The CYP2R1 Enzyme: Structure, Function, Enzymatic Properties and Genetic Polymorphism

Amelia Nathania Dong[1], Boon Hooi Tan[2], Yan Pan[3], Chin Eng Ong[4]

[1]School of Pharmacy, Monash University Malaysia, Bandar Sunway, Selangor, Malaysia; [2]Division of Applied Biomedical Sciences and Biotechnology, International Medical University, Bukit Jalil, Kuala Lumpur, Malaysia; [3]Department of Biomedical Science, University of Nottingham Malaysia Campus, Semenyih, Selangor, Malaysia; [4]School of Pharmacy, International Medical University, Bukit Jalil, Kuala Lumpur, Malaysia

Corresponding Author: Chin Eng Ong, School of Pharmacy, International Medical University, 126, Jalan Jalil Perkasa 19, Bukit Jalil, 57000 Kuala Lumpur, Malaysia; email: ceong98@hotmail.com

ABSTRACT -- Since the discovery of its role in vitamin D metabolism, significant progress has been made in the understanding of gene organisation, protein structure, catalytic function, and genetic polymorphism of cytochrome P450 2R1 (CYP2R1). Located on chromosome 11p15.2, *CYP2R1* possesses five exons, unlike most other CYP isoforms that carry nine exons. CYP2R1 crystal structure displays a fold pattern typical of a CYP protein, with 12 α-helices as its structural core, and β-sheets mostly arranged on one side, and the heme buried in the interior part of the protein. Overall, CYP2R1 structure adopts a closed conformation with the B′ helix serving as a gate covering the substrate access channel, with the substrate vitamin D_3 occupying a position with the side chain pointing toward the heme group. In liver, CYP2R1 25-hydroxylates vitamin D and serves as an important determinant of 25(OH)D level in the tissue and in circulation. While substrate profile has been well studied, inhibitor specificity for CYP2R1 requires further investigation. Both exonic and non-exonic single nucleotide polymorphisms (SNPs) have been reported in *CYP2R1*, including the *CYP2R1*2* carrying Leu99Pro exchange, and a number of non-exonic SNPs with variable functional consequences in gene regulation. A non-exonic SNP, rs10741657, has its causal relationship with diseases established, including that of rickets, ovarian cancer, and multiple sclerosis. The role of other *CYP2R1* SNPs in vitamin D deficiency and their causal link to other traits however remain uncertain currently and more studies are warranted to help identify possible physiological mechanisms underlying those complex traits.

INTRODUCTION

Cytochrome P450 (CYP) enzymes are a superfamily of haem-containing isoenzymes. These enzymes carry out the biotransformation of a diverse number of exogenous and endogenous compounds, including steroids, fatty acids, vitamins, leukotrienes, and prostaglandins, and contribute to drug metabolism and detoxification [1]. In the human genome, there are 57 genes which encode active CYP enzymes, and these are further classified into 18 gene families [2]. Of these 57 isoenzymes, 51 are found in microsomes and 6 are located in mitochondria. CYP isoforms from families 1 to 3 have major roles in the biotransformation and bioinactivation of xenobiotics, including drugs and toxins. On the other hand, another set of CYP enzymes, many of which are expressed extrahepatically, display important endogenous functions, for instance, in regulating the levels of bile acids, lipids, steroid hormones and other signalling molecules.

The Human Genome Project has allowed the identification of some previously uncharacterised extrahepatic CYPs, the so-called 'orphan' CYPs, including CYP2R1, CYP2S1, CYP2U1 and CYP2W1. Like their counterparts in the liver, these CYPs has been shown to biotransform endogenous and/or foreign compounds in a tissue-specific manner. CYP2R1 is a major vitamin D 25-hydroxylase [3] whereas CYP2S1 has been identified as the main metabolizing enzyme for all-trans retinoic acid and naphthalene [4,5]. Likewise, CYP2U1 is known to metabolise arachidonic acid, docosahexaenoic acid (DHA) and other long chain fatty acids [6] while CYP2W1 has exhibited distinctive expression pattern in fetal and tumour tissues with pivotal role in localized retinoid metabolism [7].

With full sequence of the human genome completed and execution of various initiatives related to this project, significant progress has been made in the characterization of the structure, function, catalytic properties, and genetic polymorphism of CYP2R1. This review aims to summarize the recent understanding and comprehension of the gene organisation, protein structure and catalytic function of CYP2R1, and the influence of its gene polymorphism on occurrence of inherited rickets and other diseases.

CYP2R1 gene and protein structure
Gene organisation
The *CYP2R1* gene is localized on chromosome 11p15.2 and has five exons, spanning across a region of approximately 15.5 kb (Figure 1). The exon

number is different from many other CYP2 members where nine exons is reported. The only exception to this is *CYP2U1* which also possesses 5 exons. Both *CYP2U1* and *CYP2R1* have introns 1, 3, 4, 5, and 7, as found in all other *CYP2* genes, removed, leaving only introns 2, 6 and 8, indicating that the two genes must have shared a common ancestral gene that has undergone partial processing and reinsertion in their genome structures. Since then, both *CYP2U1* and *CYP2R1* have independently acquired an additional intron, hence the unique 5-exon structure presently noted in both genes. An overview of the relationships of the 57 CYPs found in human is shown in the phylogenetic tree in Figure 2. CYP2R1 and CYP2U1 show clear orthologous relationships, and are closely related to CYP2J2, CYP2S1, CYP2W1 and CYP2D6.
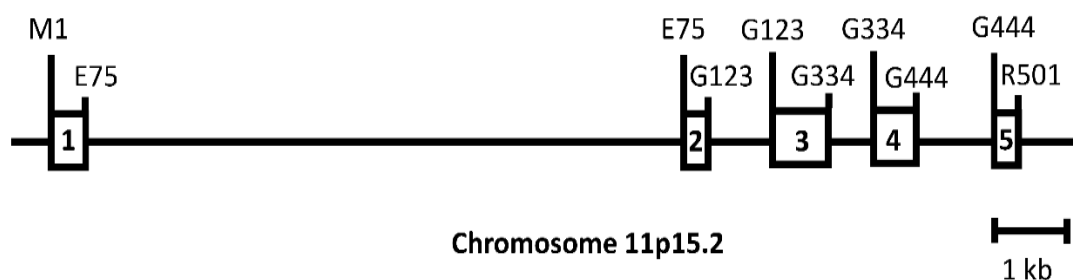

Figure 1. Structure of human *CYP2R1* gene. Introns are indicated by connecting lines and exons by boxes. Amino acids at the exon-intron junctions are indicated above the boxes. The structure of the gene and its indicated chromosomal location were deduced by comparing the sequences of the cloned cDNA to those of the genomic DNA (GenBankTM/EBI Data Bank accession no. NT_009237.15).

CYP2R1 is highly conserved among living kingdom with human CYP2R1 sharing 89% and 66% sequence homology to mouse and fugu fish CYP2R1, respectively, at primary sequence level. Highly conserved CYP enzymes are usually linked closely in metabolism of endogenous substrates, such as sterols, vitamins, and fatty acids, and CYP2R1 fits this pattern. CYP2R1 also displays high sequence homology with other CYP2 isoforms, with CYP2J2, CYP2U1 and CYP2D6 showing the highest homology among the 57 known human isoforms (41.2%, 38.6% and 38.5%, respectively). The sequence alignment of the abovementioned CYPs with CYP2R1 (Figure 3) showed that CYP2R1 has 501 residues and has similar overall sequence

features with other isoforms, including a lipophilic N-terminus, regions that interacts with the redox partner, NADPH cytochrome P450 oxidoreductase, and the cysteine ligand of the prosthetic group heme [10]. There is a region upstream of the transmembrane helix (TMH) in CYP2R1 comprising approximately 20 amino acids in its N-terminal sequence, which is not observed in other isoforms with the exception of CYP2U1 and CYP2J2 that show additional 40 and 20 pre-TMH residues, respectively (Figure 3). It is possible that these additional sequences affect intracellular targeting, membrane incorporation, or structure of CYP2R1, CYP2U1 and CYP2J2.
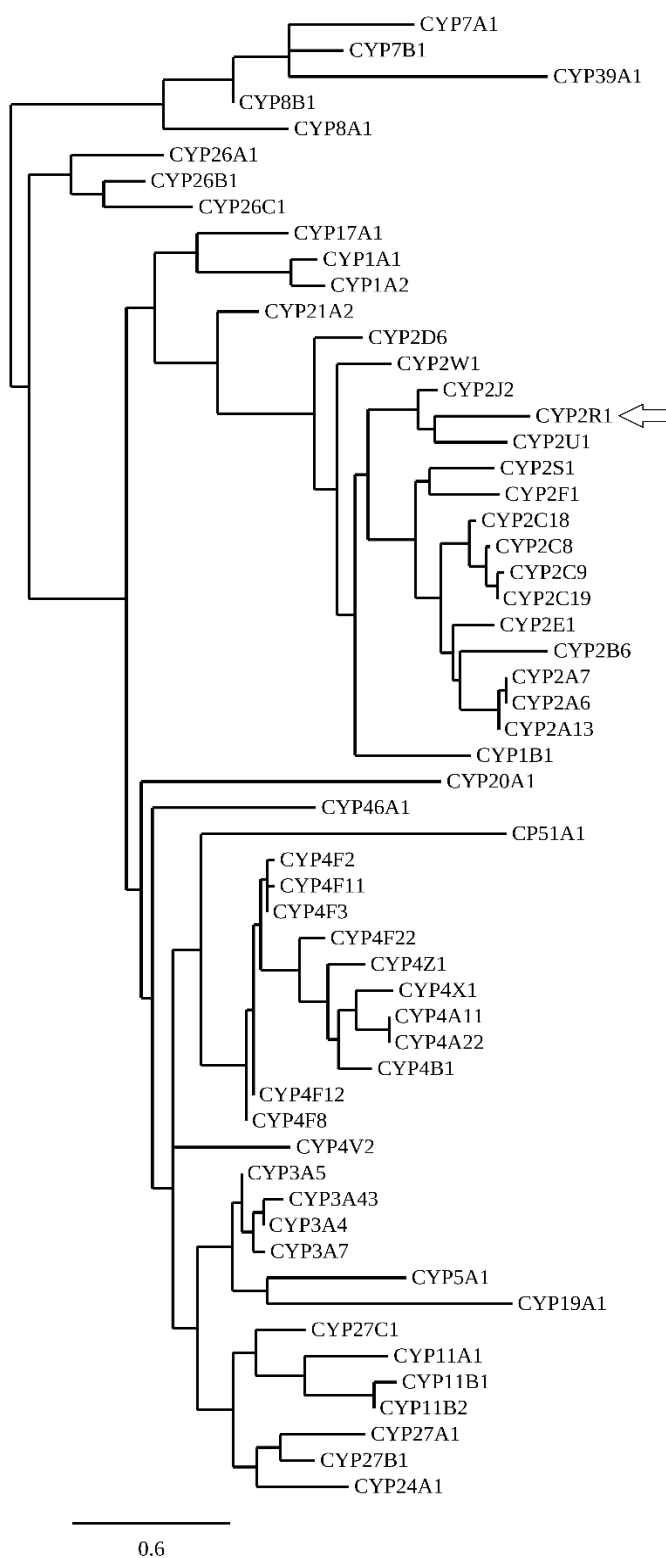
Figure 2. Phylogenetic tree of the 57 human CYPs compiled using Phylogeny.fr platform available at www.phylogeny.fr [8]. The arrow indicates the position of CYP2R1

The tissue distribution of CYP2R1 mRNA in humans and mice has been investigated. In humans, the descending order of CYP2R1 mRNA levels was pancreas, liver, and kidney. In mice, liver and testis demonstrated higher mRNA levels, with lower levels reported in other tissues such as brain, kidney, muscle, and skin [3]. A search in the web-based atlas of human gene expression, Amazonia! [11], for the distribution of CYPs in the human body, showed ubiquitous distribution of CYP2R1 in many tissues with testes recorded the highest mRNA levels, and quite high levels detected in many other organs. As for cDNA cloning, the mouse *CYP2R1* cDNA was cloned from a cDNA library constructed from CYP27A1-deficient mice. The cDNA-expressed CYP2R1 showed capability to perform 25-hydroxylation in both vitamin D_2 and vitamin D_3 [3,12].

Protein structure

A number of crystal structures of human CYP2R1 have recently been determined and published, including 3C6G (complexation with vitamin D_3), 3CZH (complexation with vitamin D_2), 2OJD (complexation with vitamin D_3), and 3DL9 (complexation with 1-α-hydroxy-vitamin D_2). The first structure (PDB code 3C6G) was solved by Strushkevich and coworkers [12]. The CYP2R1 existed as homodimer in solution and crystallized in a dimeric complex with the detergent molecule 2-hydroxypropyl-b-cyclodextrin bound. It has a fold pattern typical of a CYP protein, comprising 12 α-helices (A to L), a cluster of β-sheets mostly located on one side, and the heme group that was buried in the protein interior (Figure 4A). The core of the protein was formed by helices E, I, J, K and L. Helices F and G formed part of the homodimer interface and bordered the active site partly. The long I helix ran across the entire protein (Figure 4A), no distortion of this helix was noted in CYP2R1, even though it is observed in many mammalians and bacterial CYPs where a kink of the I-helix, a result of local distortion of intra-helical hydrogen bonding, was found to be important for CYP catalytic cycle in the delivery of proton to the O_2 molecule [13]. Nevertheless, CYP2R1 carried the highly conserved Thr314 which may be involved in proton transfer cascade to the heme-bound oxygen. The heme was tightly bound within the active site. Residues involved in this included Trp133, Arg137, Arg446

which were hydrogen bonded to the D-ring propionate, and those interacted with the A-ring propionate were Arg109, His381 and Ser442.

Overall, CYP2R1 structure adopts a closed conformation with the B′ helix serving as a gate covering the substrate access channel. The active site is bordered by B′ helix, F helix, I helix, the loop between C-terminal β-sheet 4, and the loop between helix K and β-sheet 1. The size of CYP2R1 active site, as determined from the crystal structure, is compared to a number of CYP crystal structures that have been published (Table 1). The active site volume of CYP2R1 (979 Å^3) is intermediate between those of CYP2A6, CYP2A13, CYP2B4 and CYP2E1 (<500 Å^3), that are known to only bind small substrates, and the isoforms receiving large substrates (volume in the range of 1000 - 2000 Å^3), such as CYP2C8 [19]. Its volume is comparable to those of CYP2J2 or CY2U1, therefore CYP2R1 is classified under those having a moderately sized active site cavity. Vitamin D_3 was demonstrated to enter the active site and was bound, in an elongated conformation, between the B' helix/B-C loop and helices G and I with the side chain directed toward the heme group (Figure 4B). The lining of the active site was formed mainly by hydrophobic residues. The 3β-OH group of the A-ring of vitamin D_3 faced the protein surface and adopted a chair conformation with an axial orientation of the 3β-OH group. The 3β-OH was hydrogen bonded to the amide function of Ala250 of the helix G in addition to a water-mediated hydrogen bond with Asn217 in helix F. The A-ring was surrounded by residues with aromatic side chains. Phe214 of helix F formed van der Waals interaction with the conjugated triene of vitamin D_3 and adopted a planar geometry perpendicular to the heme. Another hydrophobic residue, Phe115 of helix B' made hydrophobic contact with the C-18 methyl group of vitamin D_3. The 17β-aliphatic side-chain of vitamin D_3 was oriented within the active site cavity that placed C-25 toward the heme iron, for hydroxylation but at a distance of 6.5 Å, in contrast to a distance of 5 Å or less observed in other CYP-substrate structures such as CYP2C9 [25]. Taken together, substrate selectivity in CYP2R1 largely reflects interactions with the protein near the entrance of the substrate binding site where the secosterol portion of the substrate molecule is sequestered. Similar to other microsomal CYPs, CYP2R1 utilizes both electrostatic interactions in

complex formation with the redox partner NADPH-cytochrome P450 reductase. The interacting residues have been mapped and include residues in C helix (Arg131, Arg137, Arg138 and Arg145), meander region (Lys434, Lys435 and Arg445), as well as L helix (Arg455).

```
CYP2R1_HUMAN    ------------MWKLWRA-----EEGAAALG---GALFLLLFA-------LGV-RQLLK      32
CYP2S1_HUMAN    -ME---------ATGTWAL---------------------LLALALL----LLLTLALSG      25
CYP2U1_HUMAN    -MSSPGPSQPPAEDPPWPARLLRAPLGLLRLDPSGGALL---CGLVA---LL-GWSWLR     52
CYP2W1_HUMAN    ------------------------------------MALLLLLFLGLLGLWGLLC-ACAQD     24
CYP2D6_HUMAN    ----MC--------------------LEALVP--LAVIVAIFL-------LLVDLMHRR     26
CYP2J2_HUMAN    MLAAMGS----LAAALWAV-----VHPRTLLL---GTVAF-----------LLA-ADFLK     36
                                                                *

                              αA'                        αA
CYP2R1_HUMAN    QRRPMGFPPGPPGLPFIGNIYSLAASS--------------------ELPHVYMRRQSQ     71
CYP2S1_HUMAN    TRARGHLPPGPTPLPLLGNLLQLRPGA-----------------------LYSGLMRLSK     62
CYP2U1_HUMAN    RRRARGIPPGPTPWPLVGNFGHVLLPPFLRRRSWLSSRTRAACIDPSVIGPQVLLAHLAR    112
CYP2W1_HUMAN    PSPAARWPPGPRPLPLVGNLHLLRLSQ--------------------Q--DRSLMELSE     61
CYP2D6_HUMAN    QRWAARYPPGPLPLPGLGNLLHVDFQN--------------------T--PYCFDQLRR     63
CYP2J2_HUMAN    RRRPKNYPPGPWRLPFLCNFFLVDFE-----------------------QSHLEVQLFVK     73
                ****     *  :**:   :                             .    .   .

                      β1-1      β1-2     αB            αB'
CYP2R1_HUMAN    VYGEIFSLDLG-GISTVVLNGYDVVKECLVHQSEIFADRPCLPLFMKM---TKMGGLLNS    127
CYP2S1_HUMAN    KYGPVFTIYLGPWRPVVVLVGQEAVREALGGQAEEFSGRGTVAMLEGT---FDGHGVFF-    118
CYP2U1_HUMAN    VYGSIFSFFIG-HYLVVVLSDFHSVREALVQQAEVFSDRPRVPLISIV---TKEKGVVFA    168
CYP2W1_HUMAN    RYGPVFTVHLG-RQKTVVLTGFEAVKEALAGPGQELADRPPIAIFQLI---QRGGGIFF-    116
CYP2D6_HUMAN    RFGDVFSLQLA-WTPVVVLNGLAAVREALVTHGEDTADRPPVPITQILCFGPRSQCVFLA    122
CYP2J2_HUMAN    KYGNLFSLELG-DISAVLITGLPLIKEALIHMDQNFGNRPVTPMREHI---FKKNGIIM-    128
                                                                   SRS-1
                :* :*:. :.    .*::.    ::*.*    :  ..*     :        *:.

                    αC                    αD                αE
CYP2R1_HUMAN    RYGRGWVDHRRLAVNSFRYFGYGQKSFESKILEETKFFNDAIETYKGRPFDFKQLITNAV    187
CYP2S1_HUMAN    SNGERWRQLRKFTMLALRDLGMGKREGEELIQAEARCLVETFQGTEGRPFDPSLLLAQAT    178
CYP2U1_HUMAN    HYGPVWRQQRKFSHSTLRHFGLGKSLSEPKIIEEFKYVKAEMQXHGEDPFCPFSIISNAV    228
CYP2W1_HUMAN    SSGARWRAARQFTVRALHSLGVGREPVADKILQELKCLSGQLDGYRGRPFFL-ALLGWAP    175
CYP2D6_HUMAN    RYGPAWREQRRFSVSTLRNLGLGKKSLEQWVTEEAACLCAAFANHSGRPFRPNGLLDKAV    182
CYP2J2_HUMAN    SSGQAWKEQRRFTLTALRNFGLGKKSLEERIQEEAQHLTEAIKEENGQPFDPHFKINNAV    188
                 *   *  *:::  ::: :*   *:      :    *     .  :      **     :  *

                            αF              αF'        αG'
CYP2R1_HUMAN    SNITNLIIFGERFTYEDTDFQHMIELFSENVELAASASVFLYNAFPWIGILPFGKHQQLF    247
CYP2S1_HUMAN    SNVVCSLLFGLRFSYEDKEFQAVVRAAGGTLLGVSSQGGQTYEMFSWFLRPLPGPHXQLL    238
CYP2U1_HUMAN    SNIICSLCFGQRFDYTNSEFKKMLGFMSRGLEICLNSQVLLVNICPWLYYLPFGPFKELR    288
CYP2W1_HUMAN    SNITFALLFGRRFDYRDPVFVSLLGLIDEVMVLLGSPGLQLFNVYPWLGALLQLH-RPVL    234
CYP2D6_HUMAN    SNVIASLTCGRRFEYDDPRFLRLLDLAQEGLKEESGFLREVLNAVPVLLHIPALA-GKVL    241
CYP2J2_HUMAN    SNIICSITFGERFEYQDSWFQQLLKILLDEVTYLEASKTCQLYNVFPWIMKFLPGPHQTLE    248
                                       SRS-2
                **:   :   *  **  *  :   * ::        .          :     :     :

                  αG                    αH
CYP2R1_HUMAN    RNAAVVYDFLSRLIEKASVNRKPQL-PQHFVDAYLDEMDQ-GKNDPSSTFSKENLIFSVG    305
CYP2S1_HUMAN    HHVSTLAAFTVRQVQQHGQNLDASGPARDLVDAFLLKMAQ-EEQNPGTEFTKNMLMTVI    297
CYP2U1_HUMAN    QIEKDITSFLKKIIKDHQESLDREN-PQDFIDMYLLHMEEERKNNSNSSFDEEYLFYIIG    347
CYP2W1_HUMAN    RKIEEVRAILRTLLEARRPHVCPGDPVCSYVDALIQQ--G-QGDDPEGLFAEANAVACTL    291
CYP2D6_HUMAN    RFQKAFLTQLDELLTEHRMTWDPAQPPRDLTEAFLAEMEK-AKGNPESSFNDENLRIVVA    300
CYP2J2_HUMAN    SNWKKLKLFVSHMLDKHRKDWNPAE-TRDFIDAYLKEMSK-HTGNPTSSFHEENLICSTL    306
                SRS-3
                  .     :           : : :.       :   *  .

                  αI                  αJ               αJ'        αK
CYP2R1_HUMAN    ELIIAGTETTTNVLRWAILFMALYPNIQGQVQKEIDLIMGPNGKPSWDDKCKMPYTEAVL    365
CYP2S1_HUMAN    YLLFAGTMTVSTTVGYTLLLLMKYPHVQKWVREELNRELGAGQAPSLGDRTRLFYTDAVL    357
CYP2U1_HUMAN    DLFIAGTDTTTNSLLWCLLYMSLNPDVQEKVHEEIERVIGANRAPSLTDKAQMPYTEATI    407
CYP2W1_HUMAN    DMVMAGTETTSATLQWAALLMGRHFDVQGRVQELLDRVLCFGRIPRLEDQQALPYTSAVL    351
CYP2D6_HUMAN    DLFSAGMVTTSTTLAWGLLLMILHPDVQRRVQQEIDDVIGQVRRPEMGDQAHMPYTTAVI    360
CYP2J2_HUMAN    DLEFAGTEITSTTLRWALLYMALYPEIQEKVQAEIDRVIGQGQQPSTAARESMPYTNAVI    366
                SRS-4
                :.  **  *.:  :   *  :   *.:*  *:  *::   *     :    :*** *.:

                      β1-4     β2-1 β2-2    β1-3    αK'          αK''
CYP2R1_HUMAN    HEVLRFCNIVPLGIFHATSEDAVVRGYSIPKGTTVITNLYSVHFDEKYWRDPEVFHPERF    425
CYP2S1_HUMAN    HEAQRLLALVPMGIPRTLMRTTRFRGYTLPQGTEVFFLLGSILHDPNIFKHPEEFNPDRF    417
CYP2U1_HUMAN    MEVQRLTVVVPLATPHMTSENTVLQGYTIPKGTLILPNLWSVHRDPAIWEKPEDFYPNRF    467
CYP2W1_HUMAN    HEVQRFITLLP-HVPRCTAADTQLGGFLLPKGTFVIPLLTSVLLDETQWQTPGQFNPGHF    410
CYP2D6_HUMAN    HEVQRFGDIVPLGVTHMTSRDIEVQGFRIPKGTILITNLSSVLKDEAVWEKPFRFHPEHF    420
CYP2J2_HUMAN    HEVQRMGNIIPLNVPREVTVDTTLAGYHLPKGTMILTNLTALHRDPTEWATPDTFNPDHF    426
                         SRS-5
                *.  *:   ::*  :  :      . *:  :*:**  ::  *::  *   :  * * *:*

                                     αL            β3-1
CYP2R1_HUMAN    LDSSCYFAKKEALVPFSLGRRHCLGEHLARMEMPLFFTALLQRFHLHFPHEL-VPD--LK    482
CYP2S1_HUMAN    LDADGRFRKHEAFLPFSLGKRVCLGEGLAKAELFLFFTTILQAFSLESPCPPDTLS--LK    475
CYP2U1_HUMAN    LDDQGQLIKKETFIPFGIGKRVCMGEQLAKMELFLMFVSLMQSFAFALPEDSKKPL--LT    525
CYP2W1_HUMAN    LDANGHFVKREAFLPFSAGRRVCVGERLARTELPLLFAGLLQRYRLLPPEGVSPASLDIT    470
CYP2D6_HUMAN    LDAQGHFVKREAFLPFSAGRRACLGEPLARMELFLFFTSLLQHFSFSVPTGQPRPS--HH    478
CYP2J2_HUMAN    LEN-GQFKKREAFMPFSIGKRACLGEQLARTELFIFFTSLMQKFTFRPPNNE-KLS--LK    482
                *:  * : * *:::**. *:* *:** **: *:*::*. ::* : :  *

                β4-1   β4-2      β3-2
CYP2R1_HUMAN    PRL-GMTLQPQPYLICAFRR---------        501
CYP2S1_HUMAN    PTVSGLFNIPPAFQLQVRPTDLHSTTQTR       504
CYP2U1_HUMAN    GRF-GLTLAPHPFNITISRR---------       544
CYP2W1_HUMAN    PAR-AFTMRPRAQALCAVRP--------        490
CYP2D6_HUMAN    GVF-AFLVSPSPYELCAVPR---------       497
CYP2J2_HUMAN    FRM-GITISPVSHRLCAVPQV--------       502
                SRS-6
                 .:    *    :
```

Figure 3. Sequence alignment of human CYP2R1, CYP2S1, CYP2U1, CYP2W1, CYP2D6 and CYP2J2 performed using Clustal Omega program. The human CYP2R1 secondary structure elements are coloured (helices in blue, β-sheets in red). Predicted substrate recognition sites (SRS) of CYP2 family enzymes [9] are underlined and designated. Strongly conserved residues are indicated below the alignment with asterisks, colons indicate the residues variation occurring within the strongly conserved groups and dots indicating the residues variation occurring within weaker conserved residue groups. The sequences were taken from UniProtKB database with the following accession numbers (Q6VVX0 - CYP2R1, Q96AQ9-CYP2S1, Q7Z449-CYP2U1, Q8TAV3-CYP2W1, P10635-CYP2D6 and P51589-CYP2J2).

Table 1. Active site cavity volumes of CYP2 isoforms

Protein	PDB ID*	Active Site Volume (Å³)	Reference
CYP2A6	1Z10	250	[14]
CYP2A13	2P85	304	[15]
CYP2B4	1SUO	271	[16]
CYP2B6	31BD	636	[17]
CYP2C5	1N6B	645	[18]
CYP2C8	2NNH	1438	[19]
CYP2C9	1OG5	1667	[20]
CYP2D6	2F9Q	540	[21]
CYP2E1	3E4E	190	[22]
CYP2J2	3D model**	981	[23]
CYP2R1	3C6G	979	[12]
CYP2U1	3D model**	931	[24]

* PDB ID – identification code for the Protein Data Bank database
**3D model – the PDB ID is not available, published data from three-dimensional homology model is listed instead

(A)
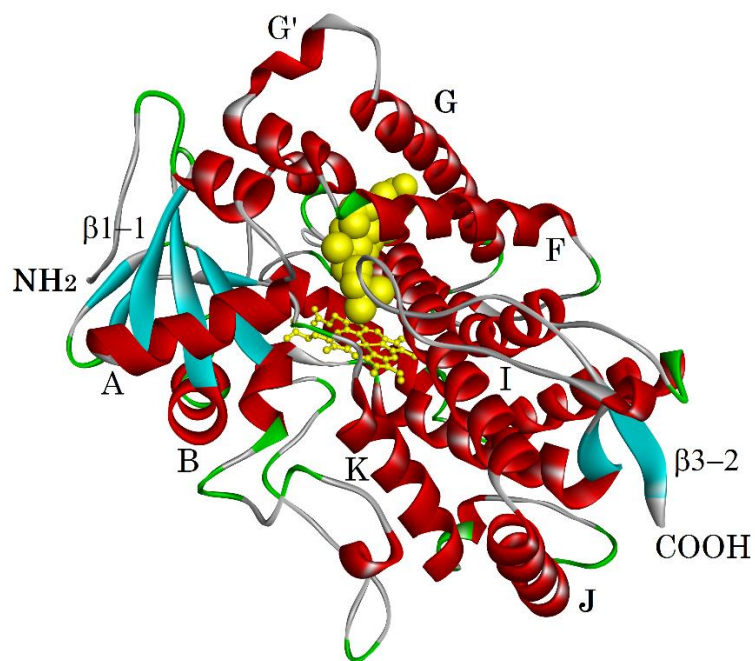
(B)
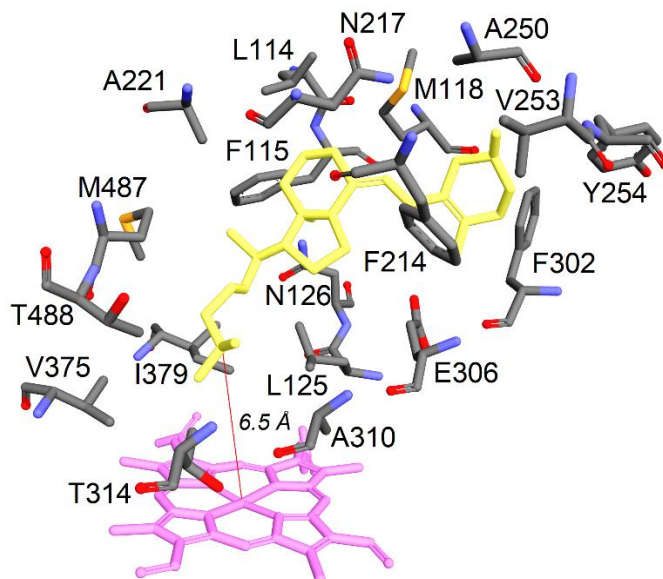
Figure 4. CYP2R1-vitamin D_3 crystal structure. **(A)** Overall structure of human CYP2R1 with helices coloured in red, beta sheets in cyan, turns in green, and coils in white. The heme and vitamin D_3 (both in yellow) are shown in ball and stick, and space-filling representations, respectively. **(B)** Close-up view of the binding of vitamin D_3 in the active site. The active site residues are shown in stick representations. Vitamin D_3 and the heme are coloured in light yellow and magenta, respectively. Red line indicates the distance between the C-25 carbon atom of vitamin D_3 from the heme.

CYP2R1 activity and function
Vitamin D metabolic pathway
Vitamin D plays a central role in calcium homeostasis and bone metabolism. Its production in the skin involves non-enzymatic reactions (Figure 5). Vitamin D_3 (cholecalciferol, the major secosteroid precursor in human body) is a breakdown product of 7-dehydrocholesterol (7-DHC) via a two-step process involving breaking down of a 7-DHC internal ring by the sun light, in particular radiation from UV light (spectrum 280-320 UVB), to form pre-vitamin D_3. The pre-cursor subsequently forms vitamin D_3 via isomerization in a thermosensitive but noncatalytic process. The rate of vitamin D_3 formation therefore depends on the intensity of UVB and skin pigmentation level [26]. An alternative source of vitamin D is dietary factor. Fatty fish is rich in vitamin D but level in most other foods is low unless fortified. The form in the fish is vitamin D_3, whereas vitamin D_2 (ergocalciferol) is usually used in fortified foods. Vitamin D_2 is produced by UVB irradiation of the ergosterol in plants and fungi such as mushrooms [27]. Both of these vitamins D lack intrinsic biological activity and are considered as hormone precursors.

Vitamin D is bioactivated through two sequential hydroxylation steps mediated by 25-hydroxylase and 1α-hydroxylase to generate 25-hydroxyvitamin D (25(OH)D_3 or 25(OH)D_2) in the liver and later, the hormonally active 1α,25-dihydroxyvitamin D (1,25(OH)$_2D_3$ or 1,25(OH)$_2D_2$) mainly in the kidney, respectively (Figure 5). 1,25(OH)$_2$D has a shorter half-life and a lower concentration than 25(OH)D (1000-fold lower), it is however the most active form of vitamin D and contributes to the bulk of the biological function of the vitamin. There is another hydroxylation step involving inactivation of 25(OH)D, by 24-hydroxylase, to form biologically inactive calcitroic acid. Distinct CYP isoforms are involved in production of the active hormone from inactive precursors and the subsequent biotransformation and inactivation. These enzymes include both microsomal (e.g., CYP2R1) or mitochondrial (e.g., CYP24A1, CYP27A1, and CYP27B1) isoforms. 25(OH)D, the major circulating form, is the most commonly used marker for vitamin D status, reflecting vitamin D input, from all sources including the diet and environment [28]. Also, in epidemiological studies, both 25(OH)D and

$1,25(OH)_2D$ are the most common reported biomarkers.

The liver is a well-established source for 25(OH)D production from vitamin D. Initial studies of the hepatic 25-hydroxlase found activity in both endoplasmic reticulum and mitochondrial fractions, and studies carried out subsequently have identified several CYPs to possess 25-hydroxylase activity. Thus far, the only mitochondrial 25-hydroxylase identified is CYP27A1, and it is widely distributed in the body inclusive of the liver [29]. Subsequent to identification of CYP27A1, CYP2R1 was identified as an active hepatic microsomal isoform in 25-hydroxylation in mice [3]. This isoform catalysed hydroxylation of both vitamin D_2 and D_3 with similar kinetics, unlike CYP27A1 that only hydroxylated vitamin D_3. Its expression was mainly in the liver and testes. The 25-hydroxylase activity has been found in CYP isoforms other than CYP27A1 and CYP2R1 (Figure 5). CYP3A4, the major drug-metabolizing isoform preferentially expressed in liver and intestine, possessed 25-hydroxylase activity [30]. CYP3A4 preferred 1(OH)D to both vitamin D_2 and D_3 as the substrate. CYP2J3 was also identified as the rat liver isoform with 25-hydroxylase activity, but CYP2J2, the human homolog to CYP2J3 and a primary isoform in the heart, was found to have lower 25-hydroxylase activity, and functioned mainly as an arachidonic acid epoxygenase [31]. CYP2C11, expressed in the liver of male rats, was demonstrated to have 25-hydroxylase activity for vitamin D_3 and D_2 as well as the 1(OH)D analogues but it was better established for its role in testosterone hydroxylations [32]. It is generally believed that CYP2R1 is the major 25-hydroxylase, but other enzymes having the same activity may play a role as determinants of the 25(OH)D level within a distinct tissue and possibly its circulating levels.

The second activation step in vitamin D metabolism involves 1α-hydroxylation by the mitochondrial CYP27B1 (Figure 5) [33,34]. CYP27B1 is widely distributed in many different tissues but generally at low levels. CYP27B1 catalyzed formation of $1,25(OH)_2D$ at the proximal tubule of the kidney, it primary site of action, from which its product entered the circulation and served its endocrine function [35]. Biologically active vitamin D forms have in general short half-lives in target cells. These active vitamin D forms attenuate their activity by auto-induction of metabolism via CYP24A1. CYP24A1 was shown to repeatedly attack the C_{20-27} side chain and was the only identified 24-hydroxylase in vitamin D metabolic pathway [36,37]. The 24-hydroxylase pathway was responsible for a cascade of sequential metabolic events resulting in a variety of products with increasing polarity, including the biologically inactive calcitroic acid, and eventual loss of hormonal activity [38].

CYP2R1 catalytic properties

As discussed earlier, CYP2R1 was first identified as microsomal vitamin D 25-hydroxylase in mouse and human via screening of a liver cDNA library from gene knockout mice for *cyp27a1* using a vitamin D receptor ligand activation assay [3]. The 25-hydroxylase activity of CYP2R1 was confirmed in the radioimmunoassay for detection of $25(OH)D_3$ as well as on thin layer chromatography plate based on co-migration with authentic standard. It was also shown that CYP2R1 worked together with CYP27B1 in generating $1,25(OH)_2D_3$ and their activity was counteracted by co-expression of CYP24A1, the isoform implicated in the degradation of the active ligand. Notably, CYP2R1 catalyzed 25 hydroxylation on both vitamin D_3 and vitamin D_2 at comparable rate, a feature that was not observed in CYP27A1 [39]. CYP2R1 has subsequently been expressed in yeast cells, where the microsomal fraction was isolated and examined for activity. This yeast-expressed CYP2R1 can hydroxylate both vitamin D_3 and vitamin D_2 at C25, with hydroxylase activity for vitamin D_3 being about double of the activity for vitamin D_2 [40]. Strushkevich and coworkers who first described CYP2R1 crystal structure (see section above) have demonstrated that the *E. coli*-expressed and purified CYP2R1 was able to metabolize vitamin D_3, $1(OH)D_3$ and $1(OH)D_2$ at similar rates to those reported for the yeast system [12]. In a more recent work, CYP2R1 was expressed in *E. coli* and purified by nickel affinity-chromatography and assayed for activity in a reconstituted membrane system comprising phospholipid vesicles supplemented with human NADPH cytochrome P450 reductase. CYP2R1 25-hydroxylated vitamin D_3 showing the typical Michaelis-Menten kinetics. The kinetic parameters for both CYP2R1 and CYP27A1 were investigated in the reconstituted mixtures under identical conditions. Lower k_{cat} and K_m than CYP27A1were observed for CYP2R1, leading to an overall 17-fold higher catalytic efficiency (k_{cat}/K_m), confirming the

major role of CYP2R1 as the vitamin D 25-hydroxylase. Moreover, CYP2R1 was able to metabolize $20(OH)D_3$, the main product of an alternative pathway mediated by CYP11A1, to $20,25(OH)_2D_3$, with catalytic efficiency similar to that for the 25-hydroxylation of vitamin D_3 [41]. It is apparent that the substrate specificity of CYP2R1 has been well characterized thus far, the inhibitor specificity however has not been investigated and reported, therefore this represents an area worthy of further exploration.

In a nutshell, several lines of evidence have substantiated the major role of CYP2R1 as vitamin D 25-hydroxylase in human: i) CYP2R1 demonstrated highest affinity for vitamin D with reported K_m values in submicromolar concentration range for both vitamin D_3 and vitamin D_2 [12,40]; ii) CYP2R1 exhibited highest specificity for vitamin D [and 1(OH)D)] as no other substrates have been described to date, and iii) a hereditary form of rickets is directly linked to a genetic defect in *CYP2R1* but not of other 25-hydroxylases (see 'Genetic Polymorphism of *CYP2R1*' section below) .
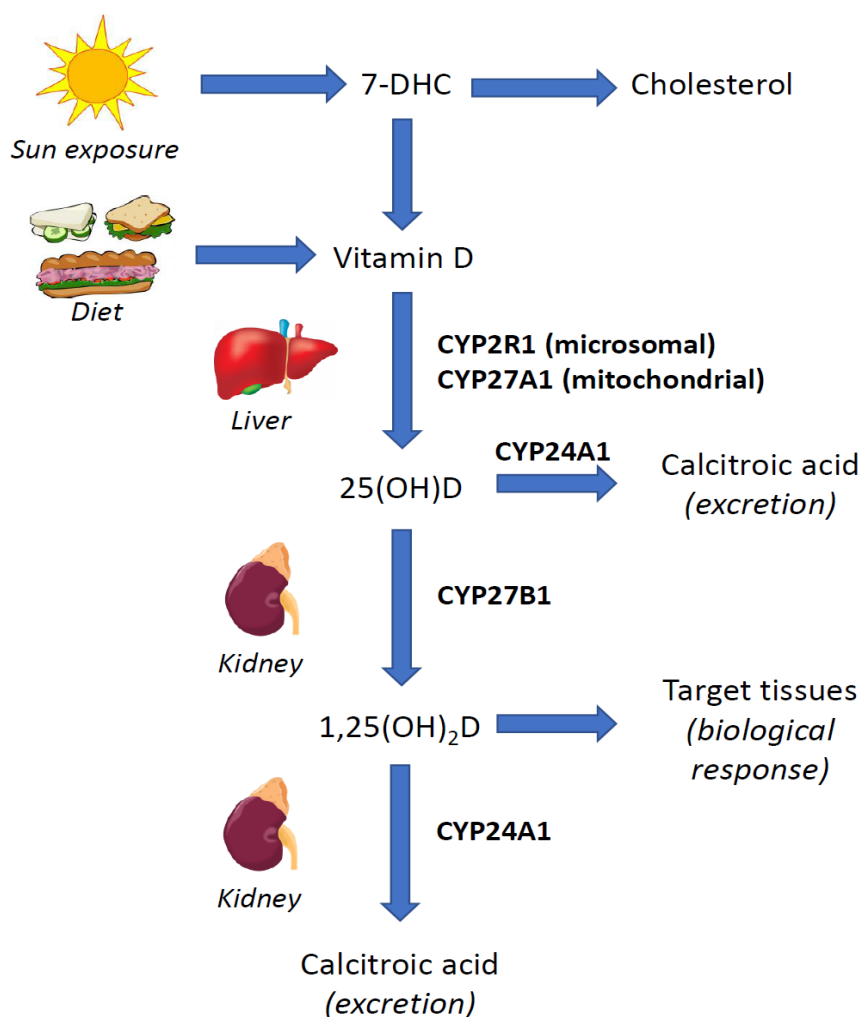
Figure 5. Vitamin D metabolic pathway.

Vitamin D signalling system

The hormonally active form 1α,25-dihydroxyvitamin D, 1,25(OH)$_2$D$_3$ or 1,25(OH)$_2$D$_2$, are similar functionally, and regulate the expression of a number of genes after binding to the vitamin D receptor (VDR). VDR is a widely expressed steroid hormone nuclear receptor family member serving as a transcription factor. Using real-time quantitative RT-PCR in mice, VDR was found to express mainly in the kidney and gastrointestinal tract, moderate levels in the thyroid, skin and bone, lower levels in most other organs, and very low levels in liver [42]. Generally, vitamin D exhibits genomic action similar to other steroid hormones. 1,25(OH)$_2$D binds, with high affinity, in the hydrophobic pocket of the VDR, thereby inducing a conformational change leading to formation of the transcriptional complex [43]. The activated complex then binds to the retinoid X receptor (RXR) and the resulting heterodimer translocates to vitamin D response elements, specific hexameric DNA sequence response elements, in the regulatory region of the target genes. This is followed by recruitment of coregulatory complexes eventually resulting in its genomic activity [44]. VDR activation leads to modulation of the gene expression in a cell-specific way. A large number of genes may be involved with up to 5% of the genome affected, and more than 900 genes are known to activated directly by vitamin D [45,46]. Among these are genes important for calcium homeostasis and skeletal integrity, including osteopontin, osteocalcin, calbindin, calcium channels TRPV5 and TRPV6, and RANKL [35]. In addition, vitamin D responsive genes are involved in modulating many physiological and pathophysiological systems including, among others, inflammation, cell proliferation and differentiation, metabolism, immune responses, response to oxidative stress, production of antibiotic peptides, as well as mental and cognitive function [46-49].

CYP2R1 polymorphism and health
Genetic Polymorphism of CYP2R1

An evidence for pivotal role of CYP2R1 in vitamin D activation and signalling as the 25-hydroxylase was the finding of a CYP2R1 polymorphism carrying Leu99Pro mutation in a Nigerian family that was presented with vitamin D-dependent rickets (VDDR) type 1B [50]. This disease was postulated four decades ago [51]. The genetic basis of the Leu99Pro mutation of CYP2R1 was however only elucidated by Cheng and colleagues [52] a decade after the initial report of the Nigerian rachitic patient [50]. This has made follow-up with the patient and family difficult. Cheng and colleagues also, in the same study, performed genetic analysis of CYP2R1 in 50 Nigerian individuals [52] and revealed one heterozygote with the Leu99Pro mutation, therefore pointing to a founder gene effect in the Nigerian population in which there was a high prevalence of vitamin D deficiency [53]. Leu99Pro exchange is the result of a single nucleotide polymorphism (SNP) involving 6359T>C mutation in exon 2 of CYP2R1 gene, and the allele has been designated as CYP2R1*2 (accessed in June 2020) on the Human Cytochrome P450 Allele Nomenclature Committee home page (https://www.pharmvar.org/genes). The molecular basis for the 25-hydroxylase deficiency involving the mutation has been elucidated in the CYP2R1 crystal structure. In the structure [12], Leu99, a B helix residue, was closely oriented to the heme binding loop but no direct substrate binding was observed. A rigid network of hydrogen bonds and structural water molecules was demonstrated to be important for structural integrity of this region in CYP proteins. The Leu99 carboxyl group was found to bond with Arg445 amide nitrogen via a water molecule. The Arg residue was located in critical part of the active site and was three residues from the heme coordinating Cys448 (Figure 6). The Leu99Pro mutation is believed to be able to disrupt the H-bond network and sterics of the helix in the region since proline residue is a helix breaker. Therefore, the mutation has likely caused protein misfolding that may adversely affect the enzymatic activity. Consistent with this notion, experiments using heterologous expression showed that the mutated CYP2R1 enzyme ~~showed~~ had negligible activity toward vitamin D$_3$ [52] or failed to express [12]. Studies using gene knockout technique revealed that serum 25(OH)D level was reduced by half compared with wild-type or heterozygous littermates in CYP2R1 knockout mice [29], implying that although CYP2R1 is a major physiologically relevant vitamin D 25-hydroxylase, redundancy exists in the vitamin D 25-hydroxylase "family" of enzymes, and there was compensation mechanism for the deletion of CYP2R1.

Subsequent to the discovery of CYP2R1*2,

over 20 SNPs of the *CYP2R1* gene associated with alteration in vitamin D levels have been identified in a number of genome-wide association studies (GWAS) carried out in European populations [54-57]. Numerous replication studies have also been conducted in populations of other ethnicity [58,59]. Among the identified SNPs, association between a SNP located in the non-coding region 5′-UTR, rs10741657, and 25(OH)D levels has been established [60-63] where individuals with the GG genotype demonstrated a decreasing trend of 25(OH)D levels in comparison to the no-risk genotype AA. It is believed that SNP in this 5′-UTR region is likely to regulate gene expression, and therefore the levels of activity and expression of 25-hydroxylase may be affected [64]. Two other SNPs,

the rs10766197 and rs12794714, have been identified in the coding region of introns, and these polymorphisms may be associated with changes in selective splicing regulation and gene regulation [64]. Subsequent GWAS however did not show any link between the SNPs and 25(OH)D levels in the total or subgroup populations. As rs12794714 and rs10766197 are in the intron region, these synonymous mutations do not change the protein sequence, hence minimal effect on the 25-hydroxylase activity and function [65]. To date, fairly large number of polymorphisms of the *CYP2R1* gene, including the above-mentioned SNPs, have been reported in various SNP databases including Sanger COSMIC, Ensembl, dbSNP, and 1000 Genomes.
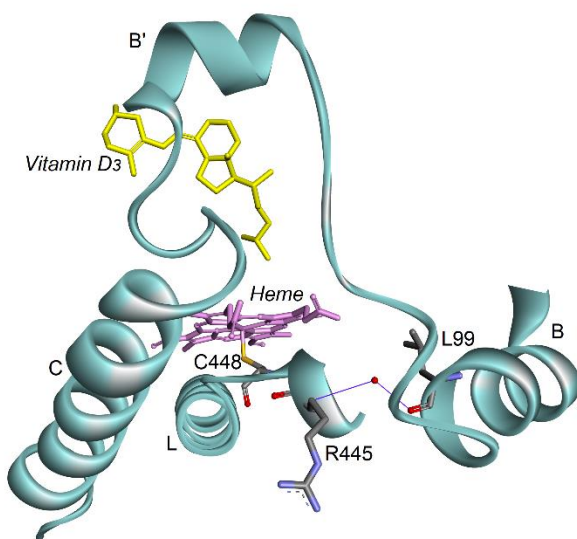
Figure 6. Positioning of Leu99 in the CYP2R1 structure. The heme is in light magenta, vitamin D_3 in yellow, and the water molecule is shown as red sphere.

Implications in health

As a steroid pro-hormone and fat-soluble vitamin, vitamin D plays a pivotal role in bone mineralization. Calcium and phosphorus levels in both blood and extracellular tissues are regulated by this fat-soluble vitamin, and adequacy of the mineral levels is vital for the deposition of calcium hydroxyapatite to the

bone matrix. Association of vitamin D deficiency with osteomalacia in adults and rickets in children has long been established [66]. Other than musculoskeletal disorders, common diseases, many of which are of important public health concerns, have been connected with vitamin D level. These include cardiovascular disease, cancer, infectious

disease, autoimmune inflammatory disorder, and diabetes [67-70]. Meta-analyses from a number of epidemiological studies showed a close relationship between levels of the precursor 25(OH)D and the incidence and severity of the abovementioned diseases that might be improved by adequate supplementation with vitamin D [71,72]. Treatment using vitamin D analogues works successfully in indications like type I rickets, secondary hyperparathyroidism, osteoporosis, and hyperproliferative skin conditions such as psoriasis [73,74]. Furthermore, experimental and clinical evidence has emerged that vitamin D analogues should work in some of the autoimmune diseases and malignancies [75,76].

The causal role of vitamin D in bone-related disorders is well established, and its importance beyond bone health has begun to emerge in recent years. Findings from GWAS, whole genome sequencing (WGS), as well as Mendelian randomization (MR) studies have shed light on genetic determinants of circulating vitamin D, including SNPs associated with *CYP2R1*, and their putative causal relationship with various disorders and traits. GWAS of the genetic determinants of serum 25(OH)D concentrations [54,60] concluded that variants at the chromosomal site for *CYP2R1* (11p15.2) (SNP rs10741657) was the second strongest association of only four sites, with vitamin D-binding protein (*DBP*, formerly known as *GC*), 7-dehydrocholesterol reductase (*DHCR7*), and CYP24A1, being the others (rs2282679, rs17216707 and rs12785878, respectively). These four susceptibility loci have exhibited a critical role whereby their common genetic variants have been associated to altered regulation of circulating 25(OH)D concentrations. This is not unexpected because the four loci are all in or near genes encoding crucial steps in vitamin D synthesis, transportation and bioinactivation. Interestingly, variants of the other 25-hydroxylases, such as CYP27A1, did not demonstrate any significant association with serum 25(OH)D levels, again arguing that it plays negligible role in 25-hydroxylation of vitamin D at physiologically relevant concentrations.

One recent GWS study has investigated the impact of low-frequency and rare variants (minor allele frequency, MAF < 5% and < 1%, respectively) in 25(OH)D, using WGS data from the UK10K program involving 2619 individuals, and deep-imputation data from 39,655 individuals with genome-wide genotypes [65]. Meta-analysis has identified a synonymous *CYP2R1* variant with low-frequency occurrence (MAF=2.5%, rs117913124) among the 19 participating cohorts. This SNP was able to modulate 25(OH)D levels substantially, showing four times stronger impact and the change was independent of the influence of another reported SNP rs10741657 in *CYP2R1*. This SNP has also been found to double the risk of vitamin D insufficiency, a magnitude that is likely to be of clinically relevance. Another recent GWS study investigated the vitamin D modulating genes in early childhood for a cohort of 761 healthy Finnish children who participated in a vitamin D intervention study [77]. The study has involved a genome-wide analysis of almost 700,000 genetic variants and demonstrated that genetic variation in *DBP* and *CYP2R1* were important determinants for serum 25(OH)D concentration. Genetic variants within the *DBP* gene also demonstrated impact on how the children responded to vitamin D supplementation. In addition, data from the study suggested that in 2-year-old children, vitamin D level, even when within the normal range, influenced bone strength as children with genetic constellations linked to lower vitamin D concentration and poorer response to vitamin D supplementation also showed lower bone strength, as judged from skeletal peripheral quantitative computed tomography parameters. These results, combined with data from other GWASs, are consistent with an oligogenetic architecture for vitamin D. This indicates that vitamin D is affected by a few genetic variants, with some having relatively large effects, therefore indicating an oligogenic feature [78].

MR is an approach that assesses the causal effect of a risk factor on an outcome from observational data where genetic variants are used as instrumental variables [79]. MR analyses have provided some evidence for a role of the genetically predicted level of 25(OH)D in certain traits or diseases. One MR study, looking into the GWAS summary statistics from ovarian cancer cases, has provided evidence for a causal effect of vitamin D in the cancer. The odds ratio was estimated to be 1.27 (1.06–1.51) per 20 nmol/L decline in 25(OH)D concentration for epithelial ovarian cancer risk, and the risk was higher (1.54 [1.19–2.01]) for high-grade serous subtype [80]. These results are supported by a Danish MR study which showed a causal effect of low circulating 25(OH)D with cancer-specific

mortality at a borderline significance [1.03 (1.00–1.06)] [81]. The investigated gene loci include *CYP2R1* SNPs, rs10741657 and rs12794714, alongside with those of *DHCR7* and *DBP*. Despite these positive relationships observed, no significant effect of 25(OH)D on the canser risk was reported in the MR studies conducted in other cancer types, including oral, oesophageal, lung, pancreatic, breast, prostate, colorectal, neuroblastoma, glioma, melanoma, and non-melanoma skin cancer [82-91]. These data imply that although some degree of causal effect of 25(OH)D with maglinancies cannot be ruled out, there may be a limit on the potential benefits of 25(OH)D on lowering cancer risk.

The causal link between vitamin D and metabolic diseases or traits remains unclear at this stage. So far, only one MR study has reported a change of 0.029 mmHg in diastolic blood pressure, and an 8.1% decreased odds ratio in hypertension [0.92 (0.87–0.97)] for each 10% increase in genetically instrumented 25(OH)D level [92]. Another MR study has demonstrated a link between 50% reduction in genetically predicted 25(OH)D level with 6.0% lower level (p=0.001) of high-density lipoprotein [93]. No obvious effect of genetically predicted vitamin D level has been found for type 2 diabetes [62], body mass index [94], cardiovascular diseases [95,96], metabolic syndrome [97] or other metabolites [98,99,100].

A causal role has been well established in the autoimmune neurodegenerative disorder, multiple sclerosis (MS), for vitamin D. Four independent MR studies have clearly demonstrated a positive link between genetically predicated 25(OH)D and MS. The first study showed the doubling risk of MS occurrence [odds ratio=2.02 (1.65–2.46); p=7.72×10^{-12}] with each standard deviation decrease in natural log-transformed 25(OH)D levels. Four SNPs were included in the analysis, namely rs10741657 in *CYP2R1*, rs2282679 in *DBP*, rs12785878 in *DHCR7*, and rs6013897 in *CYP24A1* [101]. The second study involved MR analyses in US and Swedish populations and calculation of a weighted genetic score for each individual based on the number of 25(OH)D increasing alleles at three different loci (*CYP2R1, DBP* and *DHCR7*). Meta-analysis across both populations showed that an increased genetic score was protective against both adult MS [odds ratio=0.85 (0.76–0.94); p=0.003] and pediatric-onset MS [odds ratio=0.72 (0.55–0.94); p=0.02] [102,103]. The data are supported by another

study conducted on samples of 5927 MS case and 5599 control subjects, where individuals carrying one copy of the rare variant (*CYP2R1* rs117913124) showed increased risk of MS (OR=1.4 (1.19–1.64); p=2.63×10^{-5}) [65]. A more recent MR study [104] has further confirmed the protective effect of vitamin D on MS, where a 43% decrease in the odds of MS [odds ratio=0.57 (0.41-0.81); p=0.001] was noted for each genetically determined unit increase in the natural-log-transformed vitamin D level. Similar results are also reported for Alzheimer's disease. A study has demonstrated an increased risk for the disease [odds ratio=1.25 (1.03–1.51); p=0.02] was linked to a standard deviation decrease in natural log-transformed 25(OH)D levels [105]. This is confirmed by another study [odds ratio=1.16 (1.06–1.28), p=0.002] [106]. A more recent MR study investigating the causal association between DBP levels and Alzheimer's disease showed that a standard deviation increase of the genetically DBP levels was significantly associated with reduced AD risk [odds ratio=0.63 (0.45-0.89); p=0.009 [107]. In this study, the *DBP* rs2282679 variant was selected as the instrumental variable, and the dataset was extracted from the multiple GWAS summary statistics data from International Genomics of Alzheimer's Project (IGAP). The positive associations observed above for MS and Azheimer's disease were however not evidenced across autoimmune and neurodegenerative diseases. No significant finding has been revealed for asthma [108], rheumatoid arthritis [109], Parkinson's disease [110], cognition status [111,112] or amyotrophic lateral sclerosis [113].

As a summary, examination of the genetic determinants of vitamin D could expand our knowledge to the disease itself, assist in the screening of vitamin D deficiency, and also promote personalized recommendations related to the use of supplementation. Outcomes assessed by GWAS, WGS and MR studies of vitamin D have been mixed, with both positive and negative outcomes noted. There is a need for findings from these studies to be interpreted in the context of other evidence relevant to the particular issues under investigation. The results derived from the studies may help with planning of randomised clinical trials on target populations or diseases, and also help to prioritize vitamin D treatment to the most promising target diseases. Moving forward, well-powered population-based whole genome sequencing and exome

sequencing studies are required to reveal the role of rare and low-frequency variants in vitamin D-related disorders. Identification of new promising susceptibility loci would provide an improved instrument strength and accuracy of estimation, and will further enhance our knowledge of genetic regulation of vitamin D homeostasis.

CONCLUSION

Since the identification of CYP2R1 as the major microsomal 25-hydroxylase for vitamin D [3], significant progress has been made in the characterization of gene organisation, protein structure and catalytic function of CYP2R1, and the relationship of its genetic polymorphism with inherited rickets and other vitamin D-deficiency diseases. The *CYP2R1* gene is localized on chromosome 11p15.2 and has five exons, unlike most other CYP isoforms that carry nine exons. From the X-ray crystallographic studies of the protein, CYP2R1 displays a fold pattern typical of a CYP protein, consisting of 12 α-helices, with β-sheets mostly on one side, and the heme buried deeply in the protein interior [12]. Overall, CYP2R1 structure adopts a closed conformation with the B′ helix serving as a gate covering the substrate access channel, which defines the substrate entrance point. Co-crystallized vitamin D_3 in the CYP2R1 occupied a position with the side chain pointing toward the heme group. In liver, CYP2R1 25-hydroxylates both vitamin D_2 and D_3 with comparable kinetics and plays pivotal role in determining 25(OH)D level in both the tissue and circulation. While substrate profile has been well studied, inhibitor specificity for CYP2R1 however has not been well investigated to a significant extent to date. Like most other CYP isoforms, several polymorphisms have been identified in *CYP2R1* gene. Both exonic and non-exonic SNPs have been reported, including the *CYP2R1*2* which carries a SNP resulting in Leu99Pro exchange in protein sequence [52], and a number of non-exonic SNPs with variable functional consequences in gene regulation. A non-exonic SNP, rs10741657, located in the 5'-UTR, is established as an important genetic determinant of circulating vitamin D level, and has its causal relationship with diseases established, including that of rickets, ovarian cancer and MS [54,60,80,101,102,103]. The role of other *CYP2R1* SNPs in vitamin D deficiency and their causal link to other traits however remain uncertain currently and more studies are needed to unveil the gene-disease association for *CYP2R1*.

REFERENCES

1. Nebert D.W., Russel D.W. Clinical importance of the cytochromes P450. Lancet (2002) **360** 1155-1162.
2. Nelson D.R., Zeldin D.C., Hoffman S.M.G. et al. Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants. Pharmacogenetics (2004) **14** 1-18.
3. Cheng J.B., Motola D.L., Mangelsdorf D.J., Russell D.W. Deorphanization of cytochrome P450 2R1: a microsomal vitamin D 25-hydroxilase. J. Biol. Chem. (2003) **278** 38084-38093.
4. Buckpitt A., Boland B., Isbell M. et al. Naphthalene-induced respiratory tract toxicity: metabolic mechanisms of toxicity. Drug Metab. Rev. (2002) **34** 791–820.
5. Smith G., Wolf C.R., Deeni Y.Y. et al. Cutaneous expression of cytochrome P450 CYP2S1: individuality in regulation by therapeutic agents for psoriasis and other skin diseases. Lancet (2003) **361** 1336–1343.
6. Chuang S.S., Helvig C., Taimi M. et al. CYP2U1, a novel human thymus- and brain-specific cytochrome P450, catalyzes omega- and (omega-1)-hydroxylation of fatty acids. J. Biol. Chem. (2004) **279** 6305-6314.
7. Zhao Y., Wan D., Yang J. et al. Catalytic activities of tumor-specific human cytochrome

P450 CYP2W1 toward endogenous substrates. Drug Metab. Dispos. (2016) **44** 771-780.

8. Dereeper A., Guignon V., Blanc G. et al. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res. (2008) **36**(Web Server issue) W465-9.

9. Gotoh O. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. J. Biol. Chem. (1992) **267** 83-90.

10. Peterson J.A., Graham S.E. A close family resemblance: the importance of structure in understanding cytochromes P450. Structure (1998) **6** 1079–1085.

11. Le Carrour T., Assou S., Tondeur S. et al. Amazonia!: An online resource to Google and visualize public human whole genome expression data. Open Bioinform. J. (2010) **4** 5–10.

12. Strushkevich N., Usanov S.A., Plotnikov A.N. et al. Structural analysis of CYP2R1 in complex with vitamin D_3. J. Mol. Biol. (2008) **380** 95–106.

13. Poulos T.L., Finzel B.C., Gunsalus I.C. et al. The 2.6-Å crystal structure of Pseudomonas putida cytochrome P-450. J. Biol. Chem. (1985) **260** 16122–16130.

14. Yano J.K., Hsu M.H., Griffin K.J. et al. Structures of human microsomal cytochrome P450 2A6 complexed with coumarin and methoxsalen. Nat. Struct. Mol. Biol. (2005) **12** 822-823.

15. Smith B.D., Sanders J.L., Porubsky P.R. et al. Structure of the human lung cytochrome P450 2A13. J. Biol. Chem. (2007) **282** 17306-17313.

16. Scott E.E., White M.A., He Y.A. et al. Structure of mammalian cytochrome P450 2B4 complexed with 4-(4-chlorophenyl)imidazole at 1.9-Å resolution. J. Biol. Chem. (2004) **279** 27294-27301

17. Nguyen T.A., Tychopoulos M., Bichat F. et al. Improvement of cyclophosphamide activation by CYP2B6 mutants: From *in silico* to *ex vivo.* Mol. Pharmacol. (2008) **73** 1122-1133.

18. Wester M.R., Johnson E.F., Marques-Soares C. et al. Structure of a substrate complex of mammalian cytochrome P450 2C5 at 2.3-Å resolution: Evidence for multiple substrate binding modes. Biochemistry (2003) **42** 6370-

6379.

19. Schoch G.A., Yano J.K., Sansen S. et al. Determinants of cytochrome P450 2C8 substrate binding. Structures of complexes with montelukast, troglitazone, felodipine and 9-cis-retinoic acid. J. Biol. Chem. (2008) **283** 17227-17237.

20. Williams P.A., Cosme J., Ward A. et al. Crystal structure of human cytochrome P450 2C9 with bound warfarin. Nature (2003) **424** 464-468.

21. Rowland P., Blaney F.E., Smyth M.G. et al. Crystal structure of human cytochrome P450 2D6. J. Biol. Chem. (2006) **281** 7614-7622.

22. Porubsky P.R., Meneely K.M., Scott E.E. Structure of human cytochrome P450 2E1. Insights into the binding of inhibitors and both small molecular weight and fatty acid substrates. J. Biol. Chem. (2008) **283** 33698-33707.

23. Lafite P., André F., Zeldin D.C. et al. Unusual regioselectivity and active site topology of human cytochrome P450 2J2, Biochemistry (2007) **46** 10237-10247.

24. Ducassou L., Jonasson G., Dhers L. et al. Expression in yeast, new substrates, and construction of a first 3D model of human orphan cytochrome P450 2U1: interpretation of substrate hydroxylation regioselectivity from docking studies. Biochim. Biophys. Acta. (2015) **1850** 1426-1437.

25. Wester M.R., Yano J.K., Schoch G.A. et al. The structure of human cytochrome P450 2C9 complexed with flurbiprofen at 2.0-Å resolution. J. Biol. Chem. (2004) **279** 35630-35637.

26. Holick M.F., MacLaughlin J.A., Clark M.B. et al. Photosynthesis of previtamin D3 in human skin and the physiologic consequences. Science (1980) **210** 203–205.

27. Tripkovic L., Lambert H., Hart K. et al. Comparison of vitamin D2 and vitamin D3 supplementation in raising serum 25-hydroxyvitamin D status: a systematic review and meta-analysis. Am. J. Clin. Nutr. (2012) **95** 1357–1364.

28. Holick M.F. Vitamin D deficiency. N. Engl. J. Med. (2007) **357** 266-281.

29. Zhu J.G., Ochalek J.T., Kaufmann M. et al. CYP2R1 is a major, but not exclusive, contributor to 25-hydroxyvitamin D production in vivo. Proc. Natl. Acad. Sci. USA

(2013) **110** 15650–15655.

30. Gupta R.P., Hollis B.W., Patel S.B. et al. CYP3A4 is a human microsomal vitamin D 25-hydroxylase. J. Bone Miner. Res. (2004) **19** 680–688.

31. Zhu J., DeLuca H.F. Vitamin D 25-hydroxylase - four decades of searching, are we there yet? Arch. Biochem. Biophys. (2012) **523** 30–36.

32. Rahmaniyan M., Patrick K., Bell N.H. Characterization of recombinant CYP2C11: a vitamin D 25-hydroxylase and 24-hydroxylase. Am. J. Physiol. Endocrinol. Metab. (2005) **288** E753–E760.

33. Shinki T., Shimada H., Wakino S. et al. Cloning and expression of rat 25-hydroxyvitamin D_3-1alpha-hydroxylase cDNA. Proc. Natl. Acad. Sci. USA (1997) **94** 12920-12925.

34. Takeyama K., Kitanaka S., Sato T. et al. 25-Hydroxyvitamin D_3 1alpha-hydroxylase and vitamin D synthesis. Science (1997) **277** 1827–1830.

35. Suzuki Y., Landowski C.P., Hediger M.A. Mechanisms and regulation of epithelial Ca^{2+} absorption in health and disease. Ann. Rev. Physiol. (2008) **70** 257–271.

36. Sakaki T., Sawada N., Komai K. et al. Dual metabolic pathway of 25-hydroxyvitamin D_3 catalyzed by human CYP24. Eur. J. Biochem. (2000) **267** 6158–6165.

37. Jones G., Prosser D.E., Kaufmann M. 25-Hydroxyvitamin D-24-hydroxylase (CYP24A1): its important role in the degradation of vitamin D. Arch. Biochem. Biophys. (2012) **523** 9–18.

38. Prosser D.E., Jones G. Enzymes involved in the activation and inactivation of vitamin D. Trends Biochem. Sci. (2004) **29** 664–673.

39. Guo Y.D., Strugnell S., Back D.W., Jones G. Transfected human liver cytochrome P-450 hydroxylates vitamin D analogs at different side-chain positions. Proc. Natl. Acad. Sci. USA (1993) **90** 8668-8672.

40. Shinkyo R., Sakaki T., Kamakura M. et al. Metabolism of vitamin D by human microsomal CYP2R1. Biochem. Biophys. Res. Commun. (2004) **324** 451-457.

41. Cheng C.Y.S., Kim T-K., Jeayeng S. et al. Properties of purified CYP2R1 in a reconstituted membrane environment and its

25-hydroxylation of 20-hydroxyvitamin D_3. J. Steroid Biochem. Mol. Biol. (2018) **177** 59-69.

42. Bookout A.L., Jeong Y., Downes M. et al. Anatomical, profiling of nuclear receptor expression reveals a hierarchical transcriptional network. Cell (2006) **126** 789–799.

43. Rochel N., Wurtz J.M., Mitschler A. et al. The crystal structure of the nuclear receptor for vitamin D bound to its natural ligand. Mol. Cell. (2000) **5** 173–179.

44. Pike J.W., Meyer M.B., Watanuki M. et al. Perspectives on mechanisms of gene regulation by 1, 25-dihydroxyvitamin D_3 and its receptor. J. Steroid Biochem. Mol. Biol. (2007) **103** 389–395.

45. Wang T.T., Tavera-Mendoza L.E., Laperriere D. et al. Large-scale in silico and microarray-based identification of direct 1,25-dihydroxyvitamin D_3 target genes. Mol. Endocrinol. (2005) **19** 2685–2695.

46. Bouillon R., Carmeliet G., Verlinden L. et al. Vitamin D and human health: lessons from vitamin D receptor null mice. Endocr. Rev. (2008) **29** 726–776.

47. Gombart A.F. The vitamin D – antimicrobial peptide pathway and its role in protection against infection. Future Microbiol. (2009) 4 1151–1165.

48. Kovalenko P.L., Zhang Z., Cui M. et al. 1,25 Dihydroxyvitamin D-mediated orchestration of anticancer, transcript-level effects in the immortalized, non-transformed prostate epithelial cell line, RWPE1. BMC Genomics. (2010) **11** 26.

49. Bikle D.D. Vitamin D metabolism, mechanism of action, and clinical applications, Chem. Biol. (2014) **21** 319–329.

50. Casella S.J., Reiner B.J., Chen T.C. et al. A possible genetic defect in 25-hydroxylation as a cause of rickets. J. Pediatr. (1994) **124** 929-932.

51. Fraser D., Kooh S.W., Kind H.P. et al. Pathogenesis of hereditary vitamin-D-dependent rickets. An inborn error of vitamin D metabolism involving defective conversion of 25-hydroxyvitamin D to 1α,25-dihydroxyvitamin D. N. Engl. J. Med. (1973) **289** 817-822.

52. Cheng J.B., Levine M.A., Bell N.H. et al. Genetic evidence that the human CYP2R1

enzyme is a key vitamin D 25-hydroxylase. Proc. Natl. Acad. Sci. USA (2004) **101** 7711-7715.

53. Thacher T.D., Fischer P.R., Pettifor J.M. et al. Case-control study of factors associated with nutritional rickets in Nigerian children. J. Pediatr. (2000) **137** 367-373.

54. Ahn, J., Yu K., Stolzenberg-Solomon R.et al. Genome-wide association study of circulating vitamin D levels. Hum. Mol. Genet. (2010) **19** 2739-2745.

55. Anderson D., Holt B.J., Pennell C.E. et al. Genome-wide association study of vitamin D levels in children: replication in the western Australian pregnancy cohort (Raine) study. Genes Immun. (2014) **15** 578-583.

56. O'Brien K.M., Sandler D.P., Shi M. et al. Genome-wide association study of serum 25-hydroxyvitamin D in US women. Front. Genet. (2018) **9** 67.

57. Xia J., O'Reilly P.F., Aschard H. et al. Genome-wide association study in 79,366 European-ancestry individuals informs the genetic architecture of 25-hydroxyvitamin D levels. Nat. Commun. (2018) **9** 260.

58. Robien K., Butler L.M., Wang R. et al. Genetic and environmental predictors of serum 25(OH)D concentrations among middle-aged and elderly Chinese in Singapore. Br. J. Nutr. (2013) **109** 493-502.

59. Xu X., Mao J., Zhang M. et al. Vitamin D deficiency in Uygurs and Kazaks is associated with polymorphisms in CYP2R1 and DHCR7/NADSYN1 genes. Med. Sci. Monit. (2015) **21** 1960-1968.

60. Wang T.J., Zhang F., Richards J.B. et al. Common genetic determinants of vitamin D insufficiency: a genome-wide association study. Lancet (2010) **376** 180-188.

61. Slater N.A., Rager M.L., Havrda D.E., Harralson A.F. Genetic variation in CYP2R1 and GC genes associated with vitamin D deficiency status. J. Pharm. Pract. (2015) **30** 31-36.

62. Ye Z., Sharp S.J., Burgess S. et al. Association between circulating 25-hydroxyvitamin D and incident type 2 diabetes: a mendelian randomisation study. Lancet Diabetes Endocrinol. (2015) **3** 35-42.

63. Duan L., Xue Z., Ji H. et al. Effects of CYP2R1 gene variants on vitamin D levels and status: A

64. systematic review and meta-analysis. Gene (2018) **678** 361-369.

64. Ramos-Lopez E., Brueck P., Jansen T. et al. CYP2R1 (vitamin D 25-hydroxylase) gene is associated with susceptibility to type 1 diabetes and vitamin D levels in Germans. Diabetes Metab. Res. Rev. (2007) **23** 631-636.

65. Manousaki D., Dudding T., Haworth S. et al. Low-frequency synonymous coding variation in CYP2R1 has large effects on vitamin D levels and risk of multiple sclerosis. Am. J. Hum. Genet. (2017) **101** 227-238.

66. Doppelt S.H. Vitamin D, rickets, and osteomalacia, Orthop. Clin. North Am. (1984) **15** 671-686.

67. Scaranti M., Júnior Gde C., Hoff A.O. Vitamin D and cancer: does it really matter? Curr. Opin. Oncol. (2016) **28** 205-209.

68. Al Mheid I., Quyyumi A.A. Vitamin D and cardiovascular disease: controversy unresolved. J. Am. Coll. Cardiol. (2017) **70** 89-100.

69. Altieri B., Muscogiuri G., Barrea L. et al. Does vitamin D play a role in autoimmune endocrine disorders? A proof of concept. Rev. Endocr. Metab. Disord. (2017) **18** 335-346.

70. Issa C.M. Vitamin D and type 2 diabetes mellitus. Adv. Exp. Med. Biol. (2017) **996** 193-205.

71. Kulie T., Groff A., Redmer J. et al. Vitamin D: an evidence based review. J. Am. Board Fam. Med. (2009) **22** 698-706.

72. Pittas A.G., Laskowski U., Kos L., Saltzman E. Role of vitamin D in adults requiring nutrition support. J. Parenter. Enteral Nutr. (2010) **34** 70-78.

73. Segaert S., Duvold L.B. Calcipotriol cream: a review of its use in the management of psoriasis. J. Dermatolog. Treat. (2006) **17** 327-337.

74. Brown A.J., Slatopolsky E. Drug insight: vitamin D analogs in the treatment of secondary hyperparathyroidism in patients with chronic kidney disease. Nat. Clin. Pract. Endocrinol. Metab. (2007) **3** 134-144.

75. Masuda S., Jones G. Promise of vitamin D analogues in the treatment of hyperproliferative conditions. Mol. Cancer Ther. (2006) **5** 797-808.

76. Deeb K., Trump D.L., Johnson C.S. Vitamin D signaling pathways in cancer: potential for

anticancer therapeutics. Nat. Rev. Cancer (2007) **7** 684-700.

77. Kämpe A., Enlund-Cerullo M., Valkama S. et al. Genetic variation in GC and CYP2R1 affects 25-hydroxyvitamin D concentration and skeletal parameters: A genome-wide association study in 24-month-old Finnish children. PLoS Genet. (2019) **15** e1008530.

78. Timpson N.J., Greenwood C.M.T., Soranzo N. et al. Genetic architecture: the shape of the genetic contribution to human traits and disease, Nat. Rev. Genet. (2018) **19** 110–124.

79. Davies N.M., Holmes M.V., Davey Smith G. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ (2018) **362** k601.

80. Ong J.S., Cuellar-Partida G., Lu Y. et al. Association of vitamin D levels and risk of ovarian cancer: a Mendelian randomization study. Int. J. Epidemiol. (2016) **45** 1619-1630.

81. Afzal S., Brøndum-Jacobsen P., Bojesen S.E., Nordestgaard B.G. Genetically low vitamin D concentrations and increased mortality: Mendelian randomisation analysis in three large cohorts. BMJ (2014) **349** g6330.

82. Theodoratou E., Palmer T., Zgaga L. et al. Instrumental variable estimation of the causal effect of plasma 25-hydroxy-vitamin D on colorectal cancer risk: a mendelian randomization analysis. PLoS One (2012) **7** e37662.

83. Trummer O., Langsenlehner U., Krenn-Pilko S. et al. Vitamin D and prostate cancer prognosis: a Mendelian randomization study. World J. Urol. (2016) **34** 607-611.

84. Dimitrakopoulou V.I., Tsilidis K.K., Haycock P.C. et al. Circulating vitamin D concentration and risk of seven cancers: Mendelian randomisation study. BMJ (2017) **359** j4761.

85. Dudding T., Johansson M., Thomas S.J. et al. Assessing the causal association between 25-hydroxyvitamin D and the risk of oral and oropharyngeal cancer using Mendelian randomization. Int. J. Cancer. (2018) **143** 1029-1036.

86. Takahashi H., Cornish A.J., Sud A. et al. Mendelian randomisation study of the relationship between vitamin D and risk of glioma. Sci. Rep. (2018) **8** 2339.

87. Wang S., Huo D., Kupfer S. et al. Genetic variation in the vitamin D related pathway and breast cancer risk in women of African ancestry in the root consortium. Int. J. Cancer. (2018a) **142** 36-43.

88. Winsløw U.C., Nordestgaard B.G., Afzal S. High plasma 25-hydroxyvitamin D and high risk of nonmelanoma skin cancer: a Mendelian randomization study of 97 849 individuals. Br. J. Dermatol. (2018) **178** 1388-1395.

89. Dong J., Gharahkhani P., Chow W.H. et al. No association between vitamin D status and risk of Barrett's esophagus or esophageal adenocarcinoma: a Mendelian randomization study. Clin. Gastroenterol. Hepatol. (2019) **17** 2227-2235.

90. Jiang X., Dimou N.L., Al-Dabhani K. et al. Circulating vitamin D concentrations and risk of breast and prostate cancer: a Mendelian randomization study. Int. J. Epidemiol. (2019) **48** 1416-1424.

91. Liyanage U.E., Law M.H.; Melanoma Meta-analysis Consortium et al. Is there a causal relationship between vitamin D and melanoma risk? A Mendelian randomization study. Br. J. Dermatol. (2020) **182** 97-103.

92. Vimaleswaran K.S., Cavadino A., Berry D.J. et al. Association of vitamin D status with arterial blood pressure and hypertension risk: a mendelian randomisation study. Lancet Diabetes Endocrinol. (2014) **2** 719-729.

93. Ooi E.M., Afzal S., Nordestgaard B.G. Elevated remnant cholesterol in 25-hydroxyvitamin D deficiency in the general population: Mendelian randomization study. Circ. Cardiovasc. Genet. (2014) **7** 650-658.

94. Vimaleswaran K.S., Berry D.J., Lu C. et al. Causal relationship between obesity and vitamin D status: bi-directional Mendelian randomization analysis of multiple cohorts. PLoS Med. (2013) **10** e1001383.

95. Brøndum-Jacobsen P., Benn M., Afzal S., Nordestgaard B.G. No evidence that genetically reduced 25-hydroxyvitamin D is associated with increased risk of ischaemic heart disease or myocardial infarction: a Mendelian randomization study. Int. J. Epidemiol. (2015) **44** 651-661.

96. Manousaki D., Mokry L.E., Ross S. et al. Mendelian randomization studies do not support a role for vitamin D in coronary artery disease. Circ. Cardiovasc. Genet. (2016) **9** 349-356.

97. Chen C., Chen Y., Weng P. et al. Association of 25-hydroxyvitamin D with cardiometabolic risk factors and metabolic syndrome: a mendelian randomization study. Nutr. J. (2019) **18** 61.

98. Husemoen L.L.N., Skaaby T., Martinussen T. et al. Investigating the causal effect of vitamin D on serum adiponectin using a Mendelian randomization approach, Eur. J. Clin. Nutr. (2014) **68** 189-195.

99. Liefaard M.C., Ligthart S., Vitezova A. et al. Vitamin D and C-reactive protein: a Mendelian randomization study. PLoS One (2015) **10** e0131740.

100. Wang N., Chen C., Zhao L. et al. Vitamin D and nonalcoholic fatty liver disease: bi-directional Mendelian randomization analysis. EBioMedicine. (2018b) **28** 187-193.

101. Mokry L.E., Ross S., Ahmad O.S. et al. Vitamin D and risk of multiple sclerosis: a Mendelian randomization study. PLoS Med. (2015) **12** e1001866.

102. Rhead B., Bäärnhielm M., Gianfrancesco M. et al. Mendelian randomization shows a causal effect of low vitamin D on multiple sclerosis risk. Neurol. Genet. (2016) **2** e97.

103. Gianfrancesco M.A., Stridh P., Rhead B. et al. Evidence for a causal relationship between low vitamin D, high BMI, and pediatric-onset MS. Neurology (2017) **88** 1623-1629.

104. Jacobs B.M., Noyce A.J., Giovannoni G., Dobson R. BMI and low vitamin D are causal factors for multiple sclerosis: A Mendelian Randomization study. Neurol. Neuroimmunol. Neuroinflamm. (2020) **7** e662.

105. Mokry L.E., Ross S., Morris J.A. et al. Genetically decreased vitamin D and risk of Alzheimer disease. Neurology (2016) **87** 2567-2574.

106. Larsson S.C., Traylor M., Markus H.S., Michaëlsson K. Serum parathyroid hormone, 25-hydroxyvitamin D, and risk of Alzheimer's disease: a Mendelian randomization study. Nutrients. (2018) **10 pii** E1243.

107. Zhang H., Wang T., Han Z. et al. Impact of vitamin D binding protein levels on Alzheimer's disease: a Mendelian Randomization study. J. Alzheimers Dis. (2020) **74** 1-8.

108. Hysinger E.B., Roizen J.D., Mentch F.D. et al. Mendelian randomization analysis demonstrates that low vitamin D is unlikely causative for pediatric asthma. J. Allergy Clin. Immunol. (2016) **138** 1747-1749.e4.

109. Viatte S., Yarwood A., McAllister K. et al. The role of genetic polymorphisms regulating vitamin D levels in rheumatoid arthritis outcome: a Mendelian randomization approach. Ann. Rheum. Dis. (2014) **73** 1430-1433.

110. Larsson S.C., Singleton A.B., Nalls M.A. et al. No clear support for a role for vitamin D in Parkinson's disease: a Mendelian randomization study. Mov. Disord. (2017) **32** 1249-1252.

111. Kueider A.M., Tanaka T., An Y. et al. State- and trait-dependent associations of vitamin-D with brain function during aging, Neurobiol. Aging (2016) **39** 38-45.

112. Maddock J., Zhou A., Cavadino A. et al. Vitamin D and cognitive function: a Mendelian randomisation study. Sci. Rep. (2017) **7** 13230.

113. Larsson S.C., Roos P.M. Serum 25-hydroxyvitamin D in amyotrophic lateral sclerosis: Mendelian randomization study. Neurobiol. Aging (2020) **87** 140.e1-140.e3.