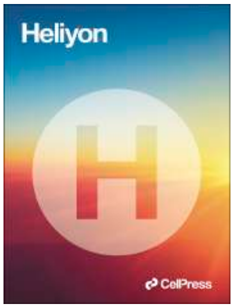


Functional characterization of the *SDR42E1* reveals its role in vitamin D biosynthesis

Nagham Nafiz Hendi, Maria Teresa Bengoechea-Alonso, Johan Ericsson, Georges Nemer

PII: S2405-8440(24)12497-8

DOI: https://doi.org/10.1016/j.heliyon.2024.e36466

Reference: HLY 36466

To appear in: *HELIYON*

Received Date: 11 July 2024

Accepted Date: 15 August 2024

Please cite this article as: Functional characterization of the *SDR42E1* reveals its role in vitamin D biosynthesis, *HELIYON*, https://doi.org/10.1016/j.heliyon.2024.e36466.

Functional characterization of the *SDR42E1* reveals its role in vitamin D biosynthesis

Nagham Nafiz Hendi[1,2,3], Maria Teresa Bengoechea-Alonso[1], Johan Ericsson[1], and Georges Nemer[2,4*]

[1] Division of Biological and Biomedical Sciences, and [2] Division of Genomics and Translational Biomedicine, College of Health and Life Sciences, Hamad Bin Khalifa University, Doha P.O. Box 34110, Qatar; pericsson@hbku.edu.qa; MAlonso@hbku.edu.qa.

[3] Systems Biology and Immunology Department, Sidra Medicine, Doha P.O. Box 26999, Qatar; nhendi@sidra.org, ORCID:0000-0003-2737-0565.

[4] Department of Biochemistry and Molecular Genetics, American University of Beirut, Beirut P.O. Box 110236, Lebanon.

* Correspondence:

Georges Nemer, College of Health and Life Sciences, Hamad Bin Khalifa University, Doha, Qatar; email: gnemer@hbku.edu.qa; Tel.: +974-445-41330; ORCID: 0000-0003-2157-5279.

Highlights

- SDR42E1 knockout alters key vitamin D synthesis regulators: EBP, DHCR7, ALPP, and CYP26A1.
- Multi-omics reveal SDR42E1's broad role in steroid synthesis and lipid metabolism.
- SDR42E1 knockout accumulates 7-dehydrocholesterol, hindering vitamin D production.
- SDR42E1 variant presents a promising target for addressing vitamin D deficiency.

Abstract

Vitamin D deficiency poses a widespread health challenge, shaped by environmental and genetic determinants. A recent discovery identified a genetic regulator, rs11542462, in the *SDR42E1* gene, though its biological implications remain largely unexplored. Our bioinformatic assessments revealed pronounced *SDR42E1* expression in skin keratinocytes and the analogous HaCat human keratinocyte cell lines, prompting us to select the latter as an experimental model. Employing CRISPR/Cas9 gene-editing technology and multi-omics approach, we discovered that depleting *SDR42E1* showed a 1.6-fold disruption in steroid biosynthesis pathway (P-value = 0.03), considerably affecting crucial vitamin D biosynthesis regulators. Notably, *SERPINB2* (P-value = 2.17 x 10^{-103}), *EBP* (P-value = 2.46 x 10^{-13}), and *DHCR7* (P-value = 8.03 x 10^{-09}) elevated by ~2-3 fold, while *ALPP* (P-value < 2.2 x 10^{-308}), *SLC7A5* (P-value = 1.96 x 10^{-215}), and *CYP26A1* (P-value = 1.06 x 10^{-08}) downregulated by ~1.5-3 fold. These alterations resulted in accumulation of 7-dehydrocholesterol and reduction of vitamin D production, as evidenced by the drug enrichment (P-value = 4.39 x 10^{-06}) and vitamin D quantification (R^2 = 0.935, P-value = 0.0016) analyses. Our investigation unveils *SDR42E1*'s significance in vitamin D homeostasis, emphasizing the potential of precision medicine in addressing vitamin D deficiency through understanding its genetic basis.

Keywords

SDR42E1; Vitamin D biosynthesis; Steroidogenesis; CRISPR/Cas9; Multi-omics; HaCat.

1. Introduction

Vitamin D deficiency, characterized by suboptimal 25-hydroxyvitamin D (25(OH)D) levels below 20 ng/mL (50 nmol/L), represents a widespread nutritional deficiency linked to critical health conditions, including osteoporosis and rickets. The intricate interplay of genetic determinants substantially influences serum 25(OH)D concentrations, with twins and familial studies revealing notable variations ranging from 23% to 90% [1]. The prevalence of vitamin D deficiency in regions with abundant sunlight, such as the Middle East and North-Africa (MENA) region, emphasizes the relevance of genetic determinants [2]. However, the complex pathways and involvement of multiple genes in vitamin D synthesis hinder our comprehensive understanding of the genetic contribution to 25(OH)D variation.

Recent genome-wide association studies (GWAS) have provided crucial insights into the genetic architecture of 25(OH)D, identifying single-nucleotide polymorphisms (SNP) statistically linked to 25(OH)D levels [3]. One notable variant is identified in the novel, uncharacterized human short-chain dehydrogenase/reductase family 42E, member 1 (*SDR42E1*) on chromosome 16q23 in exon 3, a genomic locus that has received limited attention. This variant, identified as rs11542462 in the SNP database, introduces a premature stop codon, resulting in the substitution of amino acids, specifically Glutamine, with termination at position 30 of the protein (p.Q30* GLN>*TER). This mutation potentially leads to a non-functional SDR42E1 enzyme [4-6].

SDR42E1 is a member of the extended short-chain dehydrogenase/reductase superfamily with a broad substrate specificity and potential involvement in lipid metabolism [7]. Although its precise function remains uncertain, the protein is implicated in regulating cellular processes and may impact steroid synthesis through proposed roles as an oxidoreductase and steroid delta-isomerase, potentially utilizing nicotinamide adenine dinucleotide phosphate (NAD(P)(H)) [8, 9]. Interestingly, numerous genetic investigations highlight a close interrelation between *SDR42E1* and the regulation of steroid hormone biosynthesis [10, 11]. Given the potential importance of this gene, further investigation is crucial to elucidate the structure, biological function, and impact of *SDR42E1* mutations on human health.

The nonsense variant in *SDR42E1* has been associated with serum concentrations of vitamin D precursor, 8-dehydrocholesterol (8-DHC) [12]. This steroidal compound demonstrates a close relationship with a crucial precursor in vitamin D synthesis, namely 7-dehydrocholesterol (7-DHC), both of which accumulate in patients affected by Smith-Lemli-Opitz syndrome [13, 14]. Our previous research on *in silico* characterization of the *SDR42E1* identified potential substrates of the protein as 8-DHC, 7-DHC, and 25(OH)D [8]. In this study, we expand those findings to a comprehensive profiling of *SDR42E1* functions using a multi-disciplinary approach, integrating bioinformatic analysis with the use of clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated protein 9 (Cas9) gene-editing technology to mimic the *SDR42E1* nonsense variant in the selected human keratinocyte HaCat cell line.

2. Methods

2.1. *In silico* bioinformatics analyses

The expressions of quantitative trait loci (eQTLs) of *SDR42E1* in normal and cancer human tissues were explored through Genotype-tissue Expression (GTEx, http://commonfund.nih.gov/GTEx/). The All RNA-sequencing and Combining Chromatin Immunoprecipitation-sequencing (ChIP-seq) Sample Search Space (ARCHS4) database was also used to assess the expression map of *SDR42E1* in cells and tissues, which contains data on 187,964 human and mouse samples, as previously described [15].

2.2. Cell culture

The human immortalized skin keratinocyte, HaCat (CRL-2309, Research Resource Identifier (RRID): CVCL_0038), and human embryonic kidney cell lines, HEK293T (CRL-3216), were obtained from the American Type Culture Collection (ATCC, Manassas, VA, United States) or generously provided by collaborators.

Cells were seeded in T75 tissue culture flasks (Sigma-Aldrich, United States) and cultured in Dulbecco's modified Eagle's medium (DMEM) with (4.5 g/L) high D-Glucose, (2 mM) Glutamine, and (1 mM) Sodium Pyruvate, supplemented with 10% fetal bovine serum (FBS) and (1X) antibiotic-antimycotic (Thermo Fisher Scientific Gibco, United States). To prevent contamination, the medium was changed every other day. At 70% to 80% confluency, all cells were harvested through treatment with 0.25% trypsin-ethylenediaminetetraacetic acid (EDTA, Sigma-Aldrich). The cell pellets were subsequently rinsed once with sterile (1X) Dulbecco's phosphate-buffered saline (DPBS; Gibco), underwent centrifugal precipitation at 900 revolutions per minute (rpm) for 3 minutes, and passaged at a 1:6 ratio. All cell lines were grown under sterile conditions in a monolayer culture at 37°C, with a 5% CO_2 atmosphere and 95% air humidity.

2.3. Plasmid construction and single guide RNA cloning

To generate a single-cell-derived *SDR42E1* knockout through CRISPR/Cas9 technology, GenScript (United States) designed single guide RNA (sgRNA) targeting exon 3 of the human *SDR42E1* gene, proximate to the p.Q30* GLN>*TER premature stop codon mutation, as illustrated in Figure S2 a. These sgRNAs were synthesized by

Integrated DNA Technologies (IDT, United States) with overhangs complementary to the BsmBI-digested plasmid (sgRNAs-*SDR42E1*, Table S1). After the BsmBI digestion and filler fragment excision using GeneJET Gel Extraction Kit (Thermo Scientific, United States), the sgRNA guide sequences were integrated into the BsmBI site of the LentiCRISPR v2-Puro-U6 vector (a gift from Feng Zhang, RRID: Addgene_52961; http://n2t.net/addgene:52961, Addgene plasmid, United States) according to a corresponding Addgene protocol with minor modifications [16]. Briefly, oligonucleotides for the sgRNA-*SDR42E1* were annealed through heating to 95 °C for 5 minutes and gradual cooling to 25 °C. Subsequently, they were phosphorylated using T4 Polynucleotide Kinase (New England Biolabs, United States) at 37 °C for 30 minutes and inactivated at 70 °C for 10 minutes. The annealed oligonucleotides were cloned into the digested lentiCRISPR v2 backbone using T4 DNA ligase (Invitrogen, United States) at 16 °C overnight. The three obtained CRISPR/Cas9 constructs were amplified in Stbl3 chemically competent *Escherichia coli* (C737303, Invitrogen) through heat shock transformation and cultured in Super Optimal Broth Culturing medium (SOC, Invitrogen) for 1 hour. Positive colonies were purified from (100 µg/mL) ampicillin-supplemented-LB media using QIAprep Spin Miniprep (Qiagen, Germany). The insertion of the sgRNA cassette was confirmed for several colonies through colony polymerase chain reaction (PCR) validation and DNA Sanger sequencing using the human U6 Forward primer (U6 for Lenti-CRISPRv2, Table S1). The most efficient sgRNAs were selected for the *SDR42E1* knockout experiments, namely sgRNA2-*SDR42E1* and sgRNA3-*SDR42E1* (Figure S2 b). Non-targeting lentiCRIPSR v2 plasmid (an empty vector) served as a negative control.

2.4. Lentiviral production

HEK293T cells, leveraging the crucial Simian Virus 40 (SV40) large T-antigen for viral vector production, were transiently transfected with non-targeting and *SDR42E1*-targeting CRISPR/Cas9 plasmids to produce lentiviral particles (LVPs). Briefly, HEK293T cells were seeded in a 10-cm petri dish (Sigma-Aldrich) for 20-24 hours. Cells were then supplemented with DMEM containing 10% Bovine serum (Gibco) and treated with (25 mM) chloroquine (Cayman, United States) to enhance LVP stability and transfection efficiency. At 65-70% confluency, co-transfection of (12 µg) lentiviral DNA constructs with lentivirus packaging mix (Dharmacon, TLP4606, United States) was executed using calcium phosphate precipitation as described previously [17]. At 6 hours of incubation at 37 °C, culture media was changed to standard culture medium, and cells were incubated for at least 72 hours to achieve high-titer virus production. Afterward, the viral supernatant was harvested, filtered through 0.45-µm sterile low protein binding filters (Millipore, Sigma-Aldrich), and frozen at -80°C in small aliquots, being freshly thawed for each infection cycle.

2.5. Generation of CRISPR/Cas9-mediated *SDR42E1*-edited cells

The targeted HaCat cells were infected with viruses generated from each CRISPR/Cas9-sgRNA construct at a multiplicity of infection (MOI) of 5 in 6-well plates (Sigma-Aldrich), in the presence of (10 µg/mL) polybrene to induce sgRNA expression. After 24 hours of transduction at 37°C, puromycin (A1113803, Thermo Fisher Scientific) was applied at a concentration of 2-5 µg/mL for approximately 7 days to eliminate non-transduced cells. Cells were then harvested at 80% confluency for DNA and mRNA extractions to verify the gene editing efficiency.

The array dilution method, a robust isolation technique, was utilized to generate a monoclonal homozygous *SDR42E1* gene-edited cell line from a polyclonal pool of heterozygous cells. Infected cells were isolated and sorted into single clones in a 96-well plate through a 2-fold serial dilution, first vertically, and then horizontally. Three to four weeks post-infection, individual cell clones were verified for genome editing efficiency and genotyping using T7 endonuclease 1 (T7E1) assay, reverse transcription-quantitative polymerase chain reaction (RT-qPCR), Sanger sequencing, and Western blotting. The top gene-edited clone, exhibiting homozygous knockout, was selected for further investigations into gene function and expression.

2.6. T7 endonuclease 1 mismatch assay

The efficiency of CRISPR/Cas9-mediated gene-editing of the designed sgRNAs was validated through a T7E1 assay, as per the instructions, which recognizes and cleaves non-perfectly matched DNA. Genomic DNA was extracted from infected HaCat cells using Quick Extract Genomic DNA buffer (A560001, AMPLIQON, Denmark), following the manufacturer's protocol. After DNA quantification, the sgRNA genomic target site in *SDR42E1* exon 3 was PCR-amplified with high-fidelity AccuPrime Taq DNA polymerase (NEB 2U/uL, Thermo Scientific) using primers flanking the target site (T7-*SDR42E1*, Table S1) and purified with a GeneJET PCR purification kit (Thermo Scientific). The PCR amplicons (200 ng) were denatured and reannealed to form heteroduplexes DNA, followed by digestion with T7E1 enzyme (M0302L, New England Biolabs) at 37°C for 30 minutes. The digested products were run on a 2% agarose gel at 90 volts for 40 minutes to verify the size and specificity of the products. Control samples, including empty lentiCRISPR vector and non-T7 digested-genomic samples, were used for result validation.

2.7. DNA sequencing

To determine the exact genotype of the CRISPR/Cas9-mediated *SDR42E1* gene-editing, Sanger sequencing of the purified PCR product of genomic DNA was conducted by Macrogen Inc. (http://macrogen.com, Korea), using the forward T7-*SDR42E1* primer (T7-*SDR42E1*, Table S1). For mutation identification, the obtained sequences were aligned to the human genomic reference sequence of *SDR42E1* from Ensembl (ENST00000328945) and compared to a non-targeting lentiCRISPR vector control using Snapgene version 5.3.2 (GSL Biotech LLC; Chicago, IL, United States).

2.8. Cell lysis and Western blot of immunoprecipitation

To confirm the loss of protein expression in *SDR42E1* gene-edited cells, washed cells were lysed using ice-cold radioimmunoprecipitation assay (RIPA) buffer, supplemented with (1 mM) phenylmethylsulfonyl fluoride (PMSF), (1M) dithiothreitol (DDT), and a protease inhibitor cocktail (UD282713, Thermo Fisher Scientific). After a 30-minute incubation at 4°C with rotation, the lysates were centrifuged at 10,000 ×g for 15 minutes at 4°C. The supernatants were stored at -80°C until analysis. Protein concentration was determined photometrically using a bicinchoninic acid (BCA) protein assay kit (Pierce, Thermo Fisher Scientific).

Whole-cell extracts were pre-incubated with the desired antibody, anti-SDR42E1 rabbit monoclonal antibody (Thermo Fisher Scientific Cat# PA5-53156, RRID: AB_2647060, 1:500 dilution), overnight with gentle rotation at 4 °C to generate specific immunocomplex. Protein samples were then incubated with protein A/G magnetic beads (Pierce, 88802, Thermo Scientific) for 2 hours with gentle rotation at 4 °C. After a brief centrifugation at 3000 rpm for 30 seconds to pellet the magnetic beads with bound immunocomplexes, they were placed in a magnetic separation rack and the supernatant was discarded. Magnetic beads were briefly spun down with the bound immunocomplexes at 3000 rpm for 30 seconds, and subsequently placed in a magnetic separation rack and the supernatant. The pellet underwent three washes with ice-cold RIPA buffer and was finally denatured in 5X SDS-sample buffer at 37 °C for 30 minutes.

An equal quantity of protein lysates was electrophoresed on a precast Novex NuPAGE 4-12% Bis-Tris SDS-polyacrylamide gel using the XCell SureLock electrophoresis system (Invitrogen). Resolved proteins were wet-transferred to a nitrocellulose membrane (Millipore) using a Trans-Blot Turbo transfer system (BioRad). After 1-hour blocking with 5% bovine serum albumin (BSA, Tocris, United Kingdom) buffer at room temperature, the membrane was then immunoblotted using an anti-SDR42E1-tag rabbit monoclonal antibody (1:500 dilution, WG3329739B, Invitrogen) for 3 days at 4 °C in 5% BSA buffer. Beta-actin levels were determined as loading controls using a monoclonal anti-beta-actin mouse antibody (1:5000 dilution, Sigma-Aldrich Cat# A5441, RRID: AB_476744). Following six washes with 1X Tris Buffered Saline with Tween 20 (TBST), membranes were visualized under chemiluminescence using peroxidase IgG fraction monoclonal anti-mouse IgG (H+L) secondary antibody HRP conjugate (1:5000 dilution, Thermo Fisher Scientific Cat# 62-6520, RRID: AB_2533947) or anti-rabbit IgG light chain specific (1:5000 dilution, Jackson ImmunoResearch Labs Cat# 211-032-171, RRID: AB_2339149, United Kingdom) with Pierce ECL Western blotting substrate (32106, Thermo Scientific).

2.9. RNA extraction and RT-qPCR

To assess the mRNA expression in the *SDR42E1*-edited cells, total RNA was isolated from cell lines using 1 mL TRIzol reagent (Invitrogen) according to the manufacturer's guidelines. RNA quantity and integrity were evaluated spectrophotometrically by a NanoDrop 8000 (ND-8000-GL, Thermo Scientific) and agarose gel electrophoresis. Subsequently, 2 μg of total RNA were reverse transcribed into complementary-DNA (cDNA) using the High-capacity cDNA reverse transcriptase kit (Applied Biosystems, United States) following the recommendations.

For human gene expression analysis, each 20 µL RT-qPCR reaction utilized 5 µl of the 1:5 diluted cDNA on a Quant Studio 6 Flex System (Thermo Fisher Scientific) with PowerUp SYBR Green Master Mix (Applied Biosystems). Primers designed to amplify target genes are presented in Table S1. The thermal cycling conditions included a denaturation step at 95℃ for 10 minutes, followed by 50 cycles at 95℃ for 15 seconds and 60℃ for 1 minute, concluding with a melt curve analysis. The housekeeping gene human beta-ACTIN served as an internal control to normalize variations in total RNA expression levels across each sample. Relative mRNA expression levels were assessed through comparative threshold cycle (CT) analysis, and the fold changes were calculated by the $2^{-\Delta\Delta CT}$ (delta-delta cycle threshold) method, referencing the average ΔCT value of wild-type controls and reference genes [18].

2.10. RNA sequencing

To identify differentiated genes in the relevant pathway, equivalent amounts of high-quality RNA pools (100–200 ng) were precipitated in a 75% ethanol solution and shipped to Macrogen Inc. (http://macrogen.com, Korea) for total RNA-sequencing library construction and next-generation sequencing. In summary, the RNA sequencing libraries were prepared with the TruSeq Stranded Total RNA Library Prep Kit (1000000040499, Illumina), following the instructions. Following PCR enrichment, the libraries were quality-checked and quantified before sequencing on the Illumina NovaSeq 6000 platform, yielding 101 bp paired-end reads. Raw sequence reads were trimmed to eliminate contaminant DNA, PCR duplicates, and adaptor sequences. Reads with a quality below Q20 were filtered using CLC Genomics Workbench version 22.0.2 software. The same software was employed for the assembly and alignment of all paired-end reads against the latest version of the human genome (*Homo sapiens*, GRCh38). Genes with average raw counts below 10 were excluded.

Differential expression of the genes (DEG) between control and knocked-out samples was defined as significant if the expression change was ≥ 2-fold, accompanied by an adjusted P-value < 0.05. This assessment was conducted and visualized using the R package DESeq2 version 1.34 with default parameters [19]. For gene expression heatmaps based on RNA-sequencing data, $\log_2$ of the fold-change maximum likelihood estimate (lfcMLE) values, and $-\log_{10}$ of false discovery rate (FDR, an adjusted P-value) values were used and visualized by R package pheatmap version 1.0.2. The Pathview R package was used to visualize Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of related genes [20].

We conducted Gene Set Enrichment (GSE) analysis as outlined in previous descriptions [21], against *Homo sapiens* gene sets from the Enrichment map with gene lists in descending order based on lfcMLE, the unshrunk $\log_2$ fold change generated by DESeq2. Functional annotation analyses were performed using the R package ClusterProfiler version 4.2.2, an enrichment analysis tool renowned for its comprehensive visualization capabilities that offers detailed insights into the collective functions of the input genes, encompassing Gene ontology (GO), KEGG, and enrichment analyses with default settings. The Drug Signatures Database for Gene Set Analysis (DSigDB, http://biotechlab.fudan.edu.cn/database/drugsig/) [22] was employed to identify candidate drugs associated with the DEG. Access to DSigDB was facilitated via ClusterProfiler [23].

To validate the RNA-sequencing findings, RT-qPCR was performed on the same RNA extracts, targeting five representative genes: *SDR42E1*, alkaline phosphatase placental (*ALPP*), solute carrier family 7A5 (*SLC7A5*), serine protease inhibitor B2 (*SERPINB2*), and 7-dehydrocholesterol reductase (*DHCR7*) (Table S1). Furthermore, we conducted a comparative analysis between the RNA-sequencing results from *SDR42E1* Knockout HaCat cells and common genes associated with Vitamin D traits in the NHGRI-EBI GWAS Catalog (EFO_0004631) with a statistical significance threshold of P-value $< 5.0 \times 10^{-8}$, as released on January 09, 2024, and accessed on January 27, 2024 [24]. To enhance the robustness of our analysis, we also examined genetic markers located within a 250-kilobase region upstream and downstream of the reported GWAS signals, thereby enabling a comprehensive identification of significant variants associated with vitamin D.

2.11. Proteomics

Following the whole-cell extraction procedure mentioned earlier, protein samples underwent preparation through incubation at 37°C for 30 minutes. Subsequently, 50 µg of each protein sample, encompassing six samples from three biological replicates of homozygous *SDR42E1* knockout and HaCat controls, was loaded onto a precast Novex

NuPAGE 4-12% Bis-Tris SDS-polyacrylamide gel and run at 100-voltage for approximately 60 minutes. For protein band visualization, PageBlue Protein Staining Solution was applied overnight at 4°C, followed by a 10-minute wash with sterilized distilled water. Each gel lane was individually cut, placed in tubes, and kept at 4°C for further mass spectrophotometry analysis, as detailed previously [25].

Briefly, dithioerythritol and s-carbamidomethylation with iodoacetamide were included as reducing agents before in-gel tryptic digestion. Following gel piece washing, sequencing-grade modified porcine trypsin was added for rehydration and incubated overnight at 37°C. Extracted and desalted peptides were loaded onto a nanoflow UPLC system and separated using a gradient elution solvent. The resulting peptides were then analyzed using an Orbitrap Fusion Tribrid mass spectrometer (Thermo Scientific). Relative protein abundance was calculated using precursor ion areas from non-conflicting unique peptides. Data analysis was then conducted using Progenesis QI (Version 2.2., Waters), Mascot Daemon (version 2.6.0, Matrix Science), and R packages Limma (Linear Models for Microarray Data) version 3.56.2. Benjamini–Hochberg approach was used to convert Student's t-tests-derived *P*-values to multiple test-corrected q-values, with a cut-off of <0.05. Enrichment analysis and functional annotations were performed for the ranked genes based on Limma-generated *P*-values against *Homo sapiens* protein set using ClusterProfiler with standard settings.

2.12. Vitamin D quantification

Vitamin D concentrations in the cell lysates of *SDR42E1* knockouts and wild-type HaCat cells were measured using a Human Vitamin D Enzyme-linked Immunosorbent Assay (ELISA) kit (MY BIOSOURCE, MBS735897, USA) following the manufacturer's instructions. Cell lysates were prepared by trypsinization, 3 times ultrasonication, and centrifugation at 1000×g for 15 minutes to remove cellular debris. The vitamin D ELISA employed a competitive enzyme immunoassay on a microtitre plate coated with a polyclonal anti-vitamin D antibody, utilizing a vitamin D-HRP conjugate to induce a color change inversely correlated with vitamin D levels (sensitivity: 0.1 ng/mL). Vitamin D measurements were made spectrophotometrically at 450nm in a microplate reader, with triplicate samples and a standard curve for interpolation.

2.13. Immunofluorescence

Human HaCat cells were plated on poly-L-lysine-coated coverslips in a 12-well plate. At 40-60% confluency, cells were transiently transfected with 5-microgram SDR42E1-HA tag plasmids (RRID: Addgene_55182, Addgene) using Lipofectamine 3000 (2293283, Thermo Fisher Scientific) following the manufacturer's protocol for 24 hours. After fixation with 4% paraformaldehyde, cells were blocked with 5% BSA in DPBS with 0.1% Tween-20 for 60 minutes at room temperature. Following three DPBS washes, cells were probed with primary antibody overnight at 4°C, washed three times with DPBS, and then incubated with secondary and tertiary antibodies for 60 minutes each at room temperature.

Primary antibodies were diluted at 1:100, including rabbit αHA-tag polyclonal antibody (SG77, Thermo Fisher Scientific Cat# 71-5500, RRID: AB_2533988), mouse monoclonal anti-golgi 58K antibody (Sigma-Aldrich Cat# G2404, RRID: AB_477002), HSP60 (Heat Shock Protein 60) recombinant rabbit monoclonal antibody-mitochondrial marker (HSPD1-2206R, Thermo Scientific), and rabbit calreticulin-ER marker IgG polyclonal antibody (Thermo Fisher Scientific Cat# PA5-80402, RRID: AB_2787722). For secondary antibodies, biotinylated goat anti-rabbit IgG (H+L) antibody (65-6140, Thermo Scientific) and goat anti-rabbit IgG antibody (H+L), biotinylated (Vector Laboratories Cat# BA-1000, RRID: AB_2313606) were used at a 1:250 dilution. The tertiary antibody used was Streptavidin Alexa Fluor 488 conjugate at a 1:250 dilution (Molecular Probes Cat# S32354, RRID: AB_2315383). Nuclei were stained with Hoechst (33258, Invitrogen) at a concentration of 1 μg/ml, and slides were mounted with ProlongTM Antifade Histomount (Thermo Fisher Scientific).

2.14. Confocal microscopy

Confocal microscopy was performed using a Nikon A1R confocal fluorescence microscope with a 100X oil immersion objective. Specific parameters included laser excitations at 405 nm for Hoechst, 488 nm for Alexa Fluor 488, and 561 nm for GFP, with detector settings optimized for each fluorophore. Images were captured with consistent gain, offset, and pinhole settings across all samples. Motor arm movements of 60 μm, 0.3 mm, and 6.0 mm were utilized to ensure consistent field selection. Image reconstruction and analysis were performed using Fiji-ImageJ software [26], ensuring uniform treatment of all images to prevent bias.

2.15. Statistical analyses

Statistical analyses were performed using GraphPad Prism version 9. An unpaired or paired two-tailed Student's t-test assessed significance between two groups, while Kruskal-Wallis one-way analysis of variance (ANOVA) with Dunn's post hoc test was used for three or more groups. A *P*-value cutoff of 0.05 determined significance. Results were presented as average ± standard deviation or average ± standard error of the mean (SEM). Experiments in each figure utilized at least three independent biological replicates for robust and reproducible data.

3. Results

3.1. Bioinformatic screening reveals high SDR42E1 expression in skin and intestinal epithelial cells

The GTEx database was utilized as an initial resource for analyzing the mRNA expression of *SDR42E1* across different tissues under normal physiological conditions. Our analysis revealed highest expression of *SDR42E1* in sun-exposed and non-exposed skin, followed by the esophagus with mean transcripts per million (TPM) values of 11.93, 11.62, and 5.08, respectively (Figure S1a). To explore the potential functional role of *SDR42E1* in human tissues and cell lines, we evaluated its gene expression using the RNA-sequencing public resource ARCHS4. Our analysis unveiled that *SDR42E1* exhibits the highest expression in intestinal epithelial cells (TPM = 8.7) and skin keratinocytes (TPM = 8.4) compared to other cell types present in intestinal and skin tissues (Figure S1 b). Additionally, we observed a significantly higher expression of *SDR42E1* in HaCat cells (TPM = 11.2), a spontaneously transformed aneuploid keratinocyte line from adult human skin biopsies, and the HCT116 cell line (TPM = 10), a human colorectal carcinoma cell line with a Kirsten rat sarcoma (*KRAS*) mutation initiated from an adult male (Figure S1 c).

3.2. Transcriptomic profiling identifies extensive alterations in gene expressions and pathways in the *SDR42E1* knockout model

To investigate the function of *SDR42E1*, we targeted the surrounding region of the nonsense variant with 3 different guide RNAs in HaCat cells using the LentiCRISPR v2-sgRNA system. Successful gene editing at the *SDR42E1* locus was illustrated in Figure 2 a and b, affirming the efficacy of the sgRNAs. Employing the T7E1 mismatch cleavage assay, we confirmed multiple homozygous and heterozygous gene edits through the detection of undigested and endonuclease-digested bands (Figure S2c). Subsequently, 45 gene-edited clones were subjected to genotype validation through Sanger sequencing to ascertain the specificity of the induced modifications. Interestingly, DNA sequencing of Clone 32 revealed a frame-shifting insertion of a (T) nucleotide at the Cas9 cleavage site, introducing a premature stop codon 30 amino acids downstream of the p.Q30* GLN>*TER mutation (Figure S2 d).

In parallel, to establish a comparative baseline model for validating gene-edited cells, HaCat cells were transiently transfected with an in-house constructed plasmid carrying the entire coding sequence of wild-type SDR42E1 to amplify its protein expression. This augmentation was essential given the minimal endogenous expression of SDR42E1. Wild-type SDR42E1 protein was detected at the anticipated molecular weight of approximately 44 kiloDalton (kDa) via Western blotting (Figure 1a), predominantly localized to the cytoplasm and cellular membrane in HaCat cells (Figure 1b), excluding the mitochondria, Golgi apparatus, or endoplasmic reticulum (data not shown).

Subsequent Western blot analysis of immunoprecipitated SDR42E1 on Clone 32 confirmed complete loss of its expression (Figure 1c), aligning with RT-qPCR results (Figure 1d). Consequently, this clone was identified as an *SDR42E1* homozygous knockout and selected for further validation and functional investigations.

The investigation into the transcriptomic impacts of the homozygous knockout of *SDR42E1* in HaCat cells was initially conducted through RNA-sequencing. Employing the R package DESeq2 for differential expression analysis, 5,449 DEG were identified out of 14,025 read counts in the homozygous knockout cells, exhibiting an −log of FDR less than 0.05 compared to the wild-type controls (Figure 2). The principal component analysis (PCA) and multidimensional scaling (MDS) plots revealed a distinct separation between the *SDR42E1* knockouts and the wild-type HaCat cells (Figure 2a and b). This distinction was also evident in the heatmap displaying the expression levels of the top 100 DEG (Figure 2c), with the most statistically significant mRNA observed to be downregulated, as depicted in the volcano plot (Figure 2d).

Among the significant DEG resulting from the inactivation of *SDR42E1*, 458 were significantly upregulated, with a $\log_2$ fold change (FC) ≥ 0.3, including laminin subunit gamma 2 (*LAMC2*; FC = 1.60, FDR = 4.37 x 10^{-262}, pleckstrin homology like domain family A1 (*PHLDA1*; FC = 1.93, FDR = 1.31 x 10^{-228}), keratin 6A (*KRT6A*; FC = 1.26, FDR = 8.32 x 10^{-197}), keratin 18 (*KRT18*; FC = 1.36, FDR = 2.75 x 10^{-171}), *SERPINB2* (FC = 3.17, FDR = 2.17 x 10^{-103}), and endothelial lipase G (*LIPG*; FC = 0.75, FDR = 7.31 x 10^{-68}) (Figure 2c and d). Conversely, 1,058 were downregulated with a $\log_2$ fold change ≤ −0.3, including *ALPP* (FC = −2.96, FDR < 2.2 x 10^{-308}), keratin 19 (*KRT19*; FC = −1.60, FDR < 2.2 x 10^{-308}), DNA topoisomerase II alpha (*TOP2A*; FC = −1.29, FDR = 2.4 x 10^{-284}), *SLC7A5* (FC = −1.54, FDR = 1.96 x 10^{-215}), and solute carrier family 3A2 (*SLC3A2*; FC = −1.09, FDR = 1.17 x 10^{-171}) (Figure 2c and d). Summary statistics for top RNA-sequencing data are presented in Table S2.

To gain deeper insights into the biological functions of DEG from the *SDR42E1* knockout model, an enrichment analysis was conducted using the ClusterProfiler R package. The KEGG analysis showed that the majority of activated pathways are predominantly involved in ribosome biogenesis (FC = 2.15, adjusted *P*-value = 0.0008), interleukin 17 (*IL-17*) signaling (FC = 1.6, adjusted *P*-value = 0.0008), immune disorders (FC = 1.6, adjusted *P*-value = 0.002), cellular senescence (FC = 1.5, adjusted *P*-value = 0.004), steroid biosynthesis (FC = 1.6, adjusted *P*-value = 0.03), lipid metabolism and atherosclerosis process (FC = 1.3, adjusted *P*-value = 0.04) (Figure 3a and c). On the other hand, deactivated pathways are linked to glycosphingolipid biosynthesis (FC = −1.6, adjusted *P*-value = 0.004), cardiac muscle contraction (FC = −1.6, adjusted *P*-value = 0.008), vitamins metabolism (FC = −1.6, adjusted *P*-value = 0.009), ABC transporters (FC = −1.5, adjusted *P*-value = 0.04) (Figure 3a).

Interestingly, analysis of KEGG pathways revealed significant genetic alterations in the pathways of steroid and steroid hormone biosynthesis (Figure 3b and c). This was characterized by the upregulation of emopamil binding protein (*EBP*, FC = 0.50, FDR = 2.46 x 10^{-13}), *DHCR7* (FC = 0.42, FDR = 8.03 x 10^{-09}), cytochrome P450 (*CYP*), such as *CYP51A1* (FC = 0.63, FDR = 1.16 x 10^{-27}) and *CYP27B1* (FC = 1.08, FDR = 1.11 x 10^{-06}) (Table S2). Conversely, there was a downregulation observed in several *CYP* family genes, including *CYP26A1* (FC = −1.55, FDR = 1.06 x 10^{-08}) and *CYP24A1* (FC = −0.26, FDR = 3.8 x 10^{-02}) as well as lamin B receptor (*LBR*, FC = −0.58, FDR = 5.14 x 10^{-20}) and catechol-O-methyltransferase (*COMT*, FC = −0.23, FDR = 8.80 x 10^{-04}) (Table S2).

Drug prediction enrichment analysis of the DEG utilizing the DSigDB database, revealed therapeutic agents potentially regulated by *SDR42E1*, including dinoprostone (*P*-value = 7.37 x 10^{-10}), mifepristone (*P*-value = 1.12 x 10^{-09}), 17-Ethynyl estradiol (*P*-value = 1.67 x 10^{-07}), vitamin D3 (*P*-value = 4.39 x 10^{-06}), and medroxyprogesterone acetate (*P*-value = 1.56 x 10^{-05}) (Figure 3c and Table S3).

To corroborate findings derived from the RNA-sequencing analysis of *SDR42E1* knockout HaCat cells, RT-qPCR analysis was performed on a selected set of five representative genes utilizing the identical RNA extracts. The RT-qPCR outcomes for *ALPP*, *SDR42E1*, *SLC7A5*, *DHCR7*, and *SERPINB2* exhibited a robust concordance with the RNA-sequencing data, with a substantial Spearman correlation coefficient (R^2) of approximately 0.93 when compared to wild-type HaCat cells (Table 1). These findings collectively affirm the robustness and reliability of the RNA-sequencing analysis conducted in this study.

Furthermore, we assessed the gene replication by analyzing the DEG in the *SDR42E1* Knockout HaCat RNA-sequencing data with Vitamin D-related genes reported in the GWAS Catalog [24]. We successfully replicated a considerable number of the common genes in our *SDR42E1* knockout dataset, identifying 65 out of 248 genes. Notable instances include the low-density lipoprotein receptor (*LDLR*), endothelial lipase (*LIPG*), and involucrin (*IVL*), which exhibited significant associations in the *SDR42E1* knockout data and the GWAS Catalog, with *P*-values ranging from 9 x 10^{-2305} to 5 x 10^{-08} (Table 2).

3.3. *SDR42E1* knockout profoundly alters protein expressions and pathways involved in vitamin D regulation

Employing a label-free LC-MS/MS shotgun proteomics approach, we discerned a range of differentially expressed proteins in the HaCat cells with homozygous *SDR42E1* knockout, totaling 138 out of 1,320 proteins (Figure 4 and Table S4). In comparison to wild-type HaCat cells, the *SDR42E1* knockout cells exhibited 101 downregulated proteins and 37 upregulated proteins (Figure 4a and b). Noteworthy increases were noted in SERPINB2 (molecular mass of 46851), keratin 17 (KRT17) (molecular mass of 48361), and SERPINB1 (molecular mass of 42742), exhibiting adjusted *P*-values (q-value) of 1.90×10^{-08}, 9.91×10^{-05}, and 2.38×10^{-02}, along with $\log_2$ fold changes of 3.36, 1.37, and 1.23, respectively (Figure 4c). Simultaneously, there were significant downregulations of SLC3A2 (molecular mass = 68180), SLC7A5 (molecular mass = 55659), LIM And SH3 Protein 1 (LASP1, molecular mass = 29717) with q-values of 8.41×10^{-07}, 1.88×10^{-05}, and 2.38×10^{-02}, and $\log_2$ fold changes of −1.54, −1.77, and −1.20, respectively (Figure 4c). These findings were corroborated by RNA-sequencing analysis results.

The GSE analysis linked a wide range of differentially expressed proteins to the activation of pathways involved with the regulation of epithelial cell apoptosis (GO:1904019, FC = 1.8, adjusted *P*-value = 0.002), development of skin epidermis (GO:0098773, FC = 1.7, adjusted *P*-value = 0.008), and various processes of wound healing (GO:0042060 and GO:0009611, FC = 1.7, adjusted *P*-value = 0.02) (Figure 4d). Moreover, numerous proteins contributed to the suppression of pathways related to the formation of melanosome and pigment granules (GO:0042470 and GO:0048770), cellular response to heat (GO:0034605), and various immune-related disorders (including GO:0002250, GO:0002460, and GO:0002699), with FC of −1.7 and adjusted *P*-value of 0.016 (Figure 4d).

3.4. *SDR42E1* knockout decreases vitamin d production in HaCat cells

To evaluate the direct impact of the *SDR42E1* knockout on the levels of vitamin D in cell lysates, vitamin D ELISA was conducted. These tests demonstrated a significant reduction in the vitamin D levels to 0.52-fold ($R^2 = 0.935$) in the *SDR42E1* knockout samples compared with the wild-type controls, dropping from 78.18 ng/mL to 40.60 ng/mL, yielding a *P*-value of 0.0016 (Figure 5).

4. Discussion

The widespread prevalence of vitamin D deficiency poses a significant public health concern, associated with various serious illnesses. The intricate regulation of vitamin D levels is influenced by a multifaceted interplay of genetic and environmental factors, such as limited sunlight exposure and malnutrition [2]. Recent GWAS research has linked a new non-sense variant, rs11542462, in the uncharacterized *SDR42E1* to serum 25(OH)D levels [4-6]. The novel *SDR42E1* is believed to be engaged in multiple metabolic processes that could impact lipid metabolism and steroid hormone biosynthesis [8, 9]. Our study comprehensively characterizes, for the first time, the structure and function of the *SDR42E1* gene and its variant, exploring the potential impact on biological processes related to vitamin D.

Preliminary examination of available mRNA data for *SDR42E1* from healthy subjects revealed elevated expression levels in essential tissues involved in vitamin D synthesis and metabolism [27], notably in skin keratinocytes and the corresponding HaCat cell line. This evidence suggests that *SDR42E1* could be instrumental in the regulation of vitamin D biosynthesis in the skin. Such insights are valuable for understanding the function of *SDR42E1*, guiding the development of suitable models for further investigation. Consequently, we selected the human keratinocyte HaCat cell line as our model system, reflecting the primary site of vitamin D biosynthesis. This choice is vital for understanding the proposed involvement of *SDR42E1* in vitamin D regulation, providing a relevant and accurate context for our investigations.

In the characterization study, examining the subcellular localization of uncharacterized proteins provides significant clues about their metabolic functions. Previous research has reported that the subcellular distribution of proteins in the SDR family largely depends on their enzymatic activity, predominantly localized to mitochondria, cytoplasm, plasma membrane, and endoplasmic reticulum [8, 28]. In the current research, we specifically targeted the overexpression of SDR42E1 in enriched human HaCat cell lines to delineate its subcellular distribution. Employing targeted antibodies against SDR42E1 and staining for different cellular components, we revealed a prominent localization of SDR42E1 to the plasma membrane and cytoplasm. The plasma membrane and cytosol are critical platforms for lipid and steroid metabolic processes [29]. Our observation expands our comprehension of the roles of SDR42E1 and strongly suggests its potential involvement in regulating lipid metabolism within these critical cellular compartments.

Employing CRISPR/Cas9 technology in HaCat cells, we successfully introduced a targeted and efficient modification to the *SDR42E1*, replicating the p.Q30* GLN>*TER nonsense mutation, and mimicking a functional knock-out of the gene. We conducted extensive transcriptomic and proteomic profiling in the homozygous *SDR42E1* knockout and wild-type HaCat cells, uncovering numerous differentially expressed proteins and gene enrichments in various downstream and steroid-related pathways. Key observations were the substantial downregulation of *ALPP* and *KRT19* gene expressions in the *SDR42E1* knockout model. Alkaline phosphatase (ALP) encompasses a set of enzymes, encoded by four genes, with three being tissue-specific and one ubiquitous across various body tissues. ALP plays a crucial role in the synthesis of cellular membrane phospholipids and in the mineralization of new bone, particularly via ALPP [30]. Elevated serum ALP levels are indicative of osteomalacia, often stemming from vitamin D deficiency [31]. Kover and his colleagues initially highlighted the role of ALP as a marker for vitamin D deficiency in premature infants with rickets [32]. Subsequent studies have consistently shown an inverse relationship between ALP and serum 25(OH)D levels [33], leading to the use of increased serum ALP as a diagnostic marker for vitamin D deficiency [34]. Unfortunately, there is no evidence of a functional role for *ALPP* in the skin, despite its significant expression.

Additionally, we observed notable alterations in numerous genes previously reported in the GWAS catalog for vitamin D levels [35], following the depletion of *SDR42E1* from HaCat cells. Among these DEG were key players, including *LDLR*, *LIPG*, *IVL*, *CYP26A1*, and *DHCR7*, known for their involvement in diverse biological processes related to vitamin D regulation, including synthesis, transport, and degradation. Furthermore, we noticed that a batch of important genes, including *SERPINB2*, *SLC7A5*, *CYP3A5*, and *LBR*, underwent significant modifications in the homozygous *SDR42E1* knockout model. These observations are consistent with previous studies, emphasizing the potential role of the *SDR42E1* in modulating lipid and steroid metabolism, including vitamin D in skin keratinocytes [36, 37].

Finally, our analysis of the enriched pathways has revealed a pronounced impact on the steroid biosynthesis pathway following the inactivation of the *SDR42E1* gene in HaCat cells. Notably, the upregulation of key genes involved in vitamin D synthesis in the skin, *EBP* and *DHCR7*, exhibit an enhanced production of 7-DHC and a consequent obstruction in its conversion to vitamin D3 in the absence of SDR42E1 (Figure 6). This discovery aligns with our earlier research, where we demonstrated the robust substrate affinity of vitamin D3 and its precursors, 7-DHC and 8-DHC, towards SDR42E1 through *in silico* docking studies [8]. This affinity is further evidenced by the drug enrichment analysis of our *SDR42E1* knockout model, which shows a significant association with vitamin D3 therapy. Moreover, other key enzymes implicated in vitamin D synthesis and absorption, notably CYP27B1, CYP24A1, COMT, and ABCB1, are also significantly regulated in the absence of SDR42E1, further confirming a role of SDR42E1 in maintaining vitamin D homeostasis.

Our proteomic analysis, consistent with RNA-sequencing, demonstrated marked increases in SERPINB2, KRT17, and SERPINB1, and decreases in SLC3A2, SLC7A5, and LASP1 protein levels in the *SDR42E1* knockout model. SERPINB2, encoding the plasminogen activator inhibitor-2 (PAI-2) protein, is key in regulating blood clot breakdown while SERPINB1 is an intracellular protein that shields cells from stress-induced cytoplasmic proteases [38]. A previous study showed that vitamin D regulates the production of antithrombin by affecting the SERPIN proteins, suggesting a new pathway for antithrombotic therapy development [39]. Furthermore, vitamin D influences keratinocyte behavior and gene expression, notably *SERPIN* genes, which play a role in skin differentiation and the management of skin conditions [40]. KRT17 is also connected to various cell functions, skin disorders, and bone irregularities [41, 42]. The *SLC7A5* encodes the large neutral amino acid transporter 1 (LAT1), vital for transporting large neutral amino acids across cell membranes [43]. Recent studies emphasize the critical role of LAT1 in bone homeostasis, regulated by the vitamin D receptor [44, 45]. Additionally, vitamin D regulates LAT1 expression in the placenta, possibly aiding fetal growth in vitamin D-deficient preeclampsia through vitamin D receptor and the mTOR pathways [46]. SLC3A2, a component of CD98 glycoprotein, is essential for skin health and osteoclast formation by interacting with LAT1 and vitamin D [47]. SLC3A2 is upregulated by estrogen, promoting colon health and vitamin D synthesis, thereby preventing colorectal cancer through vitamin D receptor activation [48]. These discoveries underscore the intricate role of SDR42E1 in managing vitamin D-related pathways and health outcomes, highlighting the need for further research into its specific genetic mechanisms.

While our research provides considerable new insights into the functional role of *SDR42E1* in vitamin D biosynthesis, several areas warrant further exploration. The functions and substrates of *SDR42E1* require additional characterization to fully elucidate its involvement in various biological processes, particularly in cellular senescence and DNA repair. Moreover, our study did not investigate the potential interactions between *SDR42E1*, environmental factors, and human diseases. This aspect could benefit from further investigation using diverse cellular and animal models, as well as clinical samples. These additional studies would validate our *in vitro* findings and enhance the understanding of the physiological and pathological implications of SDR42E1 in human health.

In conclusion, our research has unveiled a pioneering characterization of the *SDR42E1* gene, elucidating its relationship with key genes involved in steroid and vitamin D biosynthesis. This study also highlights the role of *SDR42E1* in lipid metabolism, cellular aging, and DNA repair, thereby expanding our knowledge of its understanding uncharacterized cellular functions.

Data availability statemen

All data generated throughout the study have been thoroughly reviewed and included in this published article or documented in the referenced data repositories. The RNA-sequencing data is accessible on the GEO database portal (https://www.ncbi.nlm.nih.gov/geo/) under the series accession number GSE262704. Bioinformatic and statistical analyses employed publicly accessible software tools, as detailed in the main text and Methods sections. For further information about resources or reagents, and access to raw data and code, please contact the lead author, Georges Nemer (gnemer@hbku.edu.qa), upon reasonable request.

Funding

Open access funding provided by the Qatar National Library.

CRediT authorship contribution

Nagham Nafiz Hendi: Writing – original draft, Visualization, Validation, Software, Investigation, Methodology, Formal analysis, Data curation, Conceptualization. **Maria Teresa Bengoechea-Alonso**: Writing – review & editing, Resources. **Johan Ericsson**: Writing – review & editing, Resources. **Georges Nemer**: Writing – review & editing, Resources, Project administration, Supervision, Conceptualization, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The York Centre of Excellence in Mass Spectrometry was created thanks to a major capital investment through Science City York, supported by Yorkshire Forward with funds from the Northern Way Initiative, and subsequent support from EPSRC (EP/K039660/1; EP/M028127/1).

Supplementary data

The following are the Supplementary data to this article:

Supplementary Figures. Figures S1–S3.

Supplementary Tables.

500

Table 1 Correlation between RNA-sequencing and RT-qPCR findings in the *SDR42E1* knockout model.

Genes	**RT-qPCR**	**RNA-Sequencing**
SLC7A5	−1.7	−1.54
ALPP	−3.59	−2.96
DHCR7	0.27	0.42
SERPINB2	1.24	3.19
SDR42E1	−2.71	−1.15
Spearman R^2	0.93	

Spearman correlation coefficient (R^2) reveals strong correlations on the $\log_2$ (fold-change) data. Beta-actin gene was utilized as an internal control for expression normalization by geometric mean. HaCat wild-type cells were functioned as a negative control. Data present the $\log_2$ (fold-change) of three replicates, demonstrating consistent outcomes. Abbreviations: RT-qPCR, reverse transcription-quantitative polymerase chain reaction.

501

Table 2 Replication of DEG reported in the GWAS Catalog for vitamin D in the *SDR42E1* knockout model.

Gene	$\log_2$ FC	*P*-value	FDR	Gene Description
LDLR	0.979	2.297E−73	4.536E−71	Low-Density Lipoprotein Receptor; is involved in the regulation of cholesterol levels in the blood.
LIPG	0.749	4.168E−70	7.305E−68	Endothelial Lipase; plays a role in lipid metabolism.
IVL	−1.937	1.471E−50	1.517E−48	Involucrin; regulates vitamin D receptor in the epidermis.
FGFBP1	0.763	1.040E−46	9.656E−45	Fibroblast Growth Factor Binding Protein 1; modulates fibroblast growth factor activity.
CXCL8	1.306	4.189E−42	3.228E−40	Chemokine (C-X-C motif) Ligand 8; is a pro-inflammatory cytokine is involved in immune response.
HERPUD1	−1.177	6.225E−37	3.950E−35	Homocysteine-Inducible, Endoplasmic Reticulum Stress-Inducible, Ubiquitin-Like Domain 1; involves in the unfolded protein response.
RETREG3	−0.867	2.697E−35	1.637E−33	Reticulophagy Regulator 3; possibly is involved in autophagy or cell survival.
PRXL2A	−0.585	7.438E−32	3.890E−30	Peroxiredoxin-Like 2A; is involved in oxidative stress response.
SDR42E1	−1.137	1.981E−29	9.168E−28	Short Chain Dehydrogenase/Reductase Family 42E, 1; potential role in steroid metabolism.
ADAR	−0.467	3.532E−27	1.448E−25	Adenosine Deaminase Acting on RNA; is involved in RNA editing.
ZPR1	0.634	1.190E−25	4.498E−24	ZPR1 Zinc Finger; essential for cell viability and may play a role in cell proliferation.
ARNT	−0.734	2.807E−19	7.197E−18	Aryl Hydrocarbon Receptor Nuclear Translocator; is involved in response to environmental toxins.
BCL11A	−1.658	5.743E−17	1.276E−15	BAF Chromatin Remodeling Complex Subunit BCL11A; a transcription factor is involved in hematopoietic development.
TRPS1	−3.028	1.862E−15	3.668E−14	TRPS1 Transcription Repressor; is involved in skeletal development.
KIF20B	−0.403	5.962E−15	1.134E−13	Kinesin Family 20B; is involved in mitosis and cell division.
RABGAP1	−0.417	1.898E−10	2.341E−09	RAB GTPase Activating Protein 1; is involved in intracellular membrane trafficking.
GNAQ	−0.395	7.397E−10	8.351E−09	G Protein Subunit Alpha Q; a component of a signaling pathway is involved in various cell processes.
CYP26A1	−1.55	9.489E−10	1.054E−08	Cytochrome P450 Family 26 Subfamily A1; is involved in vitamins metabolism.

CARMIL1	−0.390	2.820E−09	2.936E−08	Capping Protein Regulator and Myosin 1 Linker 1; is involved in cell migration and adhesion.
ATP1B3	−0.312	3.532E−09	3.618E−08	ATPase Na+/K+ Transporting Subunit Beta 3; is involved in ion transport.
DHCR7	0.406	8.034E−09	7.869E−08	7-Dehydrocholesterol Reductase; is involved in cholesterol and steroid biosynthesis.
ZNF587B	0.456	1.458E−08	1.369E−07	Zinc Finger Protein 587B; likely a transcription factor.
ASH1L	−0.337	1.510E−08	1.415E−07	ASH1 Like Histone Lysine Methyltransferase; is involved in chromatin modification.
DOCK8	−3.042	2.127E−07	1.687E−06	Dedicator Of Cytokinesis 8; is involved in immune cell signaling and function.
GALNT2	−0.307	2.218E−07	1.754E−06	Polypeptide N-Acetylgalactosaminyltransferase 2; is involved in glycosylation.
SMYD3	−0.755	5.670E−07	4.180E−06	SET And MYND Domain Containing 3; a histone methyltransferase is involved in chromatin regulation.
KLK10	−0.394	2.198E−06	1.45E-05	Kallikrein Related Peptidase 10; is involved in proteolysis and various physiological processes.
CELSR2	−0.421	2.214E−06	1.46E-05	Cadherin EGF LAG Seven-Pass G-Type Receptor 2; is involved in cell adhesion and signaling.
ZNF680	−0.820	5.506E−06	3.37E-05	Zinc Finger Protein 680; likely functions as a transcription factor.
USP3	0.330	6.418E−06	3.88E-05	Ubiquitin Specific Peptidase 3; is involved in DNA damage response and repair.
FTO	−0.344	8.016E−06	4.77E-05	FTO Alpha-Ketoglutarate Dependent Dioxygenase; associated with body mass and obesity.
MCUB	−0.599	9.294E−06	5.46E-05	Mitochondrial Calcium Uniporter, MCUb Subunit; is involved in mitochondrial calcium uptake.
MAN2A1	−0.236	1.51E-05	8.52E-05	Mannosidase Alpha Class 2A 1; is involved in glycoprotein processing.
TRMT61A	0.419	1.88E-05	0.0001046	tRNA Methyltransferase 61A; is involved in tRNA modification.
BTBD10	0.296	3.22E-05	0.0001702	BTB Domain Containing 10; potentially is involved in neuronal survival and apoptosis.

Genes listed associated with vitamin D at *P*-values < 5E−05 and reported in the GWAS Catalog for vitamin D with *P*-values < 5E−08. Abbreviations: Log_2 FC, Log_2 Fold Change estimate; FDR, false discovery rate represents adjusted *P*-value using the Benjamini-Hochberg in DESeq2.

References

1. Hypponen, E., K.S. Vimaleswaran, and A. Zhou, *Genetic Determinants of 25-Hydroxyvitamin D Concentrations and Their Relevance to Public Health.* Nutrients, 2022. **14**(20).
2. Hendi, N.N., et al., *Genetic determinants of Vitamin D deficiency in the Middle Eastern Qatari population: a genome-wide association study.* Front Nutr, 2023. **10**: p. 1242257.
3. Ogunmwonyi, I., A. Adebajo, and J.M. Wilkinson, *The genetic and epigenetic contributions to the development of nutritional rickets.* Front Endocrinol (Lausanne), 2022. **13**: p. 1059034.
4. Sinnott-Armstrong, N., et al., *Genetics of 35 blood and urine biomarkers in the UK Biobank.* Nat Genet, 2021. **53**(2): p. 185-194.
5. Revez, J.A., et al., *Genome-wide association study identifies 143 loci associated with 25 hydroxyvitamin D concentration.* Nat Commun, 2020. **11**(1): p. 1647.
6. Qiu, S., et al., *Genetic correlation, causal relationship, and shared loci between vitamin D and COVID-19: A genome-wide cross-trait analysis.* J Med Virol, 2023. **95**(5): p. e28780.
7. Persson, B., et al., *The SDR (short-chain dehydrogenase/reductase and related enzymes) nomenclature initiative.* Chem Biol Interact, 2009. **178**(1-3): p. 94-8.
8. Hendi, N.N. and G. Nemer, *In silico characterization of the novel SDR42E1 as a potential vitamin D modulator.* J Steroid Biochem Mol Biol, 2023. **238**: p. 106447.
9. Loveland, J.L., D.B. Lank, and C. Kupper, *Gene Expression Modification by an Autosomal Inversion Associated With Three Male Mating Morphs.* Front Genet, 2021. **12**: p. 641620.
10. Bouhouche, A., et al., *Identification of the novel SDR42E1 gene that affects steroid biosynthesis associated with the oculocutaneous genital syndrome.* Exp Eye Res, 2021. **209**: p. 108671.
11. Chen, S., et al., *Gonadotropin inhibitory hormone downregulates steroid hormone secretion and genes expressions in duck granulosa cells.* Anim Reprod, 2021. **18**(2): p. e20210036.
12. Stiles, A.R., et al., *Genetic, anatomic, and clinical determinants of human serum sterol and vitamin D levels.* Proc Natl Acad Sci U S A, 2014. **111**(38): p. E4006-14.
13. Liu, W., et al., *Assays of plasma dehydrocholesteryl esters and oxysterols from Smith-Lemli-Opitz syndrome patients.* J Lipid Res, 2013. **54**(1): p. 244-53.
14. Tomita, H., et al., *7-Dehydrocholesterol-derived oxysterols cause neurogenic defects in Smith-Lemli-Opitz syndrome.* Elife, 2022. **11**.
15. Lachmann, A., et al., *Massive mining of publicly available RNA-seq data from human and mouse.* Nat Commun, 2018. **9**(1): p. 1366.
16. Sanjana, N.E., O. Shalem, and F. Zhang, *Improved vectors and genome-wide libraries for CRISPR screening.* Nat Methods, 2014. **11**(8): p. 783-784.
17. Kingston, R.E., C.A. Chen, and H. Okayama, *Calcium phosphate transfection.* Curr Protoc Cell Biol, 2003. **Chapter 20**: p. Unit 20 3.
18. Livak, K.J. and T.D. Schmittgen, *Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method.* Methods, 2001. **25**(4): p. 402-8.

19. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.
20. Luo, W. and C. Brouwer, *Pathview: an R/Bioconductor package for pathway-based data integration and visualization.* Bioinformatics, 2013. **29**(14): p. 1830-1.
21. Wu, T., et al., *clusterProfiler 4.0: A universal enrichment tool for interpreting omics data.* Innovation (Camb), 2021. **2**(3): p. 100141.
22. Wu, H., et al., *DrugSig: A resource for computational drug repositioning utilizing gene expression signatures.* PLoS One, 2017. **12**(5): p. e0177743.
23. Yoo, M., et al., *DSigDB: drug signatures database for gene set analysis.* Bioinformatics, 2015. **31**(18): p. 3069-71.
24. Sollis, E., et al., *The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource.* Nucleic Acids Res, 2023. **51**(D1): p. D977-D985.
25. Dekker, J., et al., *Spatial analysis of the ancient proteome of archeological teeth using mass spectrometry imaging.* Rapid Commun Mass Spectrom, 2023. **37**(8): p. e9486.
26. Schindelin, J., et al., *Fiji: an open-source platform for biological-image analysis.* Nat Methods, 2012. **9**(7): p. 676-82.
27. Amrein, K., et al., *Vitamin D deficiency 2.0: an update on the current status worldwide.* Eur J Clin Nutr, 2020. **74**(11): p. 1498-1513.
28. He, X.Y., et al., *Characterization and localization of human type10 17beta-hydroxysteroid dehydrogenase.* Eur J Biochem, 2001. **268**(18): p. 4899-907.
29. Hu, J., et al., *Cellular cholesterol delivery, intracellular processing and utilization for biosynthesis of steroid hormones.* Nutr Metab (Lond), 2010. **7**: p. 47.
30. Shaheen, S., S.S. Noor, and Q. Barakzai, *Serum alkaline phosphatase screening for vitamin D deficiency states.* J Coll Physicians Surg Pak, 2012. **22**(7): p. 424-7.
31. Peach, H., et al., *Value of plasma calcium, phosphate, and alkaline phosphatase measurements in the diagnosis of histological osteomalacia.* J Clin Pathol, 1982. **35**(6): p. 625-30.
32. Kovar, I., P. Mayne, and D. Barltrop, *Plasma alkaline phosphatase activity: a screening test for rickets in preterm neonates.* Lancet, 1982. **1**(8267): p. 308-10.
33. Aly, Y.F., M.A. El Koumi, and R.N. Abd El Rahman, *Impact of maternal vitamin D status during pregnancy on the prevalence of neonatal vitamin D deficiency.* Pediatr Rep, 2013. **5**(1): p. e6.
34. Misra, M., et al., *Vitamin D deficiency in children and its management: review of current knowledge and recommendations.* Pediatrics, 2008. **122**(2): p. 398-417.
35. Buniello, A., et al., *The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019.* Nucleic Acids Res, 2019. **47**(D1): p. D1005-D1012.
36. Van den Broeck, T., et al., *The role of single nucleotide polymorphisms in predicting prostate cancer risk and therapeutic decision making.* Biomed Res Int, 2014. **2014**: p. 627510.
37. Hyland, P.L., et al., *Genetic variants in sex hormone metabolic pathway genes and risk of esophageal squamous cell carcinoma.* Carcinogenesis, 2013. **34**(5): p. 1062-8.
38. Choi, Y.J., et al., *Publisher Correction: SERPINB1-mediated checkpoint of inflammatory caspase activation.* Nat Immunol, 2019. **20**(5): p. 664.

39. Toderici, M., et al., *Correction: Identification of Regulatory Mutations in SERPINC1 Affecting Vitamin D Response Elements Associated with Antithrombin Deficiency.* PLoS One, 2016. **11**(7): p. e0159987.
40. Lu, J., et al., *Transcriptional profiling of keratinocytes reveals a vitamin D-regulated epidermal differentiation network.* J Invest Dermatol, 2005. **124**(4): p. 778-85.
41. Tomic, M., et al., *Nuclear receptors for retinoic acid and thyroid hormone regulate transcription of keratin genes.* Cell Regul, 1990. **1**(12): p. 965-73.
42. Yan, X., et al., *Knockdown of KRT17 decreases osteosarcoma cell proliferation and the Warburg effect via the AKT/mTOR/HIF1alpha pathway.* Oncol Rep, 2020. **44**(1): p. 103-114.
43. El Ansari, R., et al., *The amino acid transporter SLC7A5 confers a poor prognosis in the highly proliferative breast cancer subtypes and is a key therapeutic target in luminal B tumours.* Breast Cancer Res, 2018. **20**(1): p. 21.
44. Park, G., et al., *l-Type amino acid transporter 1 in hypothalamic neurons in mice maintains energy and bone homeostasis.* JCI Insight, 2023. **8**(7).
45. Ozaki, K., et al., *The L-type amino acid transporter LAT1 inhibits osteoclastogenesis and maintains bone homeostasis through the mTORC1 pathway.* Sci Signal, 2019. **12**(589).
46. Jia, X., et al., *Vitamin D stimulates placental L-type amino acid transporter 1 (LAT1) in preeclampsia.* Sci Rep, 2022. **12**(1): p. 4651.
47. Tsumura, H., et al., *Conditional deletion of CD98hc inhibits osteoclast development.* Biochem Biophys Rep, 2016. **5**: p. 203-210.
48. Li, Q., et al., *Carnobacterium maltaromaticum boosts intestinal vitamin D production to suppress colorectal cancer in female mice.* Cancer Cell, 2023. **41**(8): p. 1450-1465 e8.

Figure 1 Transient protein and transcript expression of SDR42E1 in gene-edited and wild-type HaCat cells. **a,** Immunoblotting of protein extracts from wild-type HaCat cells (lane 1) and cells with transient overexpression of wild-type SDR42E1-HA (44 kDa; lane 2, indicated by a red arrow) was performed using rabbit SDR42E1-tag polyclonal antibody at a 1:1000 dilution. Ponceau-S Red staining (PonS) served as a loading control. **b,** HaCat cells with wild-type SDR42E1 expression (green) were stained overnight with rabbit αHA-tag polyclonal antibody at a 1:100 dilution, with nuclear localization revealed by Hoechst (blue). Cells were transiently transfected with a 5-microgram SDR42E1-HA tagged plasmid for 24 hours. **c,** WB analysis of immunoprecipitated SDR42E1 in whole protein lysate derived from the gene-edited HaCat cells of clone 32 (SDR42E1-KO-32) revealed the absence of a 44 kDa band corresponding to SDR42E1 (highlighted by a red arrow) using rabbit SDR42E1 polyclonal antibody (PA5-53156, Invitrogen). Before the IP experiments, one-tenth of total lysates were subjected to the respective WB as input controls using an anti-beta-actin mouse antibody (A5441, Sigma, 1:5000 dilution). Ctrl-Cas9 is an untargeted sgRNA-Cas9 vector in HaCat as a negative control. **d,** RT-qPCR analysis revealed significantly decreased *SDR42E1* transcript expression in *SDR42E1*-KO-32 compared to controls (Ctrl-Cas9). The relative expression level of *SDR42E1* was normalized by the internal control β-actin. Data represent the mean ± standard deviation of three replicates, with similar results and significant differences relative to Cas9 control were analyzed by t-test with $p<0.0001$ (****).

Figure 2 Extensive alterations in gene expressions in the *SDR42E1* knockout model

a, A PCA plot demonstrates the clustering of three biological replicates of wild-type HaCat controls (C; in rose) and *SDR42E1* homozygous knockouts (Hom, in blue), through the major principal components of the regularized log-transformed counts. **b,** An MDS plot illustrates the correlation between $\log_2$ fold change and the mean of the normalized counts in *SDR42E1* knockouts, with significant DEG highlighted in blue (adjusted P-value < 0.05). **c**, A volcano plot depicts significant gene expression changes in *SDR42E1* knockouts compared to wild-type controls. The X-axis displays the $\log_2$ fold change (FC), with upregulated genes to the right and downregulated to the left, while the Y-axis represents the false discovery rate (FDR). Points represent individual genes with detectable expression changes, meeting the criteria of an adjusted P-value < 0.05 and a $\text{Log}_2\text{FC} > 1$, with the top 20 most significantly altered genes labeled. **d,** A cluster heatmap shows the Z-scores of regularized log-transformed counts for the top 100 DEG, with blue indicating lower and red indicating higher expression, highlighting the distinct separation of sample conditions.

The X- and Y-axes are labeled with sample names and DEGs, respectively. All analyses were conducted and visualized using R/DESeq2 and Pheatmap.

Figure 3 Extensive alterations in gene pathways in the *SDR42E1* knockout model

a, A dot plot illustrates enriched KEGG pathways for DEG. The Y-axis represents the KEGG pathways; the X-axis represents the ratio of the genes enriched in the KEGG pathway. A KEGG pathway diagram, enhanced via the R/Pathview package, shows the expression profiles of genes involved in **b**, steroid biosynthesis and **c**, in steroid hormone biosynthesis. Red indicates genes upregulated, while green denotes genes downregulated by the *SDR42E1* knockout. **d,** A bar plots of signature drugs associated with *SDR42E1* knockout through DSigDB. The color and size of dots and bars reflect the significance and count of DEG linked to KEGG pathways and drugs, respectively. *P*-value adjusted (p. adjust) < 0.05 was used as the threshold to select KEGG terms.

Figure 4 Extensive alterations in protein expressions and pathways in the *SDR42E1* knockout model

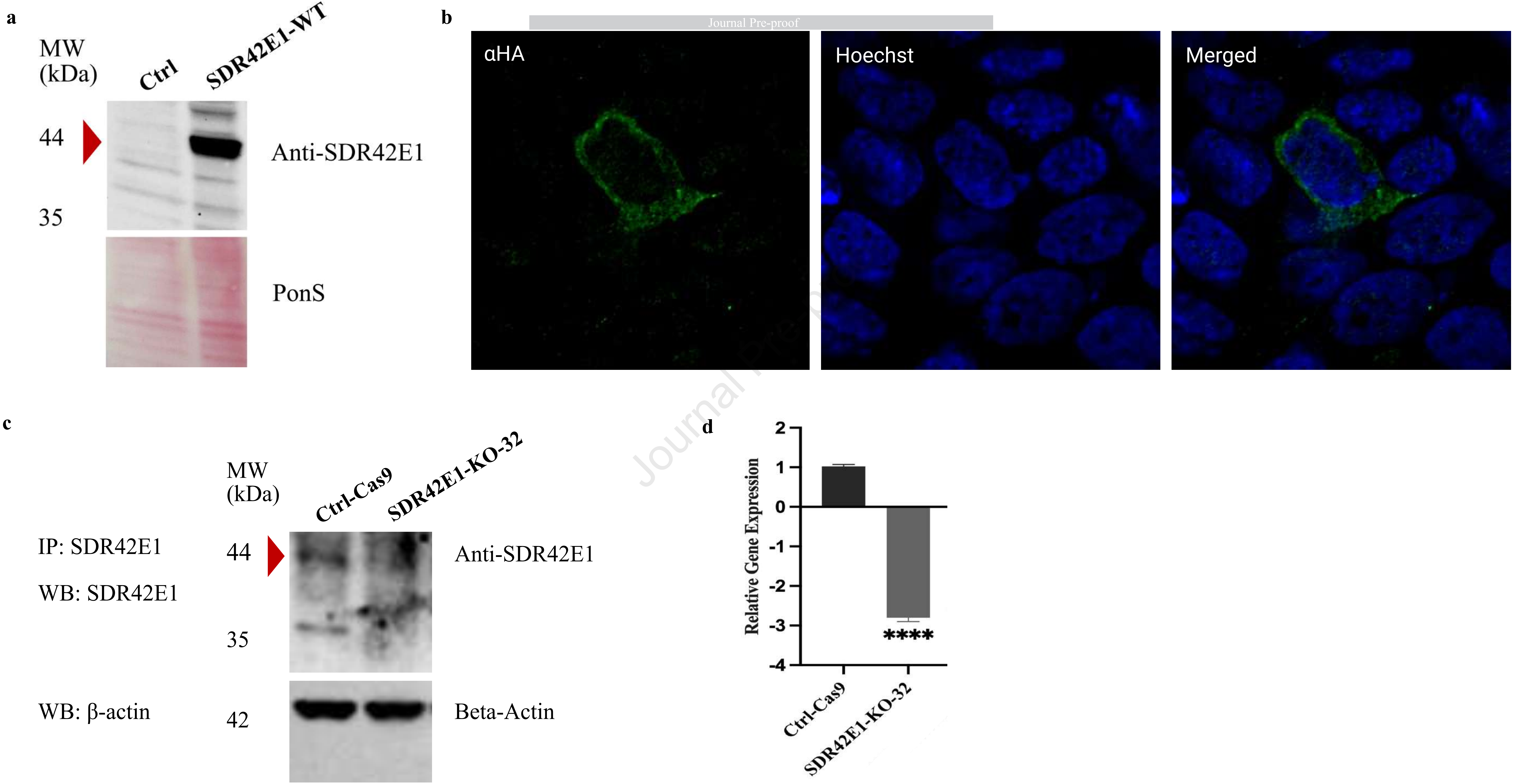
a, A PCA plot reveals the clustering of three biological replicates of wild-type HaCat controls (C; in rose) and *SDR42E1* homozygous knockouts (Hom, in blue), based on the major components of regularized log-transformed counts. **b,** A heatmap displays the Z-scores of log-transformed counts for the top 100 differentially expressed proteins, using blue to indicate lower and red for higher expression, highlighting clear separation between samples. Axes are labeled with sample names and proteins. **c,** A volcano plot compares the proteomic data of homozygous SDR42E1 knockout cells to wild-type HaCat controls. The X-axis shows $\log_2$ fold change (FC), with significant upregulation to the right (greater than 1) and downregulation to the left (less than -1). The Y-axis shows the $-\log$ of the false discovery rate (FDR), marking significance below 0.05. The top 20 significantly altered proteins are highlighted. **d,** A GSE dot plot presents pathway enrichment analysis of the differentially expressed proteins in *SDR42E1* homozygous knockouts. These analyses were conducted and visualized using R/Limma, Pheatmap, and ClusterProfiler.

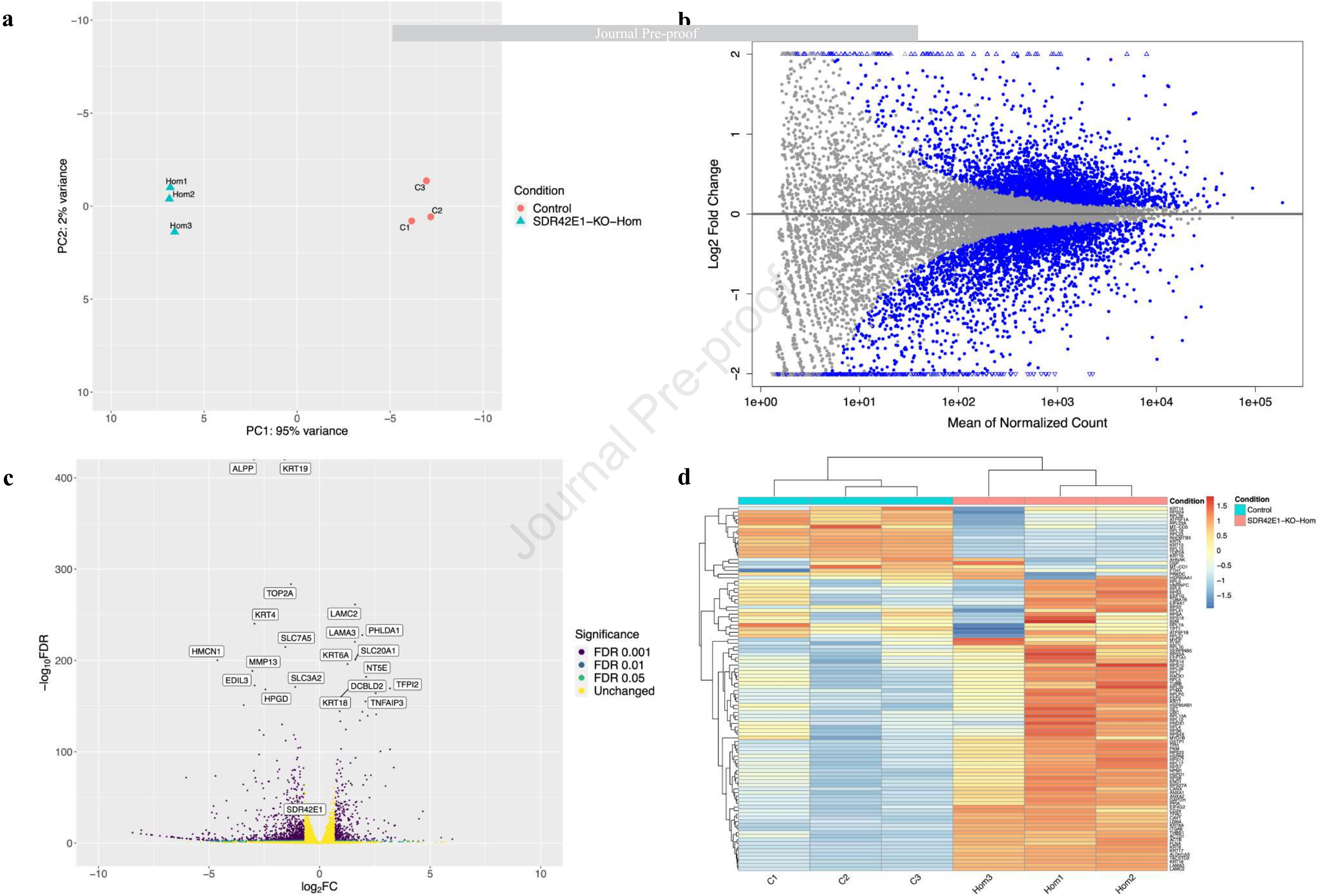
Figure 5 Decreased vitamin D levels in the *SDR42E1* knockout model.

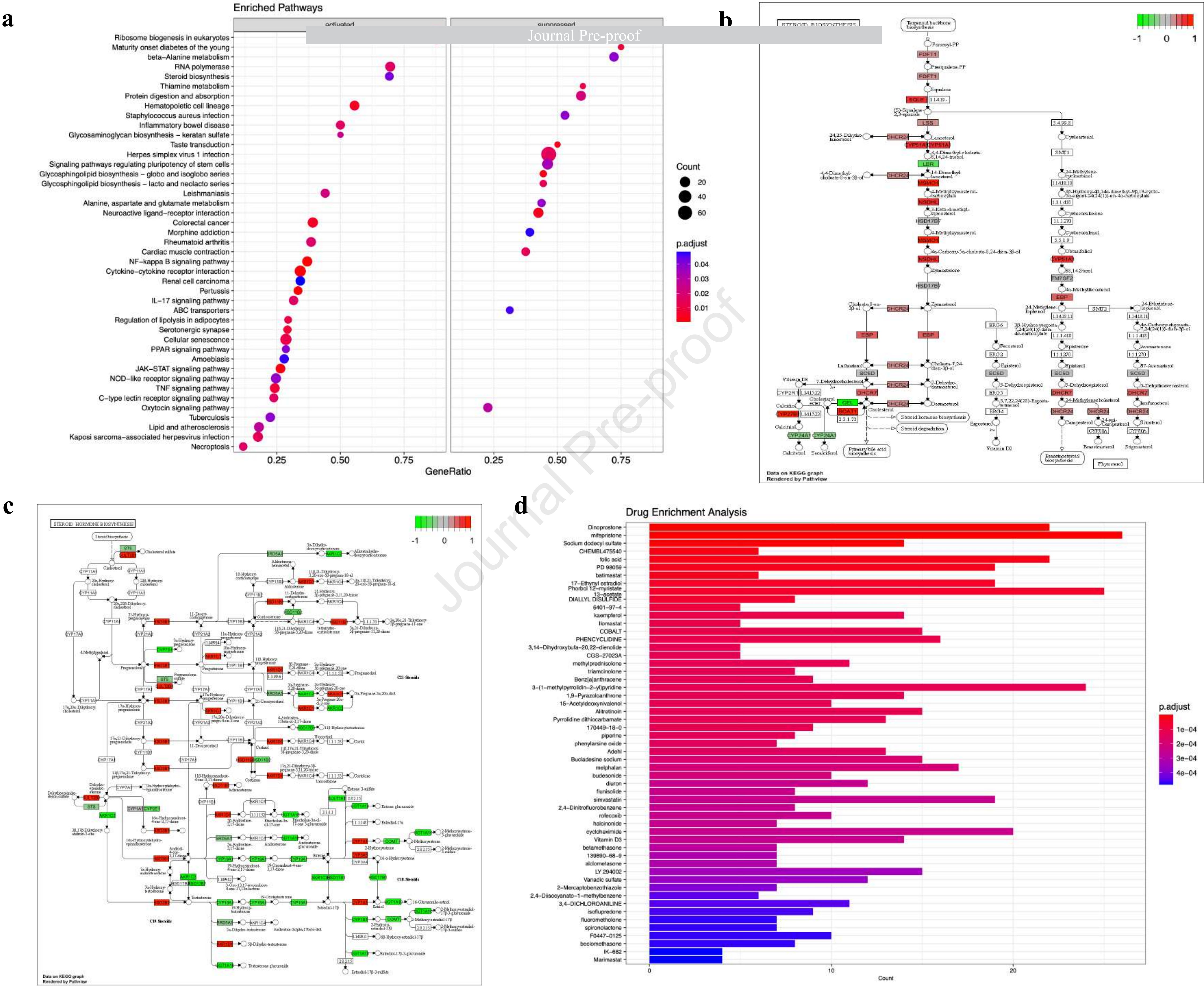
Vitamin D levels in the *SDR42E1* knockout model was measured with a vitamin D ELISA assay. Compared to wild-type HaCat cells, the knockout model showed significantly reduced vitamin D levels across three replicated experiments (**; *P*-value < 0.01).

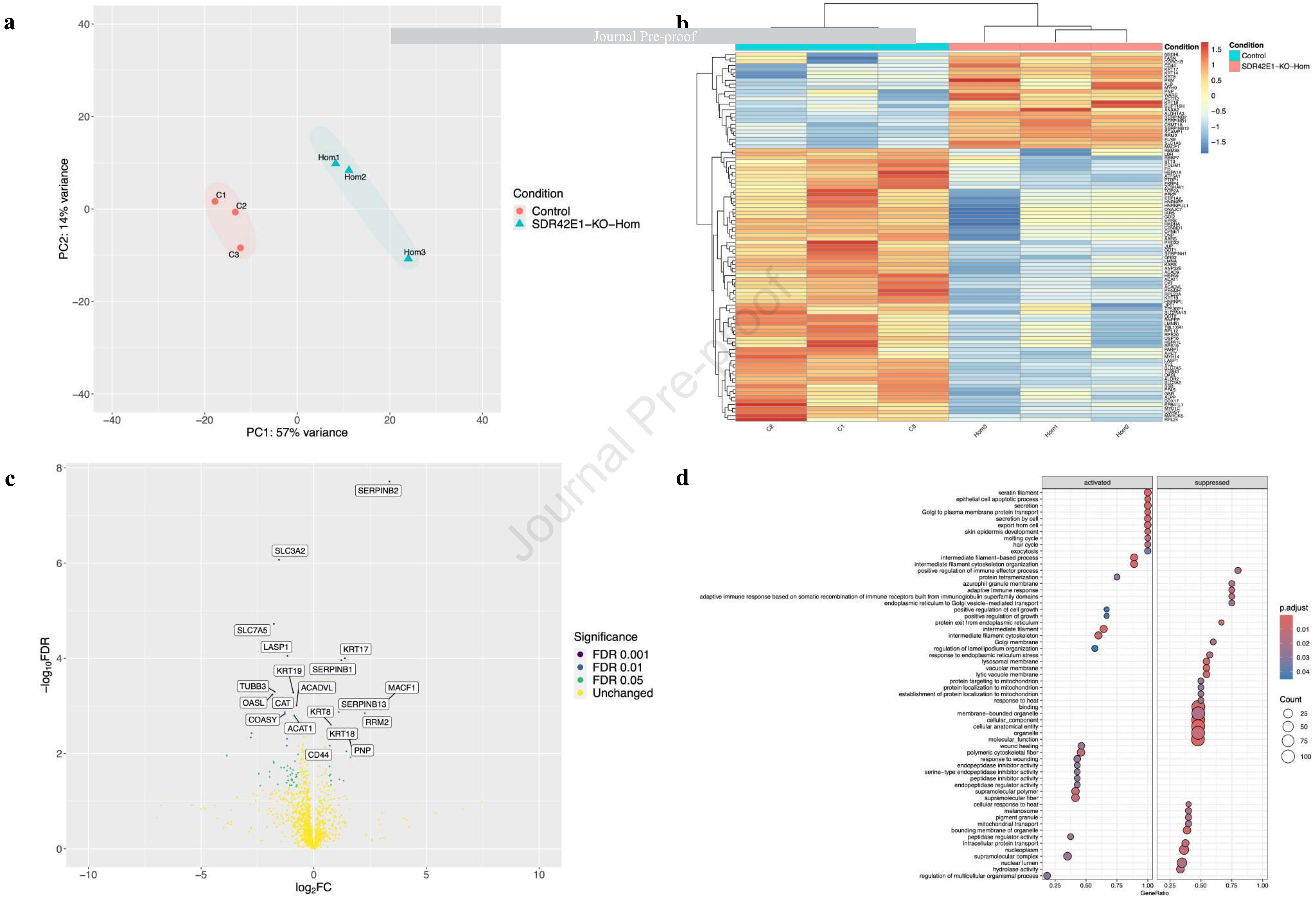
Figure 6 Potential role of *SDR42E1* in Vitamin D Biosynthesis and Regulation.

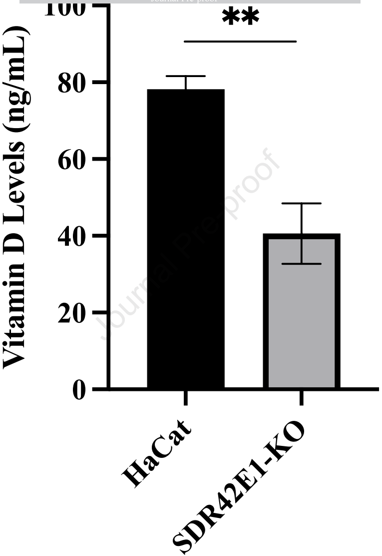
The pathway illustrates the influence of SDR42E1 absence in vitamin D skin synthesis from 7-DHC upon solar ultraviolet B (UVB) exposure. To boost vitamin D levels, the body increases the conversion of 8-DHC or cholesterol to 7-DHC by upregulating enzymes EBP or DHCR7. Intestinal absorption of vitamin D is also improved by the upregulation of ABCB1. The liver then enhances the conversion to 25(OH)D by upregulating CYP27A1 or CYP3A4, and the kidneys increase activation to 1,25-dihydroxyvitamin D via CYP27B1 upregulation, regulating related-gene expressions via vitamin D receptor (VDR)/retinoid-X receptor (RXR) complex. The inactivation and secretion process, usually facilitated by the CYP24A1 enzyme, is also diminished. Red indicates proteins upregulated, while green denotes proteins downregulated by the SDR42E1 absence. (?) indicates our proposed SDR42E1 involvement in vitamin D biosynthesis. 2D chemical structures obtained from PubChem: https://pubchem.ncbi.nlm.nih.gov. Generated with BioRender.com.

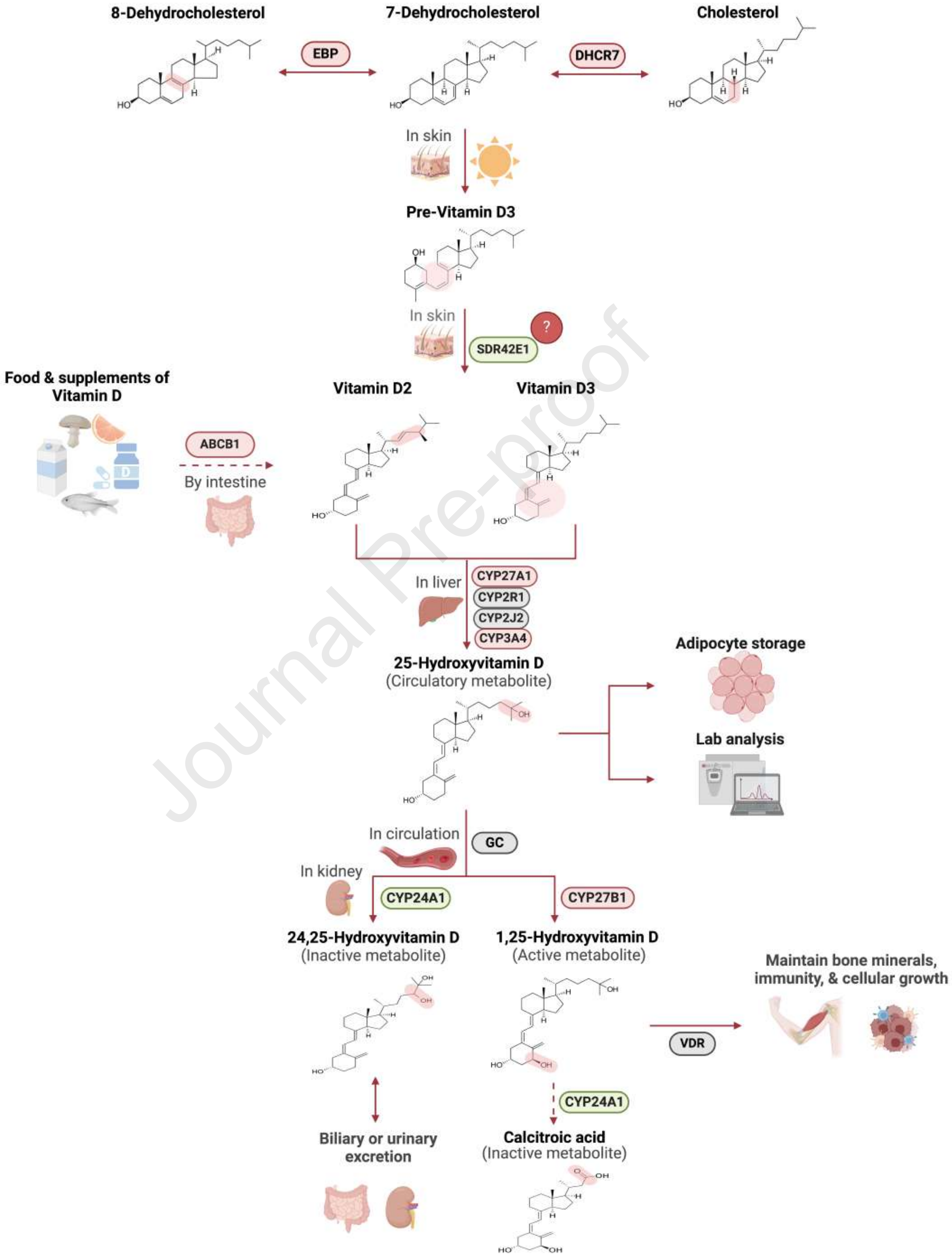

a
MW (kDa)
Ctrl
SDR42E1-WT
44
35
Anti-SDR42E1
PonS
b
αHA
Hoechst
Merged
c
MW (kDa)
Ctrl-Cas9
SDR42E1-KO-32
IP: SDR42E1
WB: SDR42E1
44
35
Anti-SDR42E1
WB: β-actin
42
Beta-Actin
d
Relative Gene Expression
2
1
0
-1
-2
-3
-4
****
Ctrl-Cas9
SDR42E1-KO-32


a
PC1: 95% variance
PC2: 2% variance
Hom1
Hom2
Hom3
C1
C2
C3
Condition
Control
SDR42E1−KO−Hom
b
Log2 Fold Change
Mean of Normalized Count
1e+00
1e+01
1e+02
1e+03
1e+04
1e+05
c
ALPP
KRT19
TOP2A
KRT4
LAMC2
LAMA3
PHLDA1
SLC7A5
HMCN1
KRT6A
SLC20A1
MMP13
NT5E
EDIL3
SLC3A2
DCBLD2
TFPI2
HPGD
KRT18
TNFAIP3
SDR42E1
-log10FDR
log2FC
Significance
FDR 0.001
FDR 0.01
FDR 0.05
Unchanged
d
Condition
Control
SDR42E1−KO−Hom
C1
C2
C3
Hom3
Hom1
Hom2


a
Enriched Pathways
activated
suppressed
Ribosome biogenesis in eukaryotes
Maturity onset diabetes of the young
beta−Alanine metabolism
RNA polymerase
Steroid biosynthesis
Thiamine metabolism
Protein digestion and absorption
Hematopoietic cell lineage
Staphylococcus aureus infection
Inflammatory bowel disease
Glycosaminoglycan biosynthesis − keratan sulfate
Taste transduction
Herpes simplex virus 1 infection
Signaling pathways regulating pluripotency of stem cells
Glycosphingolipid biosynthesis − globo and isoglobo series
Glycosphingolipid biosynthesis − lacto and neolacto series
Leishmaniasis
Alanine, aspartate and glutamate metabolism
Neuroactive ligand−receptor interaction
Colorectal cancer
Morphine addiction
Rheumatoid arthritis
Cardiac muscle contraction
NF−kappa B signaling pathway
Cytokine−cytokine receptor interaction
Renal cell carcinoma
Pertussis
IL−17 signaling pathway
ABC transporters
Regulation of lipolysis in adipocytes
Serotonergic synapse
Cellular senescence
PPAR signaling pathway
Amoebiasis
JAK−STAT signaling pathway
NOD−like receptor signaling pathway
TNF signaling pathway
C−type lectin receptor signaling pathway
Oxytocin signaling pathway
Tuberculosis
Lipid and atherosclerosis
Kaposi sarcoma−associated herpesvirus infection
Necroptosis
0.25
0.50
0.75
0.25
0.50
0.75
GeneRatio
Count
20
40
60
p.adjust
0.04
0.03
0.02
0.01
b
STEROID BIOSYNTHESIS
-1
0
1
Data on KEGG graph
Rendered by Pathview
c
STEROID HORMONE BIOSYNTHESIS
-1
0
1
C21-Steroids
C19-Steroids
C18-Steroids
Data on KEGG graph
Rendered by Pathview
d
Drug Enrichment Analysis
Dinoprostone
mifepristone
Sodium dodecyl sulfate
CHEMBL475540
folic acid
PD 98059
batimastat
17−Ethynyl estradiol
Phorbol 12−myristate 13−acetate
DIALLYL DISULFIDE
6401−97−4
kaempferol
Ilomastat
COBALT
PHENCYCLIDINE
3,14−Dihydroxybufa−20,22−dienolide
CGS−27023A
methylprednisolone
triamcinolone
Benz[a]anthracene
3−(1−methylpyrrolidin−2−yl)pyridine
1,9−Pyrazoloanthrone
15−Acetyldeoxynivalenol
Alitretinoin
Pyrrolidine dithiocarbamate
170449−18−0
piperine
phenylarsine oxide
Adehl
Bucladesine sodium
melphalan
budesonide
diuron
flunisolide
simvastatin
2,4−Dinitrofluorobenzene
rofecoxib
halcinonide
cycloheximide
Vitamin D3
betamethasone
139890−68−9
alclometasone
LY 294002
Vanadic sulfate
2−Mercaptobenzothiazole
2,4−Diisocyanato−1−methylbenzene
3,4−DICHLOROANILINE
isoflupredone
fluorometholone
spironolactone
F0447−0125
beclomethasone
IK−682
Marimastat
0
10
20
Count
p.adjust
1e−04
2e−04
3e−04
4e−04


a
PC1: 57% variance
PC2: 14% variance
C1
C2
C3
Hom1
Hom2
Hom3
Condition
Control
SDR42E1−KO−Hom
b
Condition
Control
SDR42E1−KO−Hom
1.5
1
0.5
0
−0.5
−1
−1.5
C2
C1
C3
Hom3
Hom1
Hom2
c
−log10FDR
log2FC
SERPINB2
SLC3A2
SLC7A5
LASP1
KRT17
KRT19
SERPINB1
TUBB3
ACADVL
MACF1
OASL
CAT
SERPINB13
COASY
KRT8
RRM2
ACAT1
KRT18
PNP
CD44
Significance
FDR 0.001
FDR 0.01
FDR 0.05
Unchanged
d
activated
suppressed
keratin filament
epithelial cell apoptotic process
secretion
Golgi to plasma membrane protein transport
secretion by cell
export from cell
skin epidermis development
molting cycle
hair cycle
exocytosis
intermediate filament−based process
intermediate filament cytoskeleton organization
positive regulation of immune effector process
protein tetramerization
azurophil granule membrane
adaptive immune response
adaptive immune response based on somatic recombination of immune receptors built from immunoglobulin superfamily domains
endoplasmic reticulum to Golgi vesicle−mediated transport
positive regulation of cell growth
positive regulation of growth
protein exit from endoplasmic reticulum
intermediate filament
intermediate filament cytoskeleton
Golgi membrane
regulation of lamellipodium organization
response to endoplasmic reticulum stress
lysosomal membrane
vacuolar membrane
lytic vacuole membrane
protein targeting to mitochondrion
protein localization to mitochondrion
establishment of protein localization to mitochondrion
response to heat
binding
membrane−bounded organelle
cellular_component
cellular anatomical entity
organelle
molecular_function
wound healing
polymeric cytoskeletal fiber
response to wounding
endopeptidase inhibitor activity
serine−type endopeptidase inhibitor activity
peptidase inhibitor activity
endopeptidase regulator activity
supramolecular polymer
supramolecular fiber
cellular response to heat
melanosome
pigment granule
mitochondrial transport
bounding membrane of organelle
peptidase regulator activity
intracellular protein transport
nucleoplasm
supramolecular complex
nuclear lumen
hydrolase activity
regulation of multicellular organismal process
GeneRatio
p.adjust
0.01
0.02
0.03
0.04
Count
25
50
75
100


Vitamin D Levels (ng/mL)
100
80
60
40
20
0
**
HaCat
SDR42E1-KO


8-Dehydrocholesterol
7-Dehydrocholesterol
Cholesterol
EBP
DHCR7
In skin
Pre-Vitamin D3
In skin
SDR42E1
?
Food & supplements of Vitamin D
ABCB1
By intestine
Vitamin D2
Vitamin D3
In liver
CYP27A1
CYP2R1
CYP2J2
CYP3A4
25-Hydroxyvitamin D
(Circulatory metabolite)
Adipocyte storage
Lab analysis
In circulation
GC
In kidney
CYP24A1
CYP27B1
24,25-Hydroxyvitamin D
(Inactive metabolite)
1,25-Hydroxyvitamin D
(Active metabolite)
VDR
Maintain bone minerals, immunity, & cellular growth
CYP24A1
Biliary or urinary excretion
Calcitroic acid
(Inactive metabolite)

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: