

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	<ul style="list-style-type: none">● 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)	<ul style="list-style-type: none">● 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	<ul style="list-style-type: none">● 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	<ul style="list-style-type: none">● SNP analysis revealed <i>CYP2C9</i>*2 and <i>CYP2C9</i>*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.
	<ul style="list-style-type: none">● Aims to uncover genetic factors influencing drug metabolism for better treatment strategies.

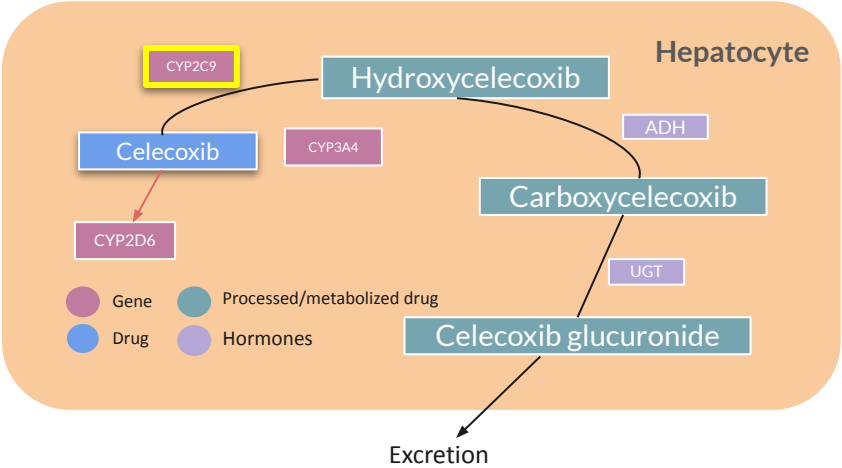
Celecoxib Overview

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.

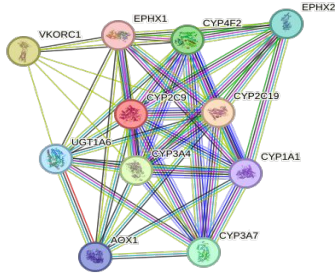


Drug Advantages	Potential Side Effects	Precautions	Undiscovered Genetic Factors
<p>Arthritis Relief: Improves joint function, reduces stiffness, lowers GI risks vs. NSAIDs.</p> <p>Benefits: Fewer stomach issues, safer for ulcers, minimal blood clotting impact.</p>	<p>Risks: High doses/long-term use may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

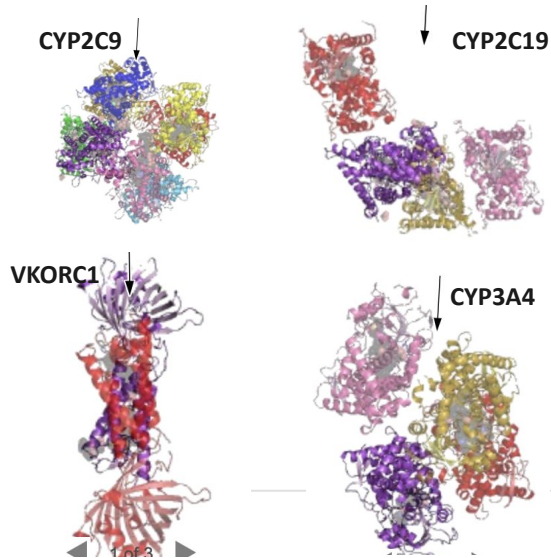
Methodology

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.



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 regulator [...] (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

- **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	GPVFTLYFGKPLVILVHGVEAVK-----EALIDLGEFSGRGIFPLAER	105
VKOR1_HUMAN	19	GLVLVSLY-----ALH-VKAARARDRDYRALCDVGTATCSRVF--SSR	58
NP_000762.2	106	ANRGFGIVFSNGKKWKEIRFSLNTLRNFGMGKRSIEDRVQEEARCLVEE	155
VKOR1_HUMAN	59	WGRGFGLV-----EHLVGDQSLNQSNIGFCIFYT	89
NP_000762.2	156	LRKTKASPCDPTFILGC	172
VKOR1_HUMAN	90	LQ-----LLLG	96

$\frac{1}{2}$ -----
 $\frac{1}{2}$ -----

CYP2C19

```
# Aligned_sequences: 2
# 1: NP_000762.2
# 2: NP_000760.1
# Matrix: EBLSUM62
# 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	MSDLVVLVLCLSCLLLLSIWRQSSGRGKLPGPPTPLPVIGNLIQTIGKDI	50
NP_000760.1	1	MDPFVVLVLCLSCLLLLSIWRQSSGRGKLPGPPTPLPVIGNLIQDIDKDV	50
NP_000762.2	51	SKSLTNLSKVYGPVFTLYFGLKPIVLHGYEAVKEALIDLGEFESSGRGIF	100
NP_000760.1	51	SKSLTNLSKVYGPVFTLYFGLERMVLHGYEVKEALIDLGEFESSGRGHF	100
NP_000762.2	101	PLAERANRGGFVFSNGKKWKEIRFSLMTLRNFMGMRKSIEDRVQEEAR	150
NP_000760.1	101	PLAERANRGGFVFSNGKRWKEIRFSLMTLRNFMGMRKSIEDRVQEEAR	150
NP_000762.2	151	CLVEELRKTASPDCPTILGCAPCNVICSIFHRFDDYKQDQFLNLMKEK	200
NP_000760.1	151	CLVEELRKTASPDCPTILGCAPCNVICSIFQRFDDYKQDQFLNLMKEK	200
NP_000762.2	201	LNENIKILSSPIWIGICNFFSPIIDYFGPTHNKKLKNVAFMKSIYLEKVK	250
NP_000760.1	201	LNENIRIVSTPVIQICNFFPTIIDYFGPTHNKKLKNLAFMESDILEKVK	250
NP_000762.2	251	HQESMDMMNPQDFIDCLFMKEKEHNQSEFFIESLENTAVDLFGAGTE	300
NP_000760.1	251	HQESMDINPPODFIDCLFMKEKEQNQSEFFIENLVITAADLLGAGTE	300
NP_000762.2	301	TTSTTLRYALLLKHPEVTKAQVEEIVGRNRSPCMQDRSHMPYDTA	350
NP_000760.1	301	TTSTTLRYALLLKHPEVTKAQVEEIVGRNRSPCMQDRSHMPYDTA	350
NP_000762.2	351	VVHEVQRYIDLPTSLPHAVTCDFKFRNYLPKGTTILSTLSVLDHONKE	400
NP_000760.1	351	VVHEVQRYIDLPTSLPHAVTCDFKFRNYLPKGTTILSTLSVLDHONKE	400
NP_000762.2	401	FPNPEMFDPHFLDEGGNFKSKSYFMPFSAGKRICVGEALAGMLFLFLT	450
NP_000760.1	401	FPNPEMFDPHFLDEGGNFKSKSYFMPFSAGKRICVGEALAGMLFLFLT	450
NP_000762.2	451	SILQNFNLKSLDVPKNLDTTPVNVNGFASVPFYQLCFIPV	490
NP_000760.1	451	FILQNFNLKSLDIPKDLDTTPVNVNGFASVPFYQLCFIPV	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
```

4

NP_000762.2		LVLGSLCLLLSLNWRQSSGRGKL--PPGPTPLVPIGNTLIGIKDTSKSL	54
AAF21034.1	13	LLAVSLVLLVLYGTHSHGLFKKLGIPGTPPLFLGNLSYH-KGFCMFD	61
NP_000762.2	55	TNLSKVYGVPTFLYFLGKPTVVLHGYEAVKCALI-DLGEFSSGRGIF-PL	102
AAF21034.1	62	MECHKYGVKVGFDYQGPVLAITDPDMKTVLVKECVSYFTNRPPGFPV	111
NP_000762.2	103	AERANRGF-----GIVFSNGKKWKEIRRFSLMTLRFNGMGK-RSIEDRVQEE	148
AAF21034.1	112	-----GFMKSAISIAEDEEWKRLR--SLLS-PTFTSGKLEKMPPIAQY	152
NP_000762.2	149	ARCLVEELRK--TKASPCDPTFILGCAPCNVICSIIFH-----	184
AAF21034.1	153	GDVLVRNLRRAEATGKPYTLKDVFGYSMDVITSTSGFVNIDSLNPNQP	202
NP_000762.2	185	-----KRFDYKQDQFLNLNMEKLNENKILSSPWIQICNMFSPIDIYF	226
AAF21034.1	203	FVENTKLLRFDFLDPPFLSI-----TVFFLPITPLEVLN-----TCVF	241
NP_000762.2	227	P-GTHNKLKLVAFMKSYILEKVKHEQESMDMMNPQDFIDCLF-----MKME	272
AAF21034.1	242	PREVTNFLRKSVKRMKESRLDEQKHRR-----VDFQLQIDQSNSKE	283
NP_000762.2	273	KEKNQPSFEFTISELNTAVDLFGAGTETTSTTLRYALLLLKHPVETAK	322
AAF21034.1	284	TESHKALSD--LELVQAISIFIF-AGVETYSVLSIYELATLPPDVQVK	330
NP_000762.2	303	VOGEIERVIGRQSCPDQSRHMPYDVAHVVEQRYDLDLPTSLPHVATC	372
AAF21034.1	331	LQEEIDAVLPNKAPPTYDTVLQMEYLDVMVNETLR--LPTIAMRLERC	377
NP_000762.2	373	--DIKFRNLPYKGTLLISLTSVLHONKEKFPNMFDPHFLDEGGNFK	420
AAF21034.1	378	KKDVEINGFMIPKGVVMIPSVLHRDPKYMEPEKFLPERF-----SK	421
NP_000762.2	422	KSQ-----YFPMFSAGKRRCVGEALAGMELFLFLSTLQNFNLK	459
AAF21034.1	422	KNKONIDPIYVTPFGSGPRNCIGRNFALNMNKLALIRVLQNFSEFK	466

Results (Finding PharmGKB Variants When Searching Celecoxib + CYP2C9)

Overview

Prescribing Info

Drug Label Annotations

Clinical Annotations

Variant Annotations

Literature

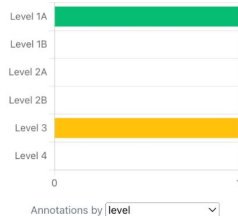
Pathways

Clinical Annotations

PharmGKB clinical annotations provide information about variant-drug pairs based primarily on variant annotations and incorporating variant-specific prescribing guidance from clinical guidelines and FDA-approved drug labels, when available. Curators manually review annotations and create genotype-based summaries describing the phenotypic impact of the variant. Each clinical annotation is assigned a [Level of Evidence](#), which is generally informed by the clinical annotation's [score](#).

Note: Alleles in PharmGKB are mapped to the positive chromosomal strand. Therefore, variants in genes on the "minus" strand (eg. VKORC1) are complemented in PharmGKB annotations.

Focus on Pediatrics ☐



[Read more about Clinical Annotations](#)

[Read more about Variant Annotations](#)

2 annotations

[Fullscreen](#) [Edit Columns](#) [Show Filters](#) [Download](#)

LEVEL	VARIANT	GENE	DRUGS	PHENOTYPE CATEGORIES	PHENOTYPE	PEDIATRI
Details	Level 1A	CYP2C9*1, CYP2C9*2, CYP2C9*3, CYP2C9*13	celecoxib	Metabolism/PK		<input checked="" type="checkbox"/>
Details	Level 3	rs1057910	Antiinflammatory agents, non-steroids, celecoxib, diclofenac	Toxicity		

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.

431 annotations											
VARIANT	LITERATURE	DRUGS	ASSOCIATION	SIGNIFICANCE	P-VALUE	# OF CASES	# OF CONTROLS	BIOGEOGRAPHICAL GROUPS	PHENOTYPE CATEGORIES	PEDIATRIC	MOI
CYP2C9*1, CYP2C9*2, CYP2C9*3	PMID: 16824278	celecoxib, ibuprofen	CYP2C9*1, *2 are not associated with concentrations of arthroic or transthoric in women with breast Neoplasms as compared to CYP2C9*1*1.	no	> 0.05	80	0	Multiple groups, mostly white		<input checked="" type="checkbox"/>	No covariates after adjustment
CYP2C9*1, CYP2C9*2, CYP2C9*3	PMID: 22538368	celecoxib, aspirin	CYP2C9*1*2 + *1*3 are associated with decreased dose of celecoxib in people with Basilar Disorder and Pharyngeal Disorders as compared to CYP2C9*1*1.	yes	< 0.01	68	181	Unknown		<input checked="" type="checkbox"/>	Storage
CYP2C9*1, CYP2C9*2, CYP2C9*3	PMID: 24636342	celecoxib	Allele 1 is associated with decreased dose of celecoxib as compared to allele 2.	not stated		0	0			<input checked="" type="checkbox"/>	Storage
CYP2C9*1, CYP2C9*2, CYP2C9*3	PMID: 26636465	celecoxib, diclofenac	CYP2C9*1*2 + *1*3 + *1*3*3 are associated with decreased dose of phenprocoumon as compared to CYP2C9*1*1.	not stated		53	0	Unknown		<input checked="" type="checkbox"/>	Storage

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

Specify a genotype for specific annotations

Pick alleles for CYP2C9

*2 *3

Alleles not present in the above pull-down menus have no guideline recommendation.

[Alternate Drug](#) [Dosing Info](#)

Submitted Genotype

CYP2C9: *2/*3

Matched Phenotype

CYP2C9: 0.5 (Poor Metabolizer)

Implications

CYP2C9: Significantly reduced metabolism and prolonged half-life; higher plasma concentrations may increase probability and/or severity of toxicities

Recommendation

Initiate therapy with 25-50% of the lowest recommended starting dose. Titrate dose upward to clinical effect or 25-50% of the maximum recommended dose with caution. In accordance with the prescribing information, use the lowest effective dosage for shortest duration consistent with individual patient treatment goals. Upward dose titration should not occur until after steady state is reached (at least 8 days for celecoxib after first dose in PMs). Carefully monitor adverse events such as blood pressure and kidney function during course of therapy. Alternatively, consider an alternate therapy not metabolized by CYP2C9 or not significantly impacted by CYP2C9 genetic variants in vivo.

Other Considerations

Alternative therapies not primarily metabolized by CYP2C9 include aspirin, ketorolac, naproxen and sulindac. Selection of therapy will depend on individual patient treatment goals and risks for toxicity.

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	

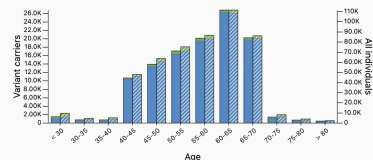
Gene: CYP2C9, cytochrome P450 family 2 subfamily C member 9 (plus strand)			
Molecule type	Change	Amino acid[Codon]	SO Term
CYP2C9 transcript	NM_000771.4:c.1075A>C	I [ATT] > L [CTT]	Coding Sequence Variant
CYP2C9 transcript	NM_000771.4:c.1075A>G	I [ATT] > V [GTT]	Coding Sequence Variant
cytochrome P450 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

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).


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	0.1182	
XY	90676	801514	5826	0.1134	
Total	185242	1613772	11971	0.1148	

■ Exome ■ Genome ■ Variant carriers ■ All individuals



Sequencing types: Exome and Genome

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	14	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



mutation t*o*sting

Prediction: Benign

Summary:

- Amino acid sequence changed
- Homozygous in TGP / EAAC / gnomAD
- Known disease mutations at this position (HGMD CM984193)
- Protein features (might be) affected

Amino acid sequence change

chr19:96702047C>T [show variant in all transcripts](#) [VDV](#)

C>T(C>G)

ExAC LOF metrics
LOF -0.00, missense -2.76, synonymous -2.56

gnomAD missense rate
[ENR13000202662.6](#)

Gerbark transcript ID
[NM_000771 \(spliced from MANE\)](#)

Unifref protein
[P11312](#)

Variant type
Single base exchange

DNA changes
c.430C>T
g.3853C>T

AA changes
R144C Score: 110 [ⓘ](#)

Frameshift
No

Predictions

Benign

- Model: simple_aaa
- Tree vote: 11/69 (se) | benign [ⓘ](#)
- Automatic classification due to TGP/EAAC

Analytic issue

Phys. location
chr19:96702047C>T [show variant in all transcripts](#) [VDV](#)

Gene symbol
[CTSC2B](#)

ExAC LOF metrics
LOF -0.00, missense -2.76, synonymous -2.56

Ensembl transcript ID
[ENST00000202662.6](#)

Gerbark transcript ID
[NM_000771 \(spliced from MANE\)](#)

Unifref peptide
[P11312](#)

Variant type
Single base exchange

DNA changes
c.430C>T
g.3853C>T

AA changes
R144S Score: 110 [ⓘ](#)

Frameshift
No

Length of protein
Unknown variant

Phylogenetic conservation
[PhyloP: PhastCons \[ⓘ\]\(#\)](#)
(Rating: 4.42 | 0.144)
0.452 | 0.953
(Rating: 2.75 | 0.987)

Splice sites
No disruption of potential splice sites

Prediction: Benign

Summary:

- Amino acid sequence changed
- Known disease mutation at this position (HGMD CM984193)
- Protein features (might be) affected

Analytic issue

Phys. location
chr19:96702047C>T [show variant in all transcripts](#) [VDV](#)

Gene symbol
[CTSC2B](#)

ExAC LOF metrics
LOF -0.00, missense -2.76, synonymous -2.56

Ensembl transcript ID
[ENST00000202662.6](#)

Gerbark transcript ID
[NM_000771 \(spliced from MANE\)](#)

Unifref peptide
[P11312](#)

Variant type
Single base exchange

DNA changes
c.430C>T
g.3853C>T

AA changes
R144S Score: 110 [ⓘ](#)

Frameshift
No

Length of protein
Unknown variant

Phylogenetic conservation
[PhyloP: PhastCons \[ⓘ\]\(#\)](#)
(Rating: 4.42 | 0.144)
0.452 | 0.953
(Rating: 2.75 | 0.987)

Splice sites
No disruption of potential splice sites

Prediction: Benign

Summary:

- Amino acid sequence changed
- Known disease mutation at this position (HGMD CM984193)
- Protein features (might be) affected

Analytic issue

Phys. location
chr19:96702047C>T [show variant in all transcripts](#) [VDV](#)

Gene symbol
[CTSC2B](#)

ExAC LOF metrics
LOF -0.00, missense -2.76, synonymous -2.56

Ensembl transcript ID
[ENST00000202662.6](#)

Gerbark transcript ID
[NM_000771 \(spliced from MANE\)](#)

Unifref peptide
[P11312](#)

Variant type
Single base exchange

DNA changes
c.430C>T
g.3853C>T

AA changes
R144C Score: 110 [ⓘ](#)

Frameshift
No

Length of protein
Unknown variant

Phylogenetic conservation
[PhyloP: PhastCons \[ⓘ\]\(#\)](#)
(Rating: 4.42 | 0.144)
0.452 | 0.953
(Rating: 2.75 | 0.987)

Splice sites
No disruption of potential splice sites

Prediction: Benign

Summary:

- Amino acid sequence changed
- Known disease mutation at this position (HGMD CM984193)
- Protein features (might be) affected

Analytic issue

Phys. location
chr19:96702047C>T [show variant in all transcripts](#) [VDV](#)

Gene symbol
[CTSC2B](#)

ExAC LOF metrics
LOF -0.00, missense -2.76, synonymous -2.56

Ensembl transcript ID
[ENST00000202662.6](#)

Gerbark transcript ID
[NM_000771 \(spliced from MANE\)](#)

Unifref peptide
[P11312](#)

Variant type
Single base exchange

DNA changes
c.430C>T
g.3853C>T

AA changes
R144C Score: 110 [ⓘ](#)

Frameshift
No

Length of protein
Unknown variant

Phylogenetic conservation
[PhyloP: PhastCons \[ⓘ\]\(#\)](#)
(Rating: 4.42 | 0.144)
0.452 | 0.953
(Rating: 2.75 | 0.987)

Splice sites
No disruption of potential splice sites

Prediction: Benign

Summary:

- Amino acid sequence changed
- Known disease mutation at this position (HGMD CM984193)
- Protein features (might be) affected

Analytic issue

Phys. location
chr19:96702047C>T [show variant in all transcripts](#) [VDV](#)

Gene symbol
[CTSC2B](#)

ExAC LOF metrics
LOF -0.00, missense -2.76, synonymous -2.56

Ensembl transcript ID
[ENST00000202662.6](#)

Gerbark transcript ID
[NM_000771 \(spliced from MANE\)](#)

Unifref peptide
[P11312](#)

Variant type
Single base exchange

DNA changes
c.430C>T
g.3853C>T

AA changes
R144C Score: 110 [ⓘ](#)

Frameshift
No

Length of protein
Unknown variant

Phylogenetic conservation
[PhyloP: PhastCons \[ⓘ\]\(#\)](#)
(Rating: 4.42 | 0.144)
0.452 | 0.953
(Rating: 2.75 | 0.987)

Splice sites
No disruption of potential splice sites

Prediction: Benign

Summary:

- Amino acid sequence changed
- Known disease mutation at this position (HGMD CM984193)
- Protein features (might be) affected

Analytic issue

Phys. location
chr19:96702047C>T [show variant in all transcripts](#) [VDV](#)

Gene symbol
[CTSC2B](#)

ExAC LOF metrics
LOF -0.00, missense -2.76, synonymous -2.56

Ensembl transcript ID
[ENST00000202662.6](#)

Gerbark transcript ID
[NM_000771 \(spliced from MANE\)](#)

Unifref peptide
[P11312](#)

Variant type
Single base exchange

DNA changes
c.430C>T
g.3853C>T

AA changes
R144C Score: 110 [ⓘ](#)

Frameshift
No

Length of protein
Unknown variant

Phylogenetic conservation
[PhyloP: PhastCons \[ⓘ\]\(#\)](#)
(Rating: 4.42 | 0.144)
0.452 | 0.953
(Rating: 2.75 | 0.987)

Splice sites
No disruption of potential splice sites

Prediction: Ben

PolyPhen-2

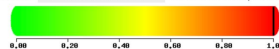
PolyPhen-2 report for P11712 R144C (rs1799853)

Query				
Protein Acc	Position	AA ₁	AA ₂	Description
P11712	144	R	C	Canonical; RecName: Full=Cytochrome P450 2C8; AltName: Full=(R)-limonene 6-monooxygenase; EC=1.14.13.80; AltName: Full=(S)-limonene 6-monooxygenase; EC=1.14.13.48 AltName: Full=(S)-limonene 7-monooxygenase; EC=1.14.13.49; AltName: Full=CYP1C3; AltName: Full=Cytochrome P450M; AltName: Full=Cytochrome P450 MP-4; AltName: Full=Cytochrome P450 MP-8; AltName: Full=Cytochrome P450 PB-1; AltName: Full=S-mephenytoin 4-hydroxylase; Length: 490

Results

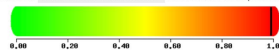
Prediction/Confidence PolyPhen-2 v2.2.3r408

This mutation is predicted to be **PROBABLY DAMAGING** with a score of 1.000 (sensitivity: 0.00; specificity: 1.00)



HumVar

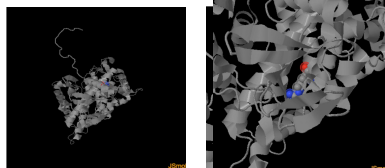
This mutation is predicted to be **PROBABLY DAMAGING** with a score of 0.991 (sensitivity: 0.50; specificity: 0.95)



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

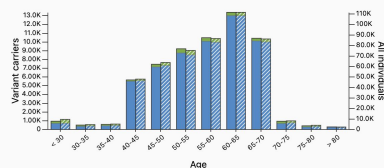
SNP	ORGANISM/BUILD	CHR	COORDINATE	REF ALLELE	ALT ALLELE	AMINO ACID CHANGE	GENE NAME	GENE ID	TRANSCRIPT ID	PROTEIN ID	REGION	SIFT SCORE	SIFT MEDIAN	NO OF SEQS AT POSITION	SIFT PREDICTION	AVG ALLELE FREQ	EAS ALLELE FREQ	AMR ALLELE FREQ	AFR ALLELE FREQ	EUR ALLELE FREQ	SAS ALLELE FREQ	
rs1057910	Homo sapiens	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
• Ashish	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

Age Distribution

Exome Genome Variant carriers All individuals



☒ Include heterozygous variant carriers
☒ Include homozygous variant carriers
☒ Compare to all individuals

Sequencing types: Exome and Genome

Prediction: Benign

Summary:

- Amino acid sequence changed
- Homozygous in TOP ExAC gnomAD
- Known disease mutation at this position (gnomAD C960441)
- Protein features (might be) affected

Analysis result:

- Model: simple_sse
- True vds: 1387 (del benign)
- Automatic classification due to TOP ExAC

Analysis result:

- Phys. location: cytochrome P450 CYP2C9
- Gene symbol: CYP2C9
- ExAC LOF metrics: LOF 0.00, missense 0.70, synonymous 0.30
| Exon/intron ID | ENSG00000138109 |
| Genebank transcript ID | NM_000771 (transcript from MIM) |
| UniProt protein | P11712 |
| Variant type | Single base exchange |
| Gene region | CDS |
| DNA changes | c.1059A>C |
| Alt changes | G96L, Score 1 |
| Frame shift | No |
| Length of protein | 359 |
| Known variant | Normal |
| Reference ID | rs1057910 |
| Reference | 10000 8 233 243 |
| ExAC | 1000 1000 1000 1000 |
| gnomAD | 1000 1000 1000 1000 |

Phylogenetic conservation:

- PhyloP: 0.000
- PhyloP: 0.000
- PhyloP: 0.000
- PhyloP: 0.000

Prediction: Benign

Summary:

- Amino acid sequence changed
- Known disease mutation at this position (gnomAD C960441)
- Protein features (might be) affected

Analysis result:

- Model: simple_sse
- True vds: 1387 (del benign)
- Automatic classification due to TOP ExAC

Analysis result:

- Phys. location: cytochrome P450 CYP2C9
- Gene symbol: CYP2C9
- ExAC LOF metrics: LOF 0.00, missense 0.70, synonymous 0.30
| Exon/intron ID | ENSG00000138109 |
| Genebank transcript ID | NM_000771 (transcript from MIM) |
| UniProt protein | P11712 |
| Variant type | Single base exchange |
| Gene region | CDS |
| DNA changes | c.1059A>C |
| Alt changes | G96L, Score 1 |
| Frame shift | No |
| Length of protein | 359 |
| Known variant | Normal |
| Reference ID | rs1057910 |
| Reference | 10000 8 233 243 |
| ExAC | 1000 1000 1000 1000 |
| gnomAD | 1000 1000 1000 1000 |

Phylogenetic conservation:

- PhyloP: 0.000
- PhyloP: 0.000
- PhyloP: 0.000
- PhyloP: 0.000

PolyPhen-2

PolyPhen-2 report for P11712 I359L (rs1057910)

Query

Protein	A1	A2	Description
P11712	359	I	L

Results

HumDiv

This mutation is predicted to be **BENIGN** with a score of 0.009 (sensitivity: 0.96; specificity: 0.77)

0.00 0.20 0.40 0.60 0.80 1.00

Details

- Multiple sequence alignment
- 3D Visualization

Software & web support: UniProtKB/UniRef100 Release 2011_12 (14-Dec-2011)
PDB/DSSP Snapshot 25-May-2021 (178229 Structures)
Web design & development: UniProt

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:** CYP2C92 *appeared more damaging based on SIFT and PolyPhen-2*, while CYP2C93, 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://www.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://www.pharmgkb.org/literature/14988371>
- MutationTaster. (n.d.). *MutationTaster - documentation*. <httpppsss://www.mutationtaster.org/info/documentation.html>
- National Human Genome Research Institute. (2025). *Bioinformatics*. National Human Genome Research Institute.
- <httpppsss://www.genome.gov/genetics-glossary/Bioinformatics>
- NCBI. (n.d.). *dbSNP (single nucleotide polymorphism database)*. NCBI. <httpppsss://www.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/>