similar Genes & Variant Impact:

- **Gene Similarity:** CYP2C19 showed high similarity to CYP2C9 in EMBOSS Water, while VKORC1 and CYP3A4 had moderate similarity. These genes likely assist CYP2C9 in Celecoxib metabolism, reinforcing CYP3A4's role, and predicting VKORC1 and CYP2C19's roles.
- **Variant Analysis:** CYP2C9*2 appeared more damaging based on SIFT and PolyPhen-2, while CYP2C9*3, due to being comparatively rarer in population frequencies, likely requires tailored treatment due to its stronger metabolic impact.
    - **Clinical Relevance:** Reduced Celecoxib metabolism in individuals with these variants poses treatment risks, emphasizing the need for genetic consideration.

Key Implications:

✔ Dose adjustments based on genotype
✔ Enables targeted treatment strategies
✔ Supports precision medicine
✔ Enhances personalized dosing

Future Directions:

- Study Celecoxib-CYP2C9 interactions using molecular docking
- Analyze variant effects on metabolism
- Integrate genetic & pharmacokinetic data and improve personalized therapy

Bibliography & Sources

Bleasby, A. (1999). *EMBOSS: water manual*. httpppsss://wwww.bioinformatics.nl/cgi-bin/emboss/help/water

gnomAD. (n.d.). *SNV : 10-94942290 - C-T (GRCh38)*. httpppsss://gnomad.broadinstitute.org/variant/10-94942290-C-T?dataset=gnomad_r4

gnomAD. (n.d.). SNV:10-94981296-A-C(GRCh38)

httpppsss://gnomad.broadinstitute.org/variant/10-94981296-A-C?dataset=gnomad_r4

Gong, L., Thorn, C., Bertagnolli, M., Grosser, T., Altman, R., & Klien, T. (2012, April). *Celecoxib Pathway, Pharmacodynamics*. PharmGKB.

httpppsss://wwww.pharmgkb.org/literature/14988371

MutationTaster. (n.d.). *MutationTaster - documentation*. httpppsss://wwww.mutationtaster.org/info/documentation.html

National Human Genome Research Institute. (2025). *Bioinformatics*. National Human Genome Research Institute.

httpppsss://wwww.genome.gov/genetics-glossary/Bioinformatics

NCBI. (n.d.). *dbSNP (single nucleotide polymorphism database)*. NCBI. httpppsss://wwww.ncbi.nlm.nih.gov/snp/

PolyPhen-2. (n.d.). PolyPhen-2: prediction of functional effects of human nsSNPs. httppp://genetics.bwh.harvard.edu/pph2/

PubChem. (2025). *Celecoxib | C17H14F3N3O2S | CID 2662 - PubChem*. PubChem.

httpppsss://pubchem.ncbi.nlm.nih.gov/compound/2662#section=3D-Conformer

Sorting Intolerant From Tolerant. (n.d.). SIFT - Predict effects of nonsynonmous / missense variants. httpppsss://sift.bii.a-star.edu.sg/

STRING. (n.d.). STRING: functional protein association networks. httpppsss://string-db.org/