The Zebrafish udu Gene Encodes a Novel Nuclear Factor and Is Essential for Primitive Erythroid Cell Development

Running head: Erythropoiesis in Zebrafish udu Mutants

Yanmei Liu1,2, Linsen Du2, Motomi Osato3, Eng Hui Teo1, Feng Qian1,2, Hao Jin1,2, Fenghua Zhen1,2, Jin Xu1,2, Lin Guo2,4, Honghui Huang2,4, Jun Chen4, Robert Geisler5, Yun-Jin Jiang6, Jinrong Peng2,4* and Zilong Wen1,2*

1Laboratory of Molecular & Developmental Immunology, Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, Singapore 138673

2Department of Biological Sciences, the National University of Singapore, Singapore 1176083

3Laboratory of molecular oncology, Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, Singapore 138673

4Functional Genomics Laboratory, Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, Singapore 138673

5Max Planck Institute fuer Entwicklungsbiologie, Tuebingen, Germany

6Laboratory of Developmental Signalling & Patterning, Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, Singapore 138673

*Correspondence should be addressed to:
Wen Z.L.
Laboratory of Molecular & Developmental Immunology, Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, Singapore 138673
Tel: 65-6586-9725
Fax: 65-6779-1117
Email: zilong@imcb.a-star.edu.sg

Peng J.R.
Functional Genomics Laboratory, Institute of Molecular and Cell Biology, 61 Biopolis Drive, Proteos, Singapore 138673
Tel: 65-6586-9729
Fax: 65-6779-1117
Email: pengjr@imcb.a-star.edu.sg

Copyright © 2007 American Society of Hematology
Hematopoiesis is a complex process which gives rise to all blood lineages in the course of an organism's lifespan. Yet the underlying molecular mechanism governing this process is not fully understood. Here we report the isolation and detailed study of a newly identified zebrafish *ugly duckling (udu)* mutant allele, *udu*<sup>ql</sup>. We show that loss-of-function mutation in the *udu* gene disrupts primitive erythroid cell proliferation and differentiation in a cell-autonomous manner, resulting in red blood cell (RBC) hypoplasia. Positional cloning reveals that the *udu* gene encodes a novel factor that contains two *Paired Amphipathic α-Helix* like (PAH-L) repeats and a putative *SW13, ADA2, N-Cor and TFIIB* like (SANT-L) domain. We further show that the Udu protein is predominantly localized in the nucleus and deletion of the putative SANT-L domain abolishes its function. Our study indicates that the Udu protein is very likely to function as a transcription modulator essential for the proliferation and differentiation of erythroid lineage.
Introduction

Hematopoiesis in vertebrate occurs in two waves, primitive and definitive.\textsuperscript{1-3} In mouse, primitive or embryonic wave of hematopoiesis occurs around embryonic day 7.5 in the yolk sac blood island, and produces primitive erythrocytes and macrophages.\textsuperscript{4,5} This primitive wave hematopoiesis lasts for a transient period of a few days and is subsequently replaced by the definitive program. Murine definitive hematopoiesis is believed to originate from a distinctive region known as the aorta-gonad-mesonephros at embryonic days 7.5-8.\textsuperscript{6,7} These definitive hematopoietic precursors, presumably the definitive hematopoietic stem cells, then migrate to the fetal liver where they undergo rapid proliferation and differentiation, and finally colonize the bone marrow for adult hematopoiesis. In contrast to primitive hematopoiesis, the definitive hematopoietic program gives rise to all the mature blood cell types and remains active throughout the lifetime of the organism.

In zebrafish, hematopoiesis also comprises both primitive and definitive programs and produces mature blood cell types similar to those found in mammals.\textsuperscript{8,9} Zebrafish primitive erythropoiesis begins at the 4-somite stage as a pair of bilateral stripes in the posterior lateral mesoderm.\textsuperscript{10} These stripes extend anteriorly and posteriorly, and then converge in the midline at the 20-somite stage to form the intermediate cell mass (ICM), where erythroid precursors further develop and enter the circulation by 24-26 hours postfertilization (hpf).\textsuperscript{10,11} On the other hand, primitive myelopoiesis originates from the rostral blood island in the anterior lateral mesoderm at around the 10-somite stage and produces mainly macrophages and possibly some neutrophils.\textsuperscript{12-15} Recent studies demonstrate that zebrafish definitive hematopoiesis
initiates in the ventral wall of dorsal aorta between 26-48hpf\textsuperscript{16,17} and then establishes in the kidney, the adult hematopoietic organ, by 5 days postfertilization (dpf).\textsuperscript{11}

The zebrafish mutant, ugly duckling (udu\textsuperscript{na24}), was first isolated from the 1996 Tuebingen large scale screen as a mutant defective in morphogenesis during gastrulation and tail formation.\textsuperscript{18} In this article, we report the isolation and detailed study of a new udu allele, udu\textsuperscript{sq1}, as a mutant with defect in blood cell development. Cell cycle, cytological, and transplantation analyses showed that the primitive erythroid cells in udu\textsuperscript{sq1/-} mutant were severely impaired in proliferation and differentiation in a cell-autonomous fashion. Positional cloning revealed that the udu gene encodes a novel zebrafish protein of 2055 amino acids (aa) containing several conserved regions, including two PAH-L repeats and a putative SANT-L or Myb-like DNA binding domain (This domain is referred as SANT-L thereafter). We further showed that the Udu protein was predominantly localized in the nucleus and injection of a truncated udu mutant RNA, in which the SANT-L domain was deleted, failed to rescue the erythroid cell defect in udu\textsuperscript{sq1/-} embryos. These data indicate that the zebrafish udu gene encodes a putative transcription modifier necessary for primitive erythroid cell proliferation and differentiation.

**Materials and Methods**

**Fish maintenance and genetic screen**

Zebrafish were maintained at 27 to 28°C as described at [http://zfin.org](http://zfin.org). Mutagenesis was performed as described previously using ethylnitrosourea (ENU) (Sigma) as a mutagen\textsuperscript{19,20} and udu\textsuperscript{sq1} mutant was isolated based on defect in blood circulation in F3.
Whole-mount in situ hybridization (WISH) was performed as described at http://zfin.org/zf_info/zfbook/chapt9/9.82.html. Digoxigenin (Roche, Germany) labeled anti-sense RNA probes were synthesized from linearized plasmids by in vitro transcription.

**TUNEL assay**

Embryos were fixed in 4% paraformaldehyde (Fluka, USA) and dehydrated in PBST/methanol series (50%, 70%, 95%, and 100%), followed by incubation in 100% acetone at -20°C (10min) and rinses in PBST. The fixed embryos were then permeabilized in fresh 0.1% sodium citrate in PBST (15min) and followed by proteinase K treatment (10µg/ml, 20min). The embryos were fixed again in 4% PFA followed by PBST rinses and finally assayed using In Situ Cell Death Detection Kit, Fluorescein (Roche) according manufacture’s protocol.

**FACS, cytology, and cell cycle analysis**

24hpf *udu*<sup>sq1-/-</sup> and sibling embryos (300-400/pool) obtained from crossing adult *udu*<sup>sq1+/-</sup>/Tg(-5.0scl:EGFP) with *udu*<sup>sq1+/-</sup> fish were separated based on the morphological phenotype and disaggregated in cold 0.9X PBS with 5% FBS (Hyclone, USA). The cell suspensions were passed through a 70µm-pore size filter and spun at 1000rpm for 5min. The pellets were re-suspended in 1ml Cell Dissociation Buffer-Free Hanks (Gibco, USA) and incubated at 37°C for 20min. After addition of 1ml washing buffer {Hanks buffered saline solution containing 20% calf serum, 5mM CaCl<sub>2</sub> and DNase (50µg/ml)}, the cell suspensions were spun at 1000rpm for 5min and re-suspended in PBS. After passing through a 40µm-pore size filter, cells (10<sup>5</sup>/ml) were analysed by FACS analysis (Beckton Dickinson, USA).
For cytological analysis, about 1-2 x 10^5 sorted GFP+ cells were
cytocentrifuged at 500rpm for 5min onto glass slides and subjected to May-Grunwald
and Giemsa staining. Images were captured on an Olympus BX51 microscope
equipped with the Olympus DP70 digital camera.

For cell cycle analysis, cell suspensions (1 x 10^6 cells) were spun at 1000rpm
for 5min, re-suspended in 1ml warmed (37°C) DMEM (Gibco, USA) supplemented
with 2% Fetal Calf Serum, 10mM HEPES buffer (Gibco, USA), and 5µg/ml Hoechst
33342 (Sigma, USA), and incubated at 37°C for 1hr. The cells were then put on ice
immediately, spun down at 1000rpm for 5min at 4°C, re-suspended in ice-cold HBSS
(Hanks Balanced Salt Solution from Gibco) with 2% Fetal Calf Serum (Hyclone,
USA) and 10mM HEPES buffer, and finally subjected to cell cycle analysis by FACS.

**Positional cloning**

Single embryo genomic DNA was prepared by incubating a single embryo in 50µl 1X
TE buffer (pH8.0) with 0.5mg/ml proteinase K (Finnzymes, Finland) at 55°C
overnight. After incubation at 98°C for 10min, 1µl of each sample was used as
template for each PCR reaction.

Initial mapping was done by the bulk segregation analysis as described
(http://134.174.23.167/zonrhmapper/positionCloningGuidenew/index.htm) and the
udu mutation was found to link with two SSLP markers, z10036 and z1215, on
linkage group 16. Two closer SSLP markers z17246 (North marker) and SSLP-S16
(5’-atagttgccagctgagggtc-3’/5’-gctgagattaggcaactgtc-3’ located at BAC clone
zk30O16) (South marker) were then used to analyze 2932 single udu^sq1/-/ embryos.
The closest North and South SSLP markers are SSLP-N6 (5’-gttactataagtgtatcgtatggggg-
3’/5’-agttagaatgatctcagtg-3’) and SSLP-S18 (5’-acatgtgtatatgacttggtgc-3’/5’-
atgttataataatcttcactcc-3’), which map the udu mutation to a region covered by three overlapping BAC clones, zC196P10, zK7E18 and zC113G11. The mutations in the udu gene were confirmed by sequencing the RT-PCR products of the total RNA extracted from the udu<sup>uu24/-</sup> and udu<sup>sq1/-</sup> embryos.

5’ and 3’ RACE

5’ and 3’ RACE were performed using the BD SMART<sup>TM</sup> RACE cDNA Amplification Kit (Clontech, USA) according to the protocol provided by the manufacturer.

<enua short="udu RNA rescue experiments"

The wild type (udu-wt), T2976A mutant (udu-T2976A) (nucleotide T2976 changed to A), and SANT-L domain deficient mutant (udu-ΔSANT-L) (deletion from aa 1963 to 2031) cDNAs were subcloned into pcDNA3.1 vector (Invitrogen, USA). The capped RNA was synthesized and polyadenylated using Ambion’s mMessage mMachine High yield capped RNA Transcription kit and Ambion’s Poly(A) Tailing Kit respectively according to the manufacturer’s protocols (Ambion, USA). For rescue experiment, each one-cell stage embryo derived from crossing adult udu<sup>sq1/+</sup> fish was injected with ~0.6ng in vitro synthesized RNA. The primer sequences for genotyping analysis are 5’-aaacacgctacccacagttcc-3’ and 5’-ttgtctgatgtctgttgctgc-3’.

Morpholino (MO) knockdown

The udu MOs (Gene Tools, USA), MO-udu-1 (5’- TAACACTACACTCACCACCC CTTTT-3’) and MO-udu-2 (5’-AAAAGGCTTGCTGACCGTCGTTGTC-3’), are designed to specifically block RNA splicing of the udu gene, whereas the p53 MO
(5’-GCGCCATTGCTTTGCAAGAATTG-3’)21 (Gene Tools, USA) is used to target the translational initiation site of the p53 gene. The Standard Control MO provided by Gene Tools was used as control. ~0.5-1pmol MO was injected into each one-cell stage embryo.

**Cell transplantation analysis**

Donor embryos derived from crossing adult *udu*<sup>aq14+</sup>-/- fish were injected at the one-cell stage with rhodamine-dextran (Invitrogen, USA). At around 3hpf, the injected donor embryos were dechorionated with forceps in an agarose- (2% agarose in 0.3X Danieau buffer) coated Petri-dish covered with 1X Danieau buffer, and 15-30 donor cells from each embryo were transplanted into the dechorionated host embryos of the same stage. The manipulated donor embryos were saved for genotyping analysis. Contribution of the rhodamine-dextran-labeled donor cells to the circulating blood cells in the host embryos was scored at around 30hpf under fluorescent microscope.

**Immunohistochemistry staining**

COS-7 cells were grown on 22 x 22mm cover slips in 35-mm wells and transfected with pcDNA3.1/*udu*-wt using SuperFect Transfection Reagent (QIAGen, Germany) as according to the manufacturer’s protocol. 24hr after transfection, cells were washed with PBS-CM (PBS with 1mM CaCl<sub>2</sub> and 1mM MgCl<sub>2</sub>) and fixed with 3% paraformaldehyde (Fluka, USA) in PBS-CM at 4°C for 30min. The fixed cells were washed twice with PBS-CM followed by 0.26% NH<sub>4</sub>Cl in PBS-CM (2 times) and finally PBS-CM. The cells were permeabilized with 0.1% saponin (Sigma, USA) in PBS-CM at room temperature for 15min. The permeabilized cells were then incubated with rabbit anti-Udu antibody (against aa 1818-1941) in FDB (PBS-CM with 5%
normal goat serum, 2% fetal bovine serum, and 2% BSA) at 4°C overnight. After rinsing three times with 0.1% saponin in PBS-CM, the cells were incubated with Alexa Fluor 488 donkey anti-rabbit IgG antibody (Molecular Probes, USA) and washed with 0.1% saponin/PBS-CM (5 times). The coverslips were mounted with Vectashield mounting medium for fluorescence with propidium iodine (PI) (Vector, USA) and imaged on a Zeiss LSM 510 confocal system.

**Affymetrix array analysis**

Total RNA was extracted from 1dpf *udu^sq1/-* and sibling embryos using TRIzol (Gibco-BRL, USA), treated with DNase I and purified using RNeasy Mini kit (QIAgen, Germany). Hybridization was performed according to the manufacturer’s instruction (Affymetrix, USA). Data were analyzed using Microarray Suite 5.0 software (Affymetrix, USA).

**Semi-quantitative and real-time RT-PCR**

Total RNA of embryos, extracted as previously described, was reversely transcribed using SuperScript™ III RNase H Reverse Transcriptase (Invitrogen, USA). The amount of reversely transcribed cDNAs was normalized with the Real-time LightCycler (Roche, Germany) using elongation factor 1 alpha (*elf1α*) as a reference. Semi-quantitative RT-PCR was performed for either 25 or 30 cycles at 94°C for 30sec; 58°C for 10sec; 72°C for 30sec. For real-time RT-PCR, total RNA of GFP+ cells sorted from 24hpf *udu^sq1/-* and wild type sibling embryos was extracted using RNeasy Micro kit (QIAgen, Germany). Reverse transcription was performed using Superscript III Kit (Invitrogen), about 10% of the RT products were amplified in the real-time PCR reaction using the LightCycler System (Roche) according to the manufacturer's
instruction. elf1α was used as an internal reference to normalize the PCR reaction. Primer sequences are listed as follows: p53 (5'-gcttttagatttagtacaaccattg-3'/5'-gcaaatgcgtgtaacagtaataag-3'), caspase8 (5'-ggattgatctggaagcctgg-3'/5'-tagcgtggttctg gcatcg-3'), fos (5'-ccaaacagagaaaagagcag-3'/5'-cagtgctcggtagttcagc-3'), gadd45αl (5'-attgaagagctgcgcgcgggtg-3'/5'-ttcggctgctctcagctc-3'), tdag51 (5'-aaacagcgacgatgc cttaca3'/5'-ccagcaacagcgggc-3'), mdm2 (5'-ctcgcagtgagggcagtgaag -3'/5'-tctaggca gttggagaggggaagc-3'), elf1α (5'-cttctcaggtgcgtgc-3'/5'-ccgctagcattaccctcc-3').

Results

Primitive Erythrocyte Hypoplasia in udu^{sq1/-} Mutant

To study the genetic programs governing vertebrate hematopoiesis, we carried out a forward genetic screen in search of zebrafish mutants affecting hematopoiesis. Among the hematopoiesis-defective mutants recovered, wz260 exhibited morphological phenotype similar to that of zebrafish mutant ugly duckling (udu^{ua24}) isolated from the 1996 Tuebingen screen.18 Complementation analysis confirmed that wz260 was a new allele of udu^{ua24}, thus this mutant was re-named as udu^{sq1}. As both udu alleles displayed similar phenotype, we used udu^{sq1} for detailed analysis in this study. The zebrafish udu^{sq1/-} embryos began to show notable morphological abnormality at around 24hpf such as short body axis, bent-down tail, irregular somites, small head and eyes and lack of blood circulation, and could not survive beyond 7-10dpf (data not shown).18 In this report, we focused our study on the primitive erythroid defect of udu^{sq1}. We first performed o-dianisidine staining to examine primitive erythropoiesis and found that hemoglobin level was severely reduced in the 2dpf udu^{sq1/-} mutant (Figure 1M-N). To explore the erythroid defects in depth, we examined the expression of two critical early hematopoietic markers lmo2 and scl.22,23
WISH showed that neither lmo2 (data not shown) nor scl exhibited apparent expression differences between udu^+/− and wild type embryos before 10-somite stage (Figure 1A-B), indicating that the udu gene is dispensable for the initiation of primitive erythropoiesis. However, by 22hpf, we found that the expression of gata1 (Figure 1C-D) was reduced and band3 transcript, an erythrocyte-specific membrane protein critical for erythrocyte maturation, was almost undetectable in udu^+/− embryos (Figure 1E-F). Intriguingly, βe3-globin expression remained relatively normal or only slightly reduced in 22hpf udu^+/− embryos (Figure 1G-H). These observations suggest that aberrance in erythrocyte differentiation is very likely to be the main reason that causes the dramatic decrease of band3 expression in udu^+/− mutant and eventually leads to RBC hypoplasia at later stage of development as indicated by reduced βe3-globin expression from 30hpf onwards (Figure 1I-L).

**Primitive Erythroid Cells in udu^+/− Mutant are Impaired in Proliferation and Differentiation**

TUNEL staining of the udu^+/− embryos revealed that, while extensive cell death was occurred in the central nervous system (CNS), no increased apoptotic cells were observed in the ICM (Figure S1). This suggests that aberrance in cell cycle rather than apoptosis contributes to the reduction of erythroid cells in udu^+/− mutant. To clarify this issue, we out-crossed the udu^+/− fish with the Tg(-5.0scl:EGFP) [referred to as Tg(-5.0scl:EGFP) in the original paper] transgenic line, in which the expression of the enhanced green fluorescent protein (EGFP) reporter gene is under the control of scl promoter. The resulting udu^+/−/Tg(-5.0scl:EGFP) was then mated with udu^+/− fish. Finally, the GFP positive (GFP+) cells of the 24hpf udu^+/−/Tg(-5.0scl:EGFP) and sibling embryos were collected by the FACS (Figure 2A and 2D)
and subjected to cell cycle analysis. DNA content examination using Hoechst 33342 staining revealed that, while wild type cells displayed 45.06%, 43.93% and 11.01% of cells in G1/G0, S and G2/M phases, respectively (Figure 2B), the udu^aq1-/- mutant cells exhibited abnormal accumulation in G2/M (43.16%) and reduction in G1/G0 (32.13%) and S (24.71%) phases (Figure 2E), indicating that the udu^aq1-/- mutant cells arrest at G2/M phase. Cytological analysis showed that the wild type GFP+ cells consisted of mainly erythroid cells in various stages of differentiation (cells with a higher degree of differentiation exhibited a round central nucleus, chromatin condensation and smaller in size) and some myeloid cells (characterized by their irregular nucleus) (Figure 2C). In contrast, the udu^aq1-/- mutant GFP+ cells, which were mainly erythroid cells, were much larger in cell size and lacked chromatin condensation (Figure 2F). Although these features resembled those erythroid progenitors found in the 16hpf wild type embryos (data not shown), the mutant GFP+ cells also displayed nuclear-cytoplasm asynchrony, which is a characteristic of abnormal megaloblastic erythroid cells found in human megaloblastic anemia patients. Taken together, these data demonstrate that the udu^aq1-/- primitive erythroid cells are defective in cell proliferation and differentiation abilities.

Identification of the udu^aq1 Gene

Positional cloning revealed that the predicted Ensemble gene ENSDARG00000005867 was the candidate for the udu^aq1 mutant gene (Figure 3A). 5’ and 3’ RACE revealed that the full-length cDNA sequence of ENSDARG00000005867 were 6787 base pairs encoding a novel protein of 2055aa (Figure 3B; Figure S2). Through sequencing, the mutations in udutu^aq24 and udu^aq1 mutants were discovered at exon 12 (T1461 to A) and exon 21 (T2976 to A)
respectively, resulting in the creation of a premature stop codon in each case (Fig. 3B). To confirm that the \( udu^{aq1/-} \) mutant phenotype is indeed caused by mutation in the ENSDARG00000005867 gene, we first performed a rescue experiment with in vitro synthesized ENSDARG00000005867 RNA. Primitive erythropoiesis was restored in 59 out of the 64 (92%) \( udu^{aq1/-} \) embryos injected with wild type RNA (\( udu\text{-}wt \)) (Figure 3E; Table S1). In contrast, injecting the \( udu^{aq1} \) mutant RNA (\( udu\text{-}T2976A \) bearing a single point nonsense mutation (nucleotide T2976 changes to A) failed to do so (Figure 3F; Table S1). As expected, embryos injected with two \( udu \) MOs, MO-\( udu \)-1 (data not shown) and MO-\( udu \)-2, displayed a phenotype mimicking the \( udu^{aq1/-} \) mutant (Figure 3G-3J and Table S2). These data demonstrate that the loss-of-function mutation in the ENSDARG00000005867 gene indeed causes the \( udu^{aq1/-} \) mutant phenotype. We herein refer ENSDARG00000005867 gene as the \( udu \) gene.

**Cell autonomous Erythroid Defect in \( udu^{aq1/-} \) Mutant**

WISH showed that the \( udu \) transcript was detected as early as one-cell stage in wild type zebrafish embryos (Figure 4A). The maternal \( udu \) mRNA retained robust expression during blastula and diminished at the onset of gastrulation (~6hpf) (Figure 4B-D). Around the time when segmentation began, the \( udu \) transcript, presumably the zygotic mRNA, began to re-appear in a ubiquitous manner, and was subsequently enriched in the CNS as well as the ICM from the 18-somite stage onwards (Figure 4E-H). At 24hpf \( udu \) mRNA was also detected in the anterior parts of kidney ducts (Figure 4I).

From the temporal and spatial expression of the \( udu \) gene, we believe that Udu possibly plays a cell-autonomous role during the primitive red blood cell development. To investigate this speculation, we performed a transplantation experiment. Around
15-30 donor cells from 3hpf $udu^{+/+}$ or sibling embryos, pre-injected with rhodamine-dextran, were transplanted into wild type host embryos of the same stage. Contribution of the donor cells to the circulating blood cells in the host embryos was scored at around 30hpf by counting the number of the rhodamine-dextran labeled cells in the circulation under a fluorescent microscope. As summarized in Table 1, when sibling donor cells were transplanted, about 42% (50/119) of the recipients had the rhodamine-dextran labeled donor cells in circulation. Of which 20% (10/50) contained 10 to 30 circulating donor cells and 22% (11/50) had more than 30 circulating donor cells. In contrast, when the $udu^{+/+}$ mutant cells were transplanted, only 26.7% (16/60) of the host embryos had the donor cells contributing to the blood circulation. More importantly, none of these embryos contained more than 10 circulating rhodamine-dextran labeled cells. The transplantation result strongly indicates that the $udu$ gene acts cell-autonomously to affect primitive red blood cell development. However, non-cell autonomous effect cannot be excluded.

**udu Gene Encodes a Putative Transcriptional Modulator**

Blast searches in database revealed that the Udu protein had the highest homology to the human and mouse GON4L in both protein sequence and gene structure (Figure S2). Sequence alignment reveals that there are several highly conserved regions (Figure 5A and Figure S2). The first three conserved regions (CR-1, CR-2 and CR-3) share no obvious similarity to any of the known domains. The fourth and fifth (from aa 1538 to 1740), which are predicted to consist of four $\alpha$-helixes each, are similar to the PAH repeats found in SIN3 proteins, and are thus designated as PAH-like (PAH-L) 1 and 2 domain (Figure 5A). Finally, the solution structure (IUG2_A) of the last conserved region (aa 1947 to aa 2039) has been solved in the
mouse udu homologue GON4L (www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=1ug2). The structure analysis reveals that this last conserved region resembles the SANT domain found in several chromatin-remodeling molecules (Figure 5A-B), suggesting that the Udu protein may be involved in transcription regulation.

To provide additional evidence to support this argument, we carried out rescue experiment with the truncated udu-\Delta SANT-L RNA, in which three \(\alpha\)-helixes (from aa 1963 to 2031) of the SANT-L domain of the Udu protein were deleted. As shown in Figure 5, the red blood cell development was restored in the mutant embryos injected with the udu-wt RNA (Figure 5E). However, the truncated udu-\Delta SANT-L RNA failed to do so (Figure 5F; Table S1), demonstrating that the SANT-L domain is critical for the function of the Udu protein. In addition, udu cDNA (pcDNA3.1/udu-wt) was transfected into COS7 cells to examine the sub-cellular localization of the Udu protein. Immunohistochemistry analysis showed that the Udu protein was predominantly localized in the nucleus (Figure 5G-I). These data strongly suggest that the Udu protein may function as a transcription regulator important for erythroid cell development.

**Elevation of p53 Activity Partially Contributes to the Erythroid Defect in udu^{eq1-/-} Mutant**

Microarray analysis was performed to identify potential target genes perturbed in udu^{eq1-/-} mutant. Results from two sets of independent experiments showed that expression of 87 and 57 genes were down- and up-regulated respectively, with 2-fold and above differences in the udu^{eq1-/-} mutant embryos (Table S3 and S4). As expected, most of the down-regulated genes are related to either hematopoiesis or neurogenesis.
Notably, several up-regulated genes including p53, fos, mdm2, gadd45α, caspase 8, and tdag51, are known to be involved in cell cycle control and apoptosis\(^{32-35}\) (Table S4; Figure S3). Considering the G2/M cell cycle arrest of the \(udu^{sq1/-}\) erythroid cells, we focused on p53 and its downstream target gadd45α. WISH showed that in the 24hpf \(udu^{sq1/-}\) embryos, the RNA levels of p53 (Figure 6A-B) and gadd45α (Figure 6I-J) were significantly elevated in the CNS and ICM regions. To confirm that p53 transcript was indeed up-regulated in \(udu^{sq1/-}\) mutant erythroid cells, real time RT-PCR was performed using GFP+ cells sorted from 24hpf wild type or \(udu^{sq1/-}\) mutant embryos. As shown in Table 2, p53 transcript was significantly increased in \(udu^{sq1/-}\) erythroid cells. These observations suggest that the up-regulation of p53 activity may contribute to the erythroid defect in \(udu^{sq1/-}\) embryos. This argument was further supported by the findings that knockdown of the p53 protein by p53 MO\(^{21}\) rescued the erythroid phenotype in the \(udu^{sq1/-}\) embryos (Figure 6C-H; Table S1). As anticipated, elevated gadd45α expression was no longer detected in these p53 morphant mutants (Figure 6I-K). Taken together, these data suggest that up-regulation of p53 activity contributes, at least partially, to the erythroid cell hypoplasia in \(udu^{sq1/-}\) mutant.

**Discussion**

In this article, we have shown that the \(udu\) gene, which encodes a novel nuclear protein containing two PAH-L repeats and a SANT-L domain, plays a critical role in regulating primitive erythroid lineage cell cycle progression and differentiation. This was indicated by the lack of erythroid specific marker band3 expression and the accumulation of G2/M phase in erythroid cells in \(udu^{sq1/-}\) embryos. We noticed that, although expressed ubiquitously throughout early development, the \(udu\) gene appears to be dispensable for early embryogenesis as well as initiation of primitive
hematopoiesis. Considering that robust maternal *udu* RNA is detected in fertilized embryos, the lack of early phenotype in *udu*<sup>eq1/-</sup> mutant embryos is possibly due to the functional compensation of maternal *udu* RNA. The primordial germ cells replacement approach described by Ciruna et al. can be used to address this issue in the future.\textsuperscript{36}

Sequence analysis reveals that the Udu protein contains several conserved regions, including two PAH-L repeats and a SANT-L or Myb-like DNA binding domain. The PAH domain, originally defined in the yeast SIN3 protein,\textsuperscript{29} is distantly related to the helix-loop-helix motif and has been shown to mediate protein-protein interaction.\textsuperscript{30} The yeast SIN3 protein and its related mammalian homologues, Sin3A and Sin3B, interact through the PAH domain, with numerous sequence-specific transcription factors and recruit histone deacetylases to suppress downstream target gene transcription.\textsuperscript{37} This implies that Udu protein may form complexes with interaction partners via PAH domain. The SANT and Myb-DNA binding (Myb-DB) domains have a similar overall structure but confer distinct functions.\textsuperscript{38} The Myb-DB domain usually contains two to three tandem repeats and recognizes specific DNA sequence, whereas the SANT domain consists of one to two repeats and plays an important role in chromatin remodeling.\textsuperscript{38} Considering the fact that the Udu protein contains only one repeat of this domain that lacks positive electrostatic surface patch as predicted by structural modeling analysis, we believe that this conserved region resembles the SANT domain. As the Udu protein has both PAH-L and SANT-L domains, it is possible that Udu may function as a chromatin-remodeling molecule involved in transcription regulation of many target genes. Further experiments need to be performed to support our hypothesis.
p53 is well recognized as one of the most important tumor suppressors in preventing cancer by modulating downstream target gene expression in response to various cell stresses, such as DNA damage and oncogene activation, resulting in cell cycle arrest or death of abnormal cells. Over-expression of p53 during embryogenesis causes various developmental defects in mouse, fly and worm. More recently, the developmental defects caused by loss-of-function mutations in zebrafish DNA polymerase delta1 or def gene have been shown to be associated with increased p53 activity. These observations imply that p53 plays an important role in maintaining normal cell growth and differentiation during animal development. Our findings show that MO knockdown of p53 protein expression partially rescues the primitive erythroid cell defect in embryos, which provides another example to demonstrate that p53 activity must be tightly controlled during embryogenesis in order to maintain normal cell growth and differentiation. At this moment it is not clear how loss-of-function mutation in the gene leads to up-regulation of p53 activity. One possible explanation is that Udu protein may function as a general transcriptional modulator and its loss-of-function cause a global change in gene expression. This remarkable change triggers stress response in mutant cells resulting in up-regulation of p53 expression. Alternatively, Udu protein may act as a suppressor that directly regulates transcription or protein activity of p53. Identifying the interaction partners or the direct target genes of Udu protein will be helpful to understand the relationship between udu and p53.

It has been recently shown that a new gene next to the human GON4L, the human counterpart of udu, is generated as a result of segmental duplication on human chromosome 1q22. Surprisingly, this segmental duplication does not exist in rat, mouse, chicken and zebrafish, suggesting that this duplication may be associated with
more recent evolutionary events specific for anthropoid primates.\textsuperscript{28} It has been known that human chromosome 1q22, where GON4L gene is located, is amplified in several human cancer types, including ovarian and breast cancer,\textsuperscript{45} sarcomas,\textsuperscript{46} and hepatocellular carcinoma.\textsuperscript{47,48} A more detailed study has further confirmed that three loci covered by YAC955E11, YAC876B11 and YAC945D5 within 1q21-q22 have the highest amplified copy number in five out of the ten hepatocellular carcinoma cases examined.\textsuperscript{49} Intriguingly, the genomic fragment covered by the YAC876B11 contains a marker D1S2140 that is 30kb apart from the GON4L gene, suggesting that perhaps amplification of the GON4L gene is associated with tumor development and this warrants future study.

**Acknowledgments**

Liu, Y.M., Du, L.S., and Osato, M. performed research and analyzed data. Teo, E.H., Qian, F., Jin, H., Zhen, F.H., Xu, J., Guo, L., Huang, H.H., and Chen, J. performed experiments. Geisler R. and Jiang, Y.J. provided critical reagent. Peng, J.R. analyzed data. Liu, Y.M. and Wen, Z.L. designed the research, analyzed data and wrote paper. We thank Dr. Haiwei Song for protein domain analysis; Ms. Meipei She and Dr. Chengjin Zhang for technical advises; Drs. Sudipto Roy, Vladimir Korzh and Mei Huang for stimulating discussions; and Dr. Dingxiang Liu for comments on the manuscript. This work was funded by the Agency for Science, Technology, and Research, Singapore.
References


(43) Plaster N, Sonntag C, Busse CE, Hammerschmidt M. p53 deficiency rescues apoptosis and differentiation of multiple cell types in zebrafish flathead


FIGURE LEGENDS

Figure 1. Primitive Erythroid Hypoplasia in udu<sup>q1/-</sup> Mutant. (A-L) WISH of scl (A,B), gata1 (C,D), band3 (E,F), and βe3-globin (G-L) expression. (M,N) O-dianisidine staining of 2dpf wild type and udu<sup>q1/-</sup> embryos. All embryos are in lateral views with anterior to the left.

Figure 2. Primitive Erythroid Cells in udu<sup>q1/-</sup> Mutant Are Defective in Cell Proliferation and Differentiation. (A,D) FACS profile of the cell suspensions from 24hpf wild type (A) or udu<sup>q1/-</sup> mutant (D) embryos. Y-axis indicates the intensity of the GFP expression and x-axis represents cell size. (B,E) The GFP-positive hematopoietic cells in A (total of 5,394 cells) (B) and D (total of 6,874 cells) (E) are subjected to DNA content analysis by Hochest33342 staining. Y-axis indicates the cell number whereas x-axis represents the DNA content. The percentages of each phase in cell cycle are given. (C,F) May-Grunwald and Giemsa staining analysis (magnification x1000) of sorted GFP+ cells in A (C) and D (F). Black and red arrows in C indicate erythroid and myeloid cell respectively.

Figure 3. Identification of the udu Gene. (A) The udu gene is mapped to LG16 within the region covered by three BACs, zC196P10, zK7E18 and zC113G11. The number in red represents the number of the recombinants over 5,864 meiosis events for each SSLP marker. Sequence analysis confirms that the udu mutation is situated in the BAC zC196P10. (B) The udu gene consists of 31 exons (solid box) and encodes a protein of 2055 aa. Both udu<sup>ua24</sup> and udu<sup>q1</sup> harbor a nonsense mutation in exon 12 and 21, respectively. (C-F) WISH of band3 expression of 22hpf wild type embryo (C), udu<sup>q1/-</sup> embryo (D), udu<sup>q1/-</sup> embryo injected with in vitro synthesized wild type udu
(udu-wt) RNA (E), and udu<sup>q1/-</sup> embryo injected with in vitro synthesized T2976A mutant (udu-T2976A) RNA (F). (G-J) <i>band3</i> (G,H) and <i>b3-globin</i> (I,J) expression by WISH in the 22hpf control morphants (G,I), and MO-udu-2 morphants (H,J). In C-J, embryos are in lateral views with anterior to the left.

**Figure 4. Temporal and Spatial Expression of the udu Gene during Early Zebrafish Development.** (A-I) Lateral views of WISH of <i>udu</i> expression in one-cell (A), 2.5hpf (B), 4hpf (C), 6hpf (D), 10hpf (E), 12hpf (F), 14hpf (G), 18hpf (H), and 24hpf (I) embryos. White and black arrows indicate the ICM and the anterior region of kidney duct respectively. Embryos in A-D are orientated with animal pole on top whereas embryos in E-I are orientated with anterior to the left.

**Figure 5. udu Gene Encodes a Putative Transcriptional Modulator.** (A) Udu protein contains six conserved regions: CR-1, CR-2, CR-3, PAH-L1, PAH-L2 and SANT-L. (B) Protein sequence alignment of SANT-L domain between zebrafish Udu (top, fish-Udu-1947-2039aa) and mouse Udu homologue GON4L (1UG2_A). (C-F) Lateral views of o-dianisidine staining of 2dpf wild type (C), udu<sup>q1/-</sup> embryo (D), the udu<sup>q1/-</sup> embryo injected with udu-wt RNA (E), and the udu<sup>q1/-</sup> embryo injected with udu-ΔSANT-L RNA (F). The embryos in C-F are in lateral views with anterior to the left. Arrows in C and E indicate o-dianisidine stained red blood cells. (G) Propidium Iodide (PI) staining of the pcDNA3.1/udu-wt transfected COS7 cells. (H) Immunohistochemistry staining of the pcDNA3.1/udu-wt transfected COS7 cells with the anti-Udu polyclonal antibody. (I) Superimposed image of G and H. Scale bars in G-I represent 10µm.
Figure 6. Up-regulation of p53 Activity Partially Contributes to the Erythroid Phenotype in udu sq1-/- Mutant. (A-B) WISH shows that p53 expression is elevated in 24hpf udu sq1-/- embryo (B) compared to wild type (A). (C-E) WISH of band3 in 22hpf wild type embryo (C) and udu sq1-/- mutant embryos injected with control MO (D) or p53 MO (E). (F-H) O-dianisidine staining of 2dpf wild type embryo (F) and udu sq1-/- mutant embryos injected with control MO (G) or p53 MO (H). (I-K) WISH of gadd45α in 24hpf wild type embryo (I) and udu sq1-/- mutant embryos injected with control MO (J) or p53 MO (K). All embryos are in lateral views with anterior to the left. Arrows in F and H indicate o-dianisidine stained red blood cells.
Figure 1

![Image of Figure 1 with panels A to N showing different stages and conditions for various genes and proteins](image-url)

Figure 2

![Image of Figure 2 with panels A to F showing cell cycle analysis and images](image-url)
Figure 5

Figure 6
Table 1. Summary of Cell Transplantation Analysis

<table>
<thead>
<tr>
<th>Donor genotype</th>
<th>Number of hosts</th>
<th>The number (percentage) of hosts with donor-derived tissue</th>
<th>Blood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Muscle</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>&gt;30 blood cells</td>
</tr>
<tr>
<td>Sibling</td>
<td>119</td>
<td>86 (72.3%)</td>
<td>50 (42%)</td>
</tr>
<tr>
<td>udusq1 -/-</td>
<td>60</td>
<td>46 (76.7%)</td>
<td>16 (26.7)</td>
</tr>
</tbody>
</table>

Table 2. Difference of p53 transcript level in erythroid cells between wild type and udusq1 -/- mutant by real-time RT-PCR

<table>
<thead>
<tr>
<th>Crossing Point</th>
<th>Wt sibling</th>
<th>udusq1 -/-</th>
<th>Cycle number difference*</th>
<th>Relative fold difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test 1</td>
<td>38.74</td>
<td>35.93</td>
<td>2.81</td>
<td>7.012845771</td>
</tr>
<tr>
<td>Test 2</td>
<td>39.15</td>
<td>36.01</td>
<td>3.14</td>
<td>8.815240927</td>
</tr>
<tr>
<td>Average</td>
<td>2.975</td>
<td></td>
<td></td>
<td>7.914043349</td>
</tr>
</tbody>
</table>

* Elongation factor 1 alpha (elf1α) was used as an internal reference to normalize the above reaction.
The zebrafish udu gene encodes a novel nuclear factor and is essential for primitive erythroid cell development

Yanmei Liu, Linsen Du, Motomi Osato, Eng Hui Teo, Feng Qian, Hao Jin, Fenghua Zhen, Jin Xu, Lin Guo, Honghui Huang, Jun Chen, Robert Geisler, Yun-Jin Jiang, Jinrong Peng and Zilong Wen